A Literature Review Paper in partial fulfillment of the Preliminary Examination

Facilitating neuroimaging marker discovery and validation:

The predictive mapping approach

Choong-Wan Woo[1,2]

[1]Department of Psychology and Neuroscience, University of Colorado, Boulder

[2]Institute of Cognitive Science, University of Colorado, Boulder

Abstract

Neuroimaging can contribute to our understanding of the human mind and its disorders *if and only if* we have neuroimaging markers that are diagnostic of particular mental processes. The predictive mapping approach is an emerging paradigm designed to produce neuroimaging markers useful for understanding brain representations and promoting translational applications. It is based on specific uses of machine learning techniques combined with experimental designs optimized for prediction, which provide well-defined models of brain-outcome relationships that can be prospectively tested in new individuals, studies, and translational applications.

For the past twenty years, neuroimaging techniques, including magnetic resonance imaging (MRI), positron emission tomography (PET), and electroencephalogram (EEG), have transformed how we study psychology and medicine. MRI in particular has become increasingly widely used, in large part due to its accessibility to researchers from many fields and the promise that it will eventually revolutionize our understanding of the human mind and brain-related illnesses. Neuroimaging has provided a principal tool for mapping mind to brain; neuroimaging data can constrain psychological theories, resolve some theoretical debates, and be used to develop new hypotheses about human cognition and emotions by providing a grounding in neurophysiology [1]. In medicine, neuroimaging provides promising measures that can serve as biomarkers for brain-related disorders, such as psychiatric and neurologic disorders, chronic pain, and fatigue [2-5]. Neuroimaging can also connect psychology to biology and medicine, which can help researchers understand how the mind and the body interact and thereby treat medical conditions more effectively (for example, understanding the placebo effect [6]).

Despite these promises, the development of neuroimaging has not followed the quick and easy path to success that was initially envisioned by some. One important reason is that too little effort has gone into developing neuroimaging markers that are sensitive and specific to particular mental processes or health-related outcomes and that can be prospectively applied to new data. The dominant paradigm in neuroimaging has focused on brain "maps", not markers. Brain maps identify anatomical regions associated with particular mental processes. This paradigm does not adequately address the many-to-many relationships between brain regions and mental processes: One brain region can be involved in multiple processes, and one process can be distributed across many regions. Thus, we cannot make inferences about which mental process is engaged based on brain mapping. Markers, by contrast, are multivariate patterns of brain activity optimized to be predictive of, and sensitive and specific to, a particular type of mental process.

Without markers with rigorously validated sensitivity and specificity, the inferences we can make about brain representations are fundamentally limited [7].

The predictive mapping approach is an emerging paradigm designed to establish neuroimaging markers that are useful for understanding brain representations and promoting translational applications. It is based on specific uses of multivariate pattern analysis (MVPA) methods [8, 9] combined with experimental designs optimized for prediction. Predictive mapping provides markers that can be prospectively applied to new individuals, studies, and clinical tests. In this paper, we describe the predictive mapping approach by contrasting it with traditional and information-based mapping, and suggest that the new approach has several advantages for both non-clinical (basic) and clinical neuroimaging research. In addition, we present findings from a literature survey of recent functional MRI (fMRI) studies that use MVPA (N = 131 studies), highlighting the fact that researchers still focus on the localization of effects with MVPA, rather than using MVPA to develop neuroimaging markers. Finally, we conclude with implications and recommendations for basic and clinical researchers about how to use the predictive mapping approach to foster their own studies and their transitional applications.

We don't have neuroimaging markers?

It might seem that neuroimaging markers for mental processes already exist, but in fact, we have been using neuroimaging findings as markers without properly assessing their sensitivity and specificity. For example, in neuroimaging studies of pain, fMRI activation within pain-processing regions, often the anterior cingulate cortex (ACC) or insula, are commonly taken as brain representations of somatic pain. In studies of emotion, amygdala activity has often been used as a brain marker for negative emotions. However, these brain regions are large anatomical brain structures comprised of heterogeneous neuronal populations that encode various physical and mental events [10-12], and therefore, averaged fMRI activity within

these regions cannot be used as brain markers for pain or negative emotions because of their low specificity [13-15].

In order to be markers for mental events or processes, brain measures should show high sensitivity and specificity to the mental events of interest. Sensitivity concerns whether a test—in this case, a brain marker—show positive results when a target psychological or behavioral process is engaged, while specificity describes whether the test shows positive results exclusively when the target process is engaged. We can express these concepts using probabilistic terms in the context of measuring mental states from the brain: Sensitivity can be expressed as the probability of positive brain findings (B) given the target mental events (M), $P(B|M)$, and specificity as the probability of negative brain findings given no mental events, $P(\sim B|\sim M)$. Sensitivity and specificity can tell us the diagnostic performance of brain measures to make inferences or predictions about mental processes or outcomes-of-interest.

Traditional brain mapping

Traditional brain mapping—often called mass-univariate analysis or statistical parametric mapping—has been extremely useful in the development of the field. However, it does not extend naturally to the idea of identifying and utilizing brain markers with established sensitivity and specificity. The main goal of the traditional approaches is to map different mental functions onto local brain regions to understand the brain-function mapping. As Figure 1A demonstrates, in this framework, tasks or conditions are used as independent variables, and each voxel's fMRI signal becomes a dependent variable. The most important question answered by the traditional approach is whether there is an effect in each voxel or region.

This traditional approach has, at best, low sensitivity to the effects of task conditions because it assumes independence among voxels and regions. However, in truth, brain regions are intricately interconnected into networks and circuits, and

psychological and behavioral processes, and related outcomes, emerge as a function of these integrated circuit dynamics. Thus, the effects of task conditions—and relationships between brain activity and behavioral/psychological outcomes—are likely to be distributed across brain regions and voxels. Analyses that consider only using information in a single voxel, as the mass-univariate approach does, are unlikely to capture the full effects of tasks. In addition, the univariate approach involves a large number of statistical tests and requires a correction for multiple comparisons [16, 17]. This correction focuses on controlling false positive findings, and in turn increases false negatives, which results in extremely low sensitivity [16]. Due to the low sensitivity, a large number of voxels activated in relation to a task or outcome could be missed, providing a poor assessment of the pattern across the brain. This, in turn, undermines efforts to establish replicability across studies and obtain a precise estimate of effect size, as each study shows activity in only a very small portion of the truly activated voxels and reports effects that survived stringent statistical tests [18, 19]. Furthermore, as illustrated in Figure 2, traditional brain mapping has a limited ability to detect the specific and diagnostic relationships between mental functions and brain voxels because it does not take other voxels into account when assessing one voxel. This limited ability undermines the specificity of the resulting brain maps.

