Commentary

PAIN®

Neuroimaging-based biomarker discovery and validation

Choong-Wan Woo[a,b], Tor D. Wager[a,b]

Although developing biological markers for chronic pain has been a major goal of the field for decades, such biomarkers have not yet made their way into clinical practice. However, given the potential uses of biomarkers in multiple aspects of prevention and treatment—such as pain and risk factor assessment, diagnosis, prognosis, treatment selection, drug discovery, and more—efforts to discover new pain biomarkers have been expanding.[5,6,8,30]

Recent advances in human neuroimaging, including functional and structural magnetic resonance imaging (fMRI, sMRI) combined with machine learning techniques, are bringing us closer to the goal of developing objective brain-based markers of the neural functions and neuropathology that underlie chronic pain.[2,7,25,33] These brain measures are particularly promising as biomarkers for chronic pain. Although pain is reliably induced by peripheral nociceptive input, many forms of chronic pain may arise from neuropathology in the supraspinal circuits that govern the construction of pain experience and long-term motivation.[1,14,26,32]

Particularly, structural neuroimaging measures could provide more stable markers of the neuropathology of chronic pain, including stable features underlying pain risk and resilience.[2,3,11,19,28,29] Gray matter changes have also been associated with a number of conditions that are often comorbid with chronic pain, including depression,[4,22,24] stress,[10,12,20] post-traumatic stress disorder,[17,21,27] and early-life adversity.[13,18,23,31] Therefore, structural measures may provide important clues about supraspinal contributions to both pain and related risk factors (**Fig. 1**).

In this issue, Labus et al.[16] developed a new neuroimaging biomarker for irritable bowel syndrome (IBS) using sMRI data, based on a relatively large sample of 80 patients with IBS and 80 healthy controls. They used sparse partial least squares-discriminant analysis, a method that allowed them to both develop a classification model based on the brain structure and identify the regions that make the most important contributions to the classification. They subsequently tested the predictive model on a "holdout" sample of 26 patients with IBS and 26 healthy controls. The model discriminated patients from controls with 70% accuracy (compared with a chance accuracy of 50%), providing a moderate but reliable morphological brain signature for IBS.

Rather than being the end of the story, this study serves as a starting point for biomarker discovery and validation. Like other brain "signatures,"[30] the signature they identified can become a "research product" that can be tested on multiple samples from different laboratories and validated or challenged in various ways. The more the marker for IBS status or IBS risk holds up to the scrutiny of being characterized across samples and populations, the more useful it will become.

Importantly, there is a set of desirable characteristics that a useful neuroimaging biomarker should demonstrate throughout the biomarker development process. We briefly describe several such characteristics (summarized in **Table 1**) and then relate them to the findings of Labus et al.[16]

1. Criterion 1: diagnosticity

Good biomarkers should produce high diagnostic performance in classification or prediction. Diagnostic performance can be evaluated by *sensitivity* and *specificity*. Sensitivity concerns whether a model can correctly detect signal when signal exists. Effect size is a closely related concept; larger effect sizes are related to higher sensitivity. Specificity concerns whether the model produces negative results when there is no signal. Specificity can be evaluated relative to a range of specific alternative conditions that may be confusable with the condition of interest.

2. Criterion 2: interpretability

Brain-based biomarkers should be meaningful and interpretable in terms of neuroscience, including previous neuroimaging studies and converging evidence from multiple sources (eg, animal models, lesion studies, etc). One potential pitfall in developing neuroimaging biomarkers is that classification or prediction models can capitalize on confounding variables that are not neuroscientifically meaningful or interesting at all (eg, in-scanner head movement[9]). Therefore, neuroimaging biomarkers should be evaluated and interpreted in the light of existing neuroscientific findings.