Developing neuroimaging markers: The predictive mapping approach

The predictive mapping approach can resolve the issues described above and provide neuroimaging markers with quantitatively characterized diagnostic performance. The predictive mapping approach aims to develop multivariate, system-level predictive models (or decoding models) that are sensitive and specific to particular outcomes of interest (e.g., mental events, or health, functional, behavioral outcomes [20, 21]). As Figure 1B shows, one of the main features that distinguishes predictive mapping from traditional approaches is that the assignment

of independent and dependent variables is reversed: Behavior is the outcome in the predictive mapping approach, whereas it is the predictor in traditional approaches.

The predictive mapping approach helps to solve the low sensitivity and specificity problem of traditional mapping in several ways. First, predictive mapping provides higher power than traditional mapping because it utilizes information at multiple spatial scales, including large-scale information distributed across multiple systems and also meso-scale information below the resolution of the imaging itself (fMRI "hyperacuity") [9]. Assessing multivariate patterns rather than individual voxels is critical if information about outcomes (e.g., mental states) is encoded in neuronal population codes: A population of neurons can encode a much larger amount of specific information than individual neurons, which can encode only a limited amount of information with less specificity and generalizability than population codes [22]. Furthermore, assessing large-scale patterns across systems is critical if mental states are encoded across systems [21]. Second, predictive mapping uses distributed, multivariate information without requiring thresholding or correction for multiple comparisons, which could cause much of the important signal to be missed. Third, predictive mapping can effectively filter voxels that have weak, spurious relationships with the outcome or identify voxels that are masked or suppressed by some correlated noise across brain regions (see Figure 2 for details).

In the predictive mapping approach, machine learning techniques become crucial because analyses based on distributed information across multiple brain systems (in other words, large-scale population codes) are subject to the high dimensionality problem. High dimensional data, where we have many more predictors than observations ($p \gg n$; often called the "curse of dimensionality"), causes problems with model optimization because the parameter space is underconstrained by the data [23]. Some machine learning algorithms, such as support vector machines and regularized regression methods, can provide stable prediction models even for the high dimensional data with a guarantee of good generalization capacity [24, 25]. Therefore, machine learning algorithms

provide critical tools for developing useful neuroimaging markers with strong diagnostic performance.

Predictive mapping versus information-based mapping (Box 1)

A related approach, called information-based mapping or searchlight MVPA [26], also uses multivariate patterns to predict outcomes, but it still focuses on localized effects using searchlights (often defined using small spheres around center voxels) or regions-of-interest (ROIs). The information-based mapping commonly consists of the following steps [27]: 1) creating a searchlight using a small sphere or cubic around a center voxel, 2) training and testing pattern classifiers on fMRI signals within the searchlight, and 3) repeating the first and second steps for each voxel within the whole brain or a ROI mask. Using these steps, information-based mapping tries to answer a similar question as the traditional mapping approaches: Is there information within each searchlight?

Though information-based mapping is currently the most popular MVPA method (Figure 3A), it has several drawbacks in developing neuroimaging markers besides some methodological issues [28, 29]. First, information-based mapping has limited sensitivity and specificity compared to the predictive mapping approach. It is because information-based mapping does not take into account the macro-scale information distributed across multiple brain systems and interactions between voxels in and outside of searchlights. The limited sensitivity and specificity of searchlight-based methods has been demonstrated in recent studies [21, 30, 31]. For example, a recent study [21] showed that the brain-wide patterns of fMRI activity explained 72.2% of variance in negative emotion ratings, whereas the maximum amount of variance explained by searchlights was only 19.4% (Figure 4). In addition, whole-brain classifiers showed better separate modifiability across different mental states than searchlight classifiers [30, 31]. Second, information-based mapping consists of massive multiple tests and requires correction for multiple comparisons, which could cause inflation of accuracy estimation and

therefore produce less generalizable predictive models. Conversely, the predictive mapping approach needs neither multiple tests nor correction for multiple comparisons. Third, information-based mapping does not provide one unified predictive model and often focus on idiographic information within each subject (Figure 3A). Therefore it is difficult to validate and test the sensitivity and specificity of predictive models derived from the information-based mapping approach. In contrast, predictive mapping provides one unified predictive model that can be easily applied to new individuals and data. Further, the predictive map from predictive mapping can be subsequently interrogated and tested for information content; for example, one can assess which regions make the most important contributions, and whether local information is sufficient to capture the outcome [21].

Through a survey of 131 fMRI MVPA studies published between January 2014 and June 2015, we found that the most popular MVPA approach is currently information-based mapping that uses local pattern information for each individual (Figure 3A): Searchlight or ROI-based approach was adopted by 69% of 131 MVPA studies, of which 92% focused on idiographic (within-subject) patterns of fMRI activity. These survey results suggest that a majority of researchers are interested in localized and idiographic effects rather than developing and validating fMRI-based markers that work across individuals.

Precisely defined model and prospective testing: Benefits for translational research

One of the most important features of the predictive mapping approach is that it can provide a precisely defined marker that can be easily tested on new individuals, scans, and studies. The ease with which a model can be applied and tested across laboratories and settings is a critical practical constraint on translational development.

Replication and hypothesis testing in neuroimaging studies have depended heavily on anatomical definitions that are often heuristic and ambiguous, leading to flexibility in how researchers identify what counts as an *a priori* hypothesis and, in turn, increases in false positive results and reduced specificity [32]. For example, "amygdala activity" does not provide a reproducible definition of precisely (a) which voxels in the amygdala should be activated (there are typically hundreds) and (b) the relative expected intensity of activity across each voxel. Any significant results *anywhere* in the amygdala can count as a positive amygdala finding, which could be spurious findings. Such flexibility also reduces the specificity of positive findings, as many tasks and outcomes may produce activation somewhere within the amygdala, even if the locations and patterns are different across tasks. In contrast, predictive mapping can minimize biases in measuring, testing, and replicating effects in new individuals and studies by providing pre-defined, *a priori* predictive models.

Precisely defined predictive models, based on multivariate patterns of neuroimaging data, provide several advantages for basic and translational research. First, because hypotheses are precisely specified in terms of spatial patterns, responses in these patterns are falsifiable and readily testable, providing a foundation for strong scientific inference [33, 34]. Second, precisely specified markers are "research products" that can be shared and tested across laboratories, enabling a cumulative understanding of their properties across test conditions and study populations. Such products, including protocols and assays, have been bedrock tools in science (e.g., molecular biology). Third, pre-defined predictive models can be prospectively applied to individual participants, which is critical for clinical and legal applications. Fourth, well-defined predictive models can serve as a means of bringing together basic and clinical research. Diverse research groups can take markers from other groups and use them to answer their own research questions, and communicate with each other through tests of the markers. This can facilitate translation of findings from one setting (e.g., basic research) into new contexts (e.g., clinical assessment).

Multi-study validation of neuroimaging markers

Establishing neuroimaging-based markers for mental processes and clinical outcomes requires a long development and validation process supported by multi-study and multi-site efforts. Most current work using the predictive mapping approach often ends with one publication demonstrating proof-of-concept. As Figure 3B presents, our survey results show that only four among 131 MVPA studies used additional independent datasets to test their predictive models, demonstrating that researchers put little effort to validate the developed markers through independent, multiple studies. However, the development of predictive markers should be the beginning rather than the end-point. Like other types of biomarkers [35], useful neuroimaging markers need to demonstrate increasing levels of supporting evidence and generalizability through multiple stages (Figure 5), as the best neuroimaging markers are moved towards larger-scale studies and clinical use [4]. The more neuroimaging markers hold up to the scrutiny of being characterized across samples, conditions, and populations, the more useful they will become.

As Figure 6 shows, the development and validation of neuroimaging markers can be an iterative process of optimization, generalization, and characterization. In the "optimization" stage, researchers identify a neuroimaging marker that is diagnostic of a particular mental process or a functional outcome. At this stage, researchers develop the neuroimaging marker as a 'research product' that can be prospectively tested on multiple samples from different laboratories, and validated or challenged in various ways. In the "generalization" stage, the marker should prove its performance to be generalizable across different individuals, scanners, scanning protocols, and populations. Along with the generalizability tests, researchers focus on "characterizing" the marker in terms of its properties across test conditions and define its boundary conditions under which the marker is valid and useful. The neuroimaging markers might need a refinement with more data during the validation process, but in that case, the improved marker should enter the iterative validation process again.

Desirable characteristics of useful neuroimaging markers (Box 2)

A neuroimaging marker should demonstrate a set of desirable characteristics throughout the validation process to be a useful marker—diagnosticity, interpretability, deployability, and generalizability [36] (Table).

First, good neuroimaging markers should produce high *diagnostic* performance (high sensitivity and specificity) in classification or prediction. Diagnostic performance can be evaluated using positive and negative control conditions. The positive control is the condition where the marker response is expected, whereas the negative control is the condition where the marker response is not expected, but which is confusable with the condition of interest. Positive findings in negative control conditions can invalidate the test or provide information on the limits of its specificity (e.g., [37]).

Second, *interpretability* concerns whether the marker is meaningful and interpretable in terms of neuroscience, and can be evaluated in the light of prior neuroimaging studies and converging evidence from multiple sources (e.g., animal models, lesion studies, etc.). One potential pitfall in developing neuroimaging markers is that classification or prediction models can capitalize on confounding variables that are not neuroscientifically meaningful or interesting at all (e.g., in-scanner head movement [38]). If results are reliable but caused by artifacts, they will not generalize well when the artifacts are not present; therefore, it is difficult to determine when a test will fail if we do not know why it succeeded.

Third, *deployability* is about how easy and feasible it is to use and test a neuroimaging marker in diverse settings across research groups and clinics. In addition to precisely defined predictive models, a simple and standardized testing procedure is a necessary condition to maximize deployability of a marker.

Fourth, a useful neuroimaging marker should prove that its performance is *generalizable* across different laboratories, different scanners, scanning protocols, different populations, and variants of testing conditions. Generalizability can be

evaluated through the generalization and characterization stages described in Figure 6 through multiple studies. If a marker is not easily deployable, it may be impractical to assess its generalizability.

Neuroimaging studies will benefit from criteria such as those described above, which provide systematic evaluation and recommendation of existing and new neuroimaging markers. When a neuroimaging marker demonstrates these four desirable characteristics, it has a good chance of being useful, and can be a good candidate for the larger-scale tests required for translation from research to clinical settings.

An example case: Neurologic Pain Signature

The neurologic pain signature (NPS) provides an example of a promising neuroimaging marker, which is predictive of pain induced by noxious input [20]. As shown in Figure 7, the NPS is based on brain-wide, meso-scale patterns of fMRI activity across multiple pain-related regions. The NPS can be easily applied to new individuals and datasets for prospective testing and validation. Across four independent studies, Wager et al. showed that the NPS response discriminated somatic pain from non-painful warmth, pain anticipation, painful emotions induced by social rejection, and pain recall with 90-100% sensitivity and specificity ("diagnosticity") [20]. Importantly, the NPS does not simply measure the intensity of noxious input: It tracked subjective pain more closely than the noxious stimulus itself, responded more strongly when pain was more intense at a fixed noxious stimulus intensity, and was substantially reduced by the opiate analgesic drug [20].

Wager et al. identified brain regions that reliably contribute to the prediction using bootstrap tests (Figure 7A). Also, to maximize "interpretability" of predictive weights, the NPS was developed using a linear predictive model; the brain regions identified have a linear relationship with pain. The regions that have a positive relationship with pain included well-defined targets of ascending nociceptive brain pathways [39, 40]. For example, positive predictive weights were reliably found in

the dosal posterior insula (dpINS), secondary somatosensory cortex, ventrolateral and medial thalamus, anterior insula, dorsal anterior cingulate cortex; these regions have been associated with pain and noxious stimulus intensity in prior animal and human literature [39-45]. In addition, negative predictive weights were observed in regions associated with pain regulation (e.g., pregenual cingulate and ventrolateral prefrontal cortex) [6, 46] and non-somatic sensory and cognitive processes (visual cortex, precuneus) [47].

The NPS has been tested across multiple studies using a simple and easy testing procedure, demonstrating its high "deployability" (Figure 8) [20, 21, 31, 48-50]. As shown in Figure 7B, the NPS can be expressed as a linear weight vector ($\vec{w}$) on a standard brain space (e.g., the Montreal Neurological Institute space), and a dot-product between two linear vectors—the NPS and a new brain activation map ($\vec{\beta}^T \vec{w}$, resulting in a scalar value)—can define signature response. The signature response then can be used in a two-choice (forced-choice) classification or receiver-operating-characteristic analysis for a test of high versus low pain or in calculating an outcome-prediction correlation for predicting of continuous pain experience. Based on this simple and standardized testing procedure, the NPS can be easily shared across laboratories and tested on new datasets to answer different research questions (Figure 8).

The NPS has proved to be generalizable across different laboratories, scanning protocols, and different pain stimuli including thermal, pressure, and electrical shock pain [20, 31, 48]. There are also other ongoing studies to examine the signature's boundary condition, such as clinical pain and different types of aversive stimuli. Through multi-study and multi-site efforts like this, neuroimaging markers will be able to be better characterized with increasing levels of supporting evidence for its use in research and clinical settings.

Implications for basic and clinical neuroimaging research

As shown in Figure 3C, our survey results show a discrepancy between basic (non-clinical) and clinical research in how they are using MVPA. Among 15 clinical MVPA studies, 73.3% ($n = 11$) adopted the predictive mapping approach; they used brain-wide pattern information to identify fMRI-based markers that can be generalizable across subjects. In contrast, only 15.5% ($n = 18$) among 116 non-clinical studies used the predictive mapping approach. These results suggest that many clinical researchers are interested in developing and utilizing neuroimaging markers for individual-level predictions, whereas non-clinical researchers focus more on local mapping of brain functions.