3. Criterion 3: deployability

Once the classification or outcome-prediction model has been developed as a neuroimaging biomarker, the model and the testing procedure should be precisely defined so that it can be prospectively applied to new data. Any flexibility in the testing procedures could introduce potential overoptimistic biases into test results, rendering them useless and potentially misleading. For example, "amygdala activity" cannot be a good neuroimaging biomarker without a precise definition of which "voxels" in the amygdala should be activated and the relative expected intensity of activity across each voxel. A well-defined model and

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, USA,* [b] *Institute of Cognitive Science, University of Colorado, Boulder, CO, USA*

PAIN 156 (2015) 1379–1381

http://dx.doi.org/10.1097/j.pain.0000000000000223

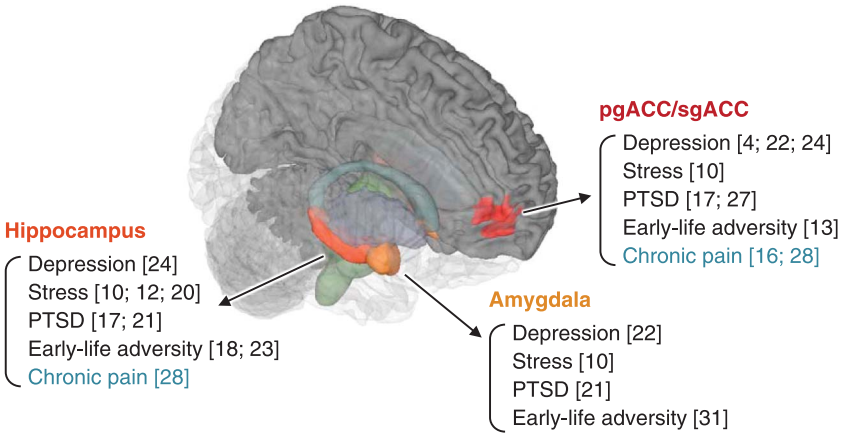

Figure 1. Key common brain regions that show structural changes across different conditions related to chronic pain, including depression, stress, posttraumatic stress disorder (PTSD), and early-life adversity.

standardized testing procedure are crucial aspects of turning neuroimaging results into a "research product," a biomarker that can be shared and tested across laboratories.

4. Criterion 4: generalizability

Clinically useful neuroimaging biomarkers aim to provide predictions about new individuals. Therefore, they should be validated through prospective testing to prove that their performance is generalizable across different laboratories, different scanners or scanning procedures, different populations, and variants of testing conditions (eg, other types of chronic pain). Generalizability tests inherently require multistudy and multisite efforts. With a precisely defined model and standardized testing procedure (criterion 3), we can easily test the generalizability of biomarkers and define the boundary conditions under which they are valid and useful.

5. Evaluating the neuroimaging biomarker for IBS by Labus et al.

We hope that more studies will use criteria such as those described above to evaluate existing and new biomarkers. Here, we apply our criteria to the new neuroimaging biomarker for IBS of Labus et al, and in so doing point towards some opportunities for future development.

Table 1
Desirable characteristics of neuroimaging biomarkers.

Development stages	Criteria		Definition
Discovery	1	Diagnosticity	Sensitivity: positive results when there is signal, effect size Specificity: negative results when there is no signal
	2	Interpretability	Neuroscientifically interpretable model
Validation	3	Deployability	Precisely defined model and standardized testing procedure (well-described, clear, and easy to deploy across research groups and clinics)
	4	Generalizability	Generalizable results across different laboratories, scanners, populations, and variants of testing conditions

5.1. Criterion 1

The sMRI signature for IBS by Labus et al. showed 68% sensitivity, 71% specificity, and 70% classification accuracy in holdout test data. Although this accuracy level is similar to other brain structure-based tests (eg, 73% accuracy[2]), it is not high enough to be used as a biomarker for IBS, as Labus et al. acknowledged. However, the signature could be still useful as a marker for a potential risk factor for IBS, in combination with other measures, or as a probe for resilience given a brain propensity for IBS. There are avenues for potential improvement, including refinement of the algorithm, generation and selection of important brain features, data quality control, multimodal assessment, and refined phenotyping (ie, using multiple functional or symptomatic outcomes rather than diagnostic categories). Labus et al. tested healthy controls, but later studies could also evaluate specificity relative to other types of chronic pain or other mental health conditions that may share similar brain features (eg, depression).