Despite different preferences of researchers, the predictive mapping approach has a potential to benefit both basic and clinical research. For basic research, the predictive mapping approach provides an analysis framework that has high sensitivity to the effects of interest. In addition, a specific and diagnostic relationship between brain markers and mental states (or processes) can provide a good proxy for neural representations of particular mental states. Predictive mapping can also be used to assess localized effects (but in the context of whole brain) through supplemental analyses such as bootstrap tests [20, 21, 30]. Finally, neuroimaging markers as research products provide an effective and robust way to establish replicability and falsifiability, which ensure good scientific practice. For the clinical researchers, predictive mapping can provide clinical markers that can be used to assess risk factors, stratify patients into biological subgroups (diagnosis), predict illness trajectory (prognosis), help treatment choice (prediction of treatment effects), and more [2, 51, 52]. In addition, those clinical biomarkers can be used to assess treatment effects and study the underlying mechanisms of pharmacological, cognitive, and behavioral interventions [53-55].

Predictive mapping could also facilitate translations in both directions—from clinical to basic, and from basic to clinical—and integration of findings from both basic and clinical research. By virtue of deployability of neuroimaging markers, researchers can easily examine the relationships between basic cognitive, affective, and social processes and health-related outcomes. Ideally, we hope that the

predictive mapping approach could start active conversation and collaboration between basic and clinical groups where a wide gap currently exists [56, 57]. For example, pharmacological and psychological influences on pain have been studied separately. However, those influences might interact with one another in more profound ways than we have thought. Possibly, there might be a particular mental state or social support to help a particular drug to work more effectively. Using brain signatures developed by pharmacology or psychology research groups, we might be able to start comparing two different influences and building integrative models about the interactions [6, 55].

For effective translation and communication between basic and clinical research groups, we have a specific recommendation for each of the groups. First, we encourage basic researchers (e.g., cognitive, affective, and social neuroscientists) to develop *deployable* neuroimaging markers for their research domain and share them with clinical research groups. Consistent with the idea of research domain criteria (RDoC), utilizing basic neurobiological and cognitive dimensions could lead us to a better understanding and treatment of brain-related illnesses than relying on current diagnostic system [3]. Therefore, markers for basic building blocks of mental processes could have a big impact on how we understand, evaluate, and treat psychiatric and neurologic disorders. Second, for clinical researchers, we recommend focusing more on *interpretability* of neuroimaging markers. Clinical researchers and clinicians tend to be more interested in predictive performance (e.g., when predicting treatment response using fMRI [58]), whereas basic researchers are more interested to know its neuroscientific meaning (e.g., why the fMRI signature is predictive of treatment response?). For better communication between two groups, each group needs to pay attention to what the other group is interested in.

Broader implications for legal and commercial use of neuroimaging (Box 3)

The predictive mapping approach has broader implications for legal and commercial use of neuroimaging. One critical example is the problem of unresolved pain. A large number of lawsuits involve claims about pain, and needs for corroboration with objective measures is often pressing [59]. Based on the potential use of neuroimaging in proving unresolved pain, there have already been a few attempts to use neuroimaging in real legal or commercial settings, and some companies have already started to use neuroimaging in a for-profit setting [59]. However, it is unethical to use neuroimaging methods without proper validation, including establishing sensitivity, specificity, and generalizability as described above. There is not yet any set standard for how far along the validation process a marker needs to be done, or how well it needs to work. However, diagnostic and treatment procedures that are spurious or do not work as claimed are clearly unethical.

Therefore, for future medical, legal, and commercial applications of neuroimaging, we need a well-established validation system for neuroimaging markers. As a good example, MammaPrint® (Agendia), a gene expression signature for breast cancer, has been through more than 20 validation studies by different laboratories from more than 5 different countries, and eventually received clearance from the Food and Drug Administration (FDA) in 2007 [60]. Likewise, before being able to use neuroimaging markers in real-world applications, we need concentrated efforts to validate and test potential candidate markers.

Conclusions

Recent advances in neuroimaging have provided promise and hope that we can use neuroimaging to better understand the human mind, including understanding of the neurophysiology that underlies behavior and brain-related illnesses. However, a wide gap still exists between neuroimaging data and the mental processes we want to measure. Part of the problem is that we do not have neuroimaging markers that are sensitive and specific enough to inform us that

some particular class of mental processes is engaged. The predictive mapping approach we outline here can be used to develop neuroimaging markers that have better sensitivity and specificity, compared to traditional and information-based mapping approaches. The predictive mapping approach can also provide precisely defined predictive models that can be prospectively tested on new individuals and studies, and thereby turn the predictive models into research products. This characteristic can help neuroimaging markers to be easily accessible and testable by other researchers and laboratories, and promote replicability and facilitate translations. All together, the predictive mapping approach has potential to facilitate neuroimaging marker discovery and validation for both basic and clinical science.
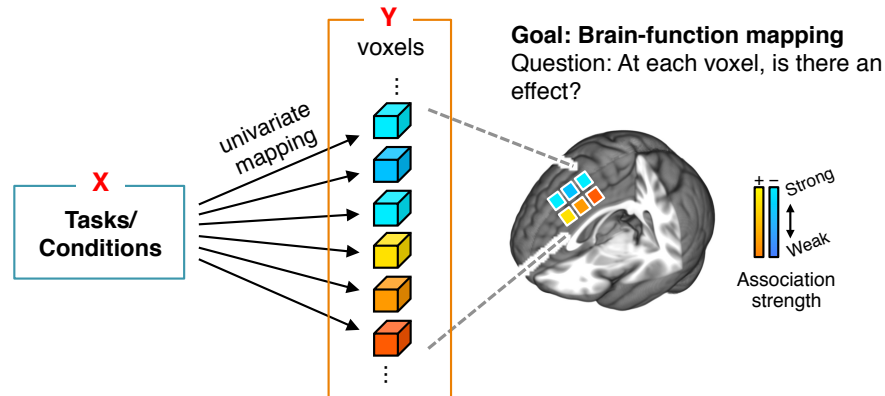
Acknowledgments

References

1 Mather, M.*, et al.* (2013) How fMRI Can Inform Cognitive Theories. *Perspect Psychol Sci* 8, 108-113

2 Singh, I. and Rose, N. (2009) Biomarkers in psychiatry. *Nature* 460, 202-207

3 Insel, T.R. and Cuthbert, B.N. (2015) Brain disorders? Precisely. *Science* 348, 499-500

4 Borsook, D.*, et al.* (2011) Biomarkers for chronic pain and analgesia. Part 1: the need, reality, challenges, and solutions. *Discovery medicine* 11, 197-207

5 Borsook, D.*, et al.* (2011) Biomarkers for chronic pain and analgesia. Part 2: how, where, and what to look for using functional imaging. *Discov. Med.* 11, 209-219

6 Wager, T.D. and Atlas, L.Y. (2015) The neuroscience of placebo effects: connecting context, learning and health. *Nat Rev Neurosci* 16, 403-418

7 Poldrack, R.A. (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* 72, 692-697

8 Haxby, J.V.*, et al.* (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425-2430

9 Kamitani, Y. and Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8, 679-685

10 Kvitsiani, D.*, et al.* (2013) Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* 498, 363-366

11 Sheth, S.A.*, et al.* (2012) Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* 488, 218-221

12 Paton, J.J.*, et al.* (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865-870

13 Yarkoni, T.*, et al.* (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8, 665-U695

14 Shackman, A.J.*, et al.* (2011) The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat Rev Neurosci* 12, 154-167

15 Cunningham, W.A. and Brosch, T. (2012) Motivational Salience: Amygdala Tuning From Traits, Needs, Values, and Goals. *Curr Dir Psychol Sci* 21, 54-59

16 Nichols, T. and Hayasaka, S. (2003) Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research* 12, 419-446

17 Woo, C.W.*, et al.* (2014) Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 91, 412-419

18 Yarkoni, T. (2009) Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). *Perspect Psychol Sci* 4, 294-298

19 Button, K.S.*, et al.* (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14, 365-376

20 Wager, T.D.*, et al.* (2013) An fMRI-Based Neurologic Signature of Physical Pain. *New Engl J Med* 368, 1388-1397

21 Chang, L.J.*, et al.* (2015) A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. *PLoS biology* 13, e1002180

22 Georgopoulos, A.P.*, et al.* (1986) Neuronal Population Coding of Movement Direction. *Science* 233, 1416-1419

23 Clarke, R.*, et al.* (2008) The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. *Nat Rev Cancer* 8, 37-49

24 Vapnik, V. (1995) *The nature of statistical learning theory*. Springer

25 Hastie, T.*, et al.* (2009) *The elements of statistical learning : data mining, inference, and prediction*. Springer

26 Kriegeskorte, N.*, et al.* (2006) Information-based functional brain mapping. *P Natl Acad Sci USA* 103, 3863-3868

27 Kahnt, T.*, et al.* (2010) The neural code of reward anticipation in human orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107, 6010-6015

28 Etzel, J.A.*, et al.* (2013) Searchlight analysis: promise, pitfalls, and potential. *Neuroimage* 78, 261-269

29 Todd, M.T.*, et al.* (2013) Confounds in multivariate pattern analysis: Theory and rule representation case study. *Neuroimage* 77, 157-165

30 Woo, C.W.*, et al.* (2014) Separate neural representations for physical pain and social rejection. *Nat Commun* 5

31 Krishnan, A.*, et al.* (under review) Somatic and vicarious pain are represented by dissociable multivariate brain patterns.

32 Simmons, J.P.*, et al.* (2011) False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol Sci* 22, 1359-1366

33 Popper, K.R. (1959) *The logic of scientific discovery*. Basic Books

34 Platt, J.R. (1964) Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science* 146, 347-353

35 Wagner, J.A. (2008) Strategic approach to fit-for-purpose biomarkers in drug development. *Annual review of pharmacology and toxicology* 48, 631-651

36 Woo, C.W. and Wager, T.D. (2015) Neuroimaging-based biomarker discovery and validation. *Pain*

37 Uncapher, M.R.*, et al.* (2015) Goal-Directed Modulation of Neural Memory Patterns: Implications for fMRI-Based Memory Detection. *J Neurosci* 35, 8531-8545

38 Eloyan, A.*, et al.* (2012) Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging. *Frontiers in systems neuroscience* 6, 61

39 Willis, W.D. and Westlund, K.N. (1997) Neuroanatomy of the pain system and of the pathways that modulate pain. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society* 14, 2-31

40 Fields, H. (2004) State-dependent opioid control of pain. *Nat Rev Neurosci* 5, 565-575

41 Apkarian, A.V.*, et al.* (2005) Human brain mechanisms of pain perception and regulation in health and disease. *European journal of pain* 9, 463-484

42 Coghill, R.C.*, et al.* (1999) Pain intensity processing within the human brain: A bilateral, distributed mechanism. *Journal of neurophysiology* 82, 1934-1943

43 Bushnell, M.C.*, et al.* (2013) Cognitive and emotional control of pain and its disruption in chronic pain. *Nat Rev Neurosci* 14, 502-511

44 Price, D.D. (2000) Psychological and neural mechanisms of the affective dimension of pain. *Science* 288, 1769-1772

45 Segerdahl, A.R.*, et al.* (2015) The dorsal posterior insula subserves a fundamental role in human pain. *Nat Neurosci* 18, 499-500

46 Wiech, K.*, et al.* (2008) Neurocognitive aspects of pain perception. *Trends in cognitive sciences* 12, 306-313

47 Bingel, U.*, et al.* (2007) fMRI reveals how pain modulates visual object processing in the ventral visual stream. *Neuron* 55, 157-167

48 Lopez-Sola, M.*, et al.* (under review) A neurophysiological signature for fibromyalgia.

49 Vachon-Presseau, E.*, et al.* (under review) Multiple faces of pain: Specific effects of chronic pain on the brain regulation of facial expression.

50 Woo, C.W.*, et al.* (2015) Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain. *PLoS biology* 13, e1002036

51 Risacher, S.L. and Saykin, A.J. (2013) Neuroimaging and other biomarkers for Alzheimer's disease: the changing landscape of early detection. *Annual review of clinical psychology* 9, 621-648

52 Kapur, S.*, et al.* (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatr* 17, 1174-1179

53 Hargreaves, R.J.*, et al.* (2015) Optimizing CNS drug development using molecular imaging. *Clinical Pharmacology & Therapeutics*, n/a-n/a

54 Duff, E.P.*, et al.* (2015) Learning to identify CNS drug action and efficacy using multistudy fMRI data. *Science translational medicine* 7, 274ra216

55 Wager, T.D. and Woo, C.W. (2015) fMRI in analgesic drug discovery. *Science Tanslational Medicine* 7, 274fs276

56 Butler, D. (2008) Translational research: Crossing the valley of death. *Nature* 453, 840-842

57 Holmes, E.A.*, et al.* (2014) A call for mental-health science. *Nature* 511, 287-289

58 Doehrmann, O.*, et al.* (2013) Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA psychiatry* 70, 87-97

59 Reardon, S. (2015) The Painful Truth. *Nature* 518, 474-476

60 Delahaye, L.J.M.*, et al.* (2013) Performance characteristics of the MammaPrint (R) breast cancer diagnostic gene signature. *Pers Med* 10, 801-811

A Traditional brain mapping approach (univariate analysis)

Y
voxels

univariate mapping

X
Tasks/ Conditions

Goal: Brain-function mapping
Question: At each voxel, is there an effect?

± − Strong
Weak
Association strength

B Predictive mapping approach (multivariate analysis)

X
voxels

brain-wide pattern

Machine Learning

Goal: Developing brain markers
Question: Is a pattern sensitive and specific to a mental process or an outcome?