5.2. Criterion 2

Through stability analysis and variable importance in projection scores, Labus et al. tried to obtain an interpretable classification model and discussed the brain findings based on the literature. However, we still need more evidence to fully understand the roles of these brain structures in IBS or chronic pain broadly and to know which features are related to pain vs other comorbid risk factors. Converging evidence from different approaches (eg, functional magnetic resonance imaging, animal models) and across patient groups will be helpful.

5.3. Criterion 3

Labus et al. developed their model on 160 participants, and then they applied the a priori model on new holdout test data. They also reported precise model weights. These are strong features of the study. The research community could facilitate deployment across laboratories and patient groups using new innovations in technology. For example, Labus et al. could provide an online platform where researchers can upload structural images that they want to test the signature on, and signature scores can be sent to the researchers. In this way, we can maximize ease of deployment and minimize the number of choices researchers can make in applying the signature of Labus et al. to their data.

5.4. Criterion 4

Like the majority of studies, Labus et al. used data only from one laboratory and one scanner. However, importantly, they used data obtained from 6 different acquisition sequences, which could help generalize their findings across different sequences. They included only female participants in this study, so the results cannot be generalized to males and/or different types of visceral pain disorders yet. Therefore, the next step could include testing their a priori signature on new data from different laboratories and different scanners, and also on male participants and patients with other chronic pain conditions (including other types of chronic *visceral* pain and other types of chronic pain, such as chronic low back pain).

6. Conclusions

Labus et al.[16] took an exciting step toward a neuroimaging biomarker for IBS, and more broadly, chronic visceral pain. Considering the study by Labus et al. as a starting point, collaborative multisite efforts will help facilitate the biomarker development process, particularly focusing on the criteria above. We believe that the Pain and Interoception Imaging Network (*PAIN*®; painrepository.org)[15] repository will provide great resources for the biomarker discovery and validation process.

Conflict of interest statement

The authors have no conflicts of interest to declare.

This work was funded by NIDA R01DA035484-01 (T.D.W., PI).

Article history:
Received 30 April 2015
Accepted 1 May 2015
Available online 12 May 2015

References

[1] Apkarian AV, Baliki MN, Geha PY. Towards a theory of chronic pain. Prog Neurobiol 2009;87:81–97.
[2] Bagarinao E, Johnson KA, Martucci KT, Ichesco E, Farmer MA, Labus J, Ness TJ, Harris R, Deutsch G, Apkarian AV, Mayer EA, Clauw DJ, Mackey S. Preliminary structural MRI based brain classification of chronic pelvic pain: a MAPP network study. PAIN 2014;155:2502–9.
[3] Baliki MN, Schnitzer TJ, Bauer WR, Apkarian AV. Brain morphological signatures for chronic pain. PLoS One 2011;6:e26010.
[4] Bora E, Fornito A, Pantelis C, Yucel M. Gray matter abnormalities in major depressive disorder: a meta-analysis of voxel based morphometry studies. J Affect Disord 2012;138:9–18.
[5] Borsook D, Becerra L, Hargreaves R. Biomarkers for chronic pain and analgesia. Part 1: the need, reality, challenges, and solutions. Discov Med 2011;11:197–207.
[6] Borsook D, Becerra L, Hargreaves R. Biomarkers for chronic pain and analgesia. Part 2: how, where, and what to look for using functional imaging. Discov Med 2011;11:209–19.
[7] Callan D, Mills L, Nott C, England R, England S. A tool for classifying individuals with chronic back pain: using multivariate pattern analysis with functional magnetic resonance imaging data. PLoS One 2014;9: e98007.
[8] Duff