Y
Tasks/ Conditions/ Outcomes

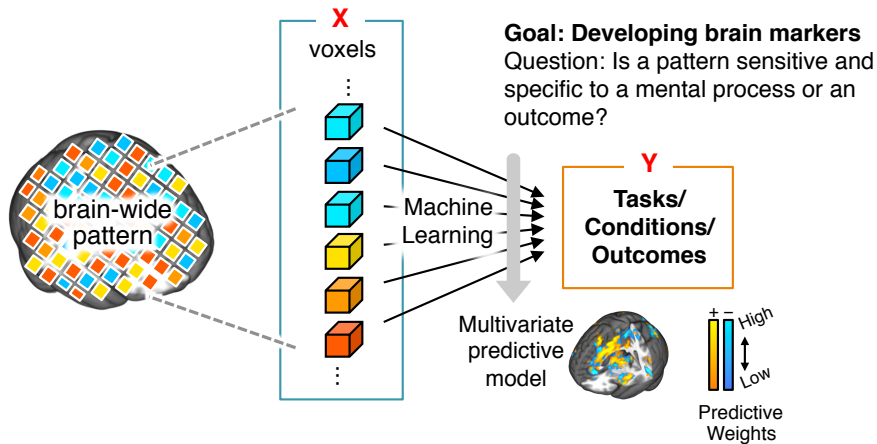Multivariate predictive model

± − High
Low
Predictive Weights

Figure 1. Traditional versus predictive mapping approach. A. Traditional mapping (univariate analysis) aims to obtain the functional architecture of the brain by localizing the effects in the brain. This approach is often suffers from low sensitivity and specificity. **B.** The predictive mapping approach aims to develop a multivariate, brain-wide prediction (decoding) model that is sensitive and specific to the outcomes-of-interest.
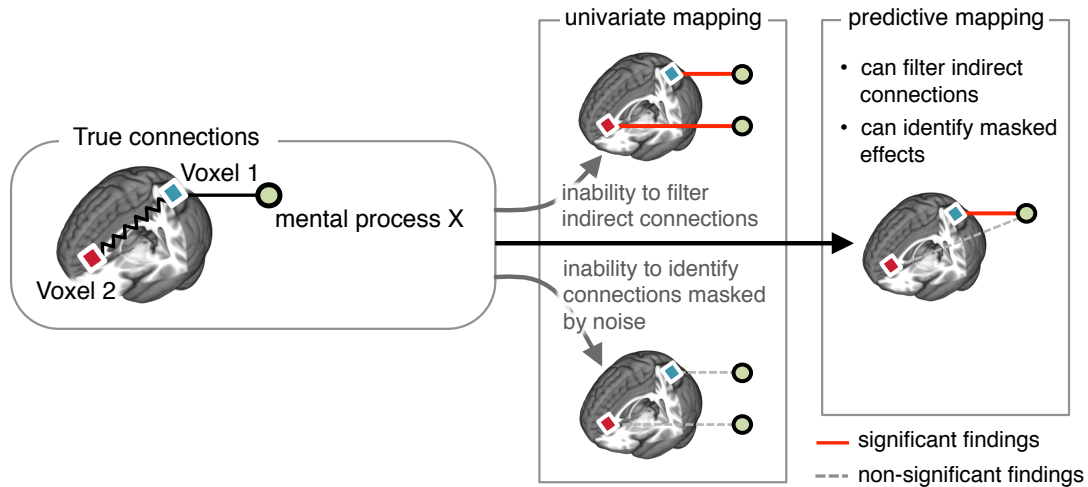
Figure 2. Benefits of predictive mapping. There are a number of scenarios under which predictive mapping yields higher power and more accurate representation of brain-outcome relationships than traditional (univariate) mapping. We illustrate one such scenario here, in which there is a direct relationship between Voxel 1 and the outcome (mental process X), and a correlation between Voxel 1 and Voxel 2 that reflects a common source of noise unrelated to the outcome. In the top panel, univariate mapping may identify significant effects in both voxels, as it cannot separate regions with indirect connections to X from those that have more direct relationships. The predictive mapping approach controls for Voxel 1 when assessing the effects of Voxel 2, and so will not spuriously identify Voxel 2 as significant. In addition, controlling for Voxel 2 can remove some of the noise in Voxel 1, which may otherwise mask the relationship between Voxel 1 and X and prevent it from reaching significance (bottom panel). For both reasons, under this scenario, the predictive mapping approach has a greater chance of identifying the true relationships.
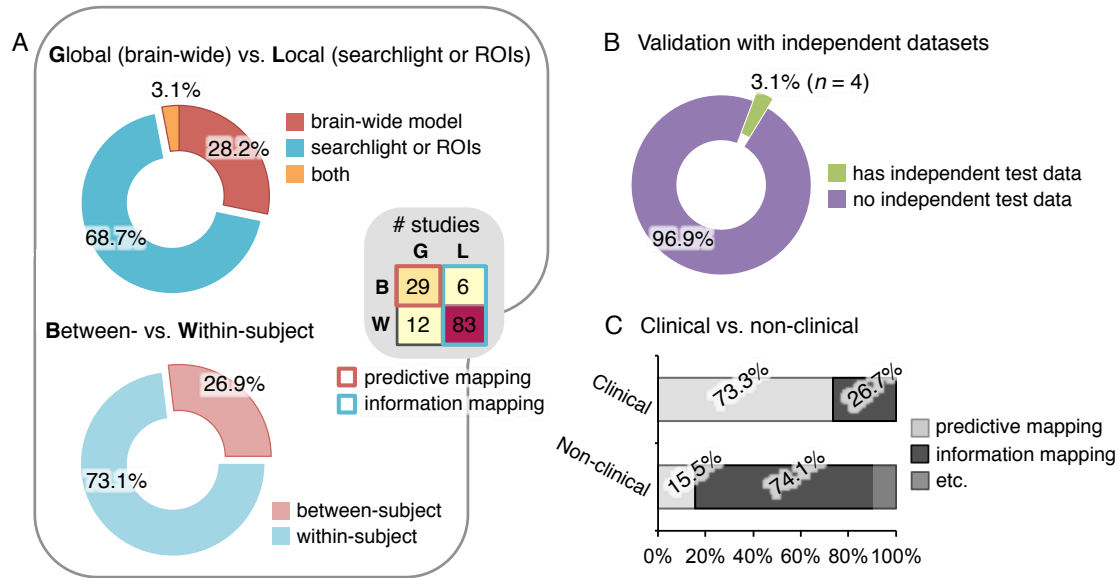
Figure 3. Survey results. We conducted a literature survey of recent fMRI studies that used multivariate pattern analysis (*N* = 131 studies published between Jan 2014 and Jun 2015). **A.** The proportion of studies using brain-wide vs. local region information (top) and between- vs. within-subject information (bottom). The table in the middle panel shows the number of studies for each combination of conditions. The cell within the red outlined box indicates predictive mapping (global and between-subjects), while the cells within the blue outlined box indicates information mapping (local). **B.** The proportion of studies that have independent test datasets. **C.** The proportion of studies using the predictive mapping approach for clinical vs. non-clinical (basic) studies.
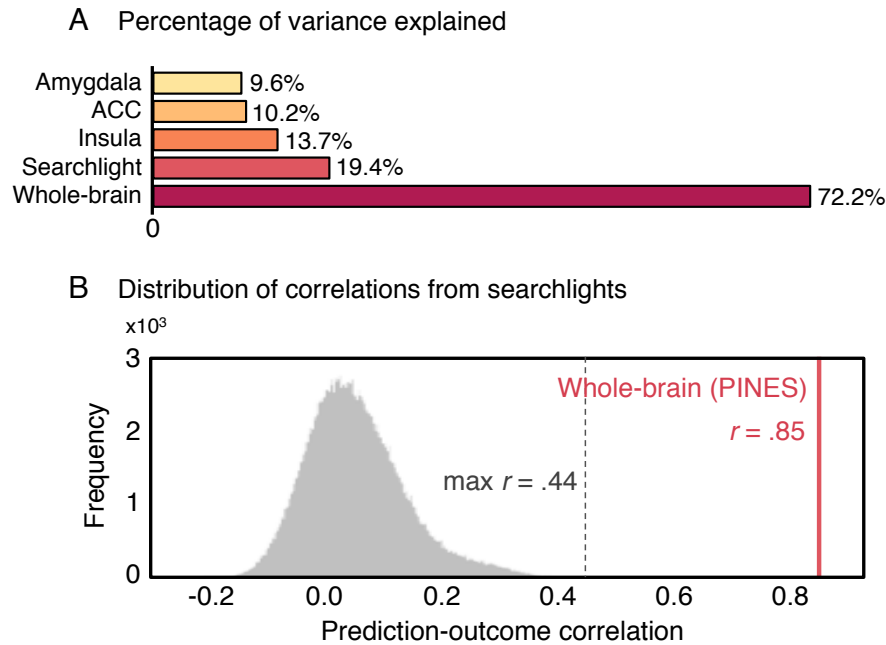
A   Percentage of variance explained

Amygdala 9.6%
ACC 10.2%
Insula 13.7%
Searchlight 19.4%
Whole-brain 72.2%

B   Distribution of correlations from searchlights

Whole-brain (PINES)
$r = .85$
max $r = .44$

Figure 4. An example of predictive performance with the whole brain versus local regions (Reproduced from [21]). **A.** The percentage of variance explained by predictive models based on local regions vs. the whole brain when predicting negative emotion ratings induced by aversive pictures. Local regions include amygdala, anterior cingulate cortex (ACC), insula, and searchlights (sphere with 5 voxel radius). **B.** Distribution of prediction-outcome correlations from searchlights. The dashed line represents the maximum correlation when searchlights used, and the red line shows the prediction-outcome correlation with the whole brain.
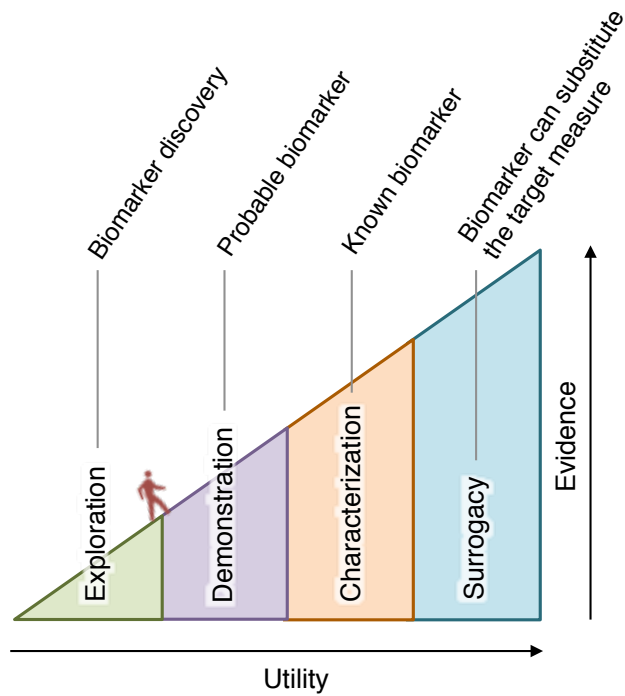
Figure 5. To be a useful neuroimaging marker, it needs to demonstrate increasing levels of supporting evidence and generalizability through multiple stages. Therefore, establishing neuroimaging markers requires a long development and validation process supported by multi-study and multi-site efforts (Reproduced from [35]).
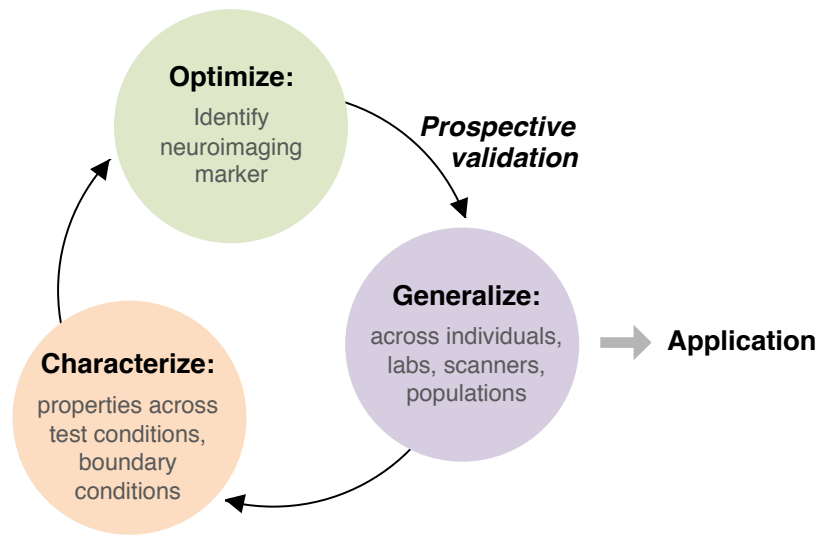
Figure 6. Neuroimaging marker development and validation process. This can be described as an iterative process of optimization, generalization, and characterization (for details, see text).
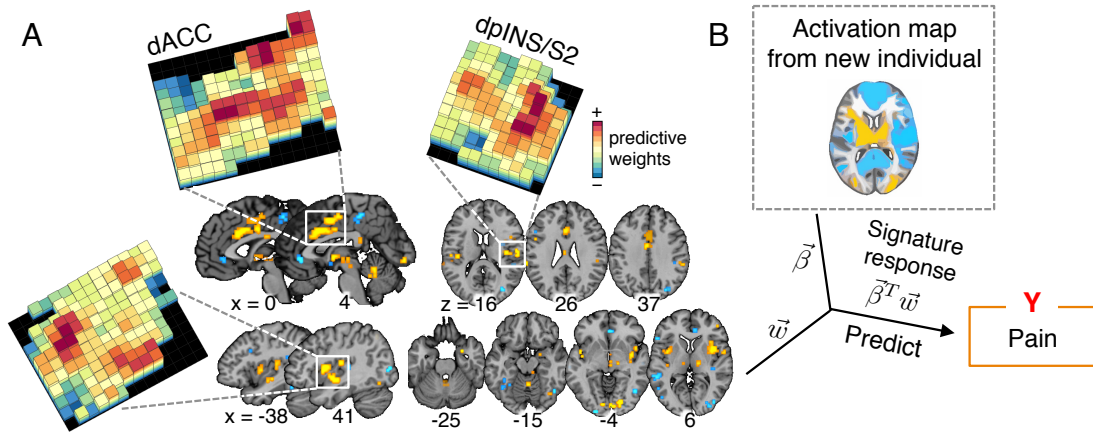
Figure 7. The Neurologic Pain Signature (NPS). A. The NPS is defined by brain-wide, meso-scale patterns of fMRI activity across multiple pain-related regions, and can be prospectively tested on new individuals and datasets. The brain map shown here is the thresholded pattern map ($q < 0.05$, false discovery rate [FDR]) for display only. All voxels within NPS should be used to predict pain in new individuals. Some examples of unthresholded patterns are presented in the insets; each 3-d bar represents one voxel. dACC, dorsal anterior cingulate cortex; INS, insula; dpINS, dorsal posterior insula; S2, secondary somatosensory cortex. Reproduced with permission from [20]. **B.** Signature response can be calculated using a dot-product between the NPS and an activation map.
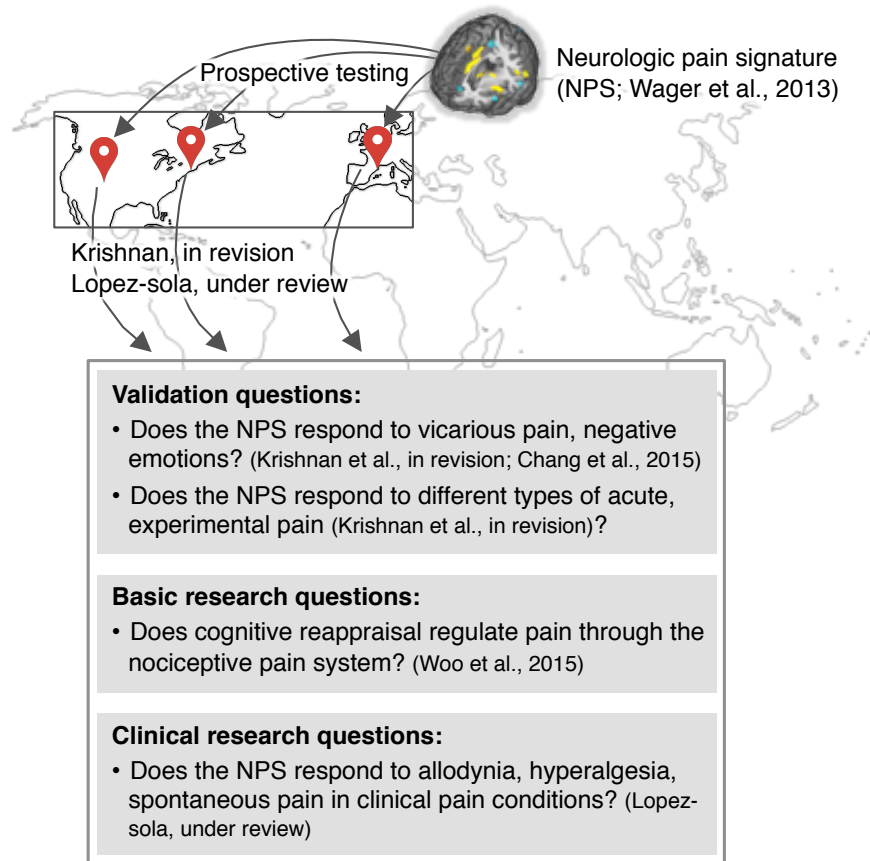
Neurologic pain signature
(NPS; Wager et al., 2013)

Prospective testing

Krishnan, in revision
Lopez-sola, under review

Validation questions:
• Does the NPS respond to vicarious pain, negative emotions? (Krishnan et al., in revision; Chang et al., 2015)
• Does the NPS respond to different types of acute, experimental pain (Krishnan et al., in revision)?

Basic research questions:
• Does cognitive reappraisal regulate pain through the nociceptive pain system? (Woo et al., 2015)

Clinical research questions:
• Does the NPS respond to allodynia, hyperalgesia, spontaneous pain in clinical pain conditions? (Lopez-sola, under review)

Figure 8. Deployability of the NPS. The NPS has been applied and tested on datasets from different laboratories around the world to answer different types of research questions.

Table. Desirable characteristics of neuroimaging biomarkers

Development Stages	Criteria	Definition	Test setting
Discovery	1 Diagnosticity	Sensitivity: positive results when a target psychological or behavioral process is engaged	Positive control
		Specificity: positive results exclusively when the target process is engaged	Negative control
	2 Interpretability	Neuroscientifically interpretable model	Neuroscience literature, meta-analysis, animal models, lesion studies
Validation	3 Deployability	Easy to apply the marker across different research groups and clinics	Well-specified predictive model, simple and standardized testing procedure
	4 Generalizability	Generalizable across different laboratories, scanners, populations, and variants of testing conditions	New test studies (with multi-study, multi-site efforts)

Supplemental methods

Literature survey

We conducted a literature survey to examine how fMRI researchers are using multivariate pattern analysis (MVPA) methods. PubMed was searched for original fMRI research articles published between January 2014 and June 2015 using the following two search queries:

(fMRI[Title/Abstract]) AND (decoding[Title/Abstract] OR multivariate[Title/Abstract] OR multivoxel[Title/Abstract] OR mvpa[Title/Abstract]) AND ("2014/01/01"[Date - Create] : "3000"[Date - Create])

(fMRI[Title/Abstract]) AND (biomarker*[Title/Abstract] OR marker*[Title/Abstract] OR signature*[Title/Abstract]) AND (pattern*[Title/Abstract] OR decoding[Title/Abstract] OR multivariate[Title/Abstract] OR multivoxel[Title/Abstract] OR mvpa[Title/Abstract]) AND ("2014/01/01"[Date - Create] : "3000"[Date - Create])

Exclusion criteria included (1) method or review papers and (2) non-human studies. In addition, one paper was excluded because its fulltext was not available. The survey data are available at
http://wanirepo.github.io/data/Woo_comps_lit_survey_share.xlsx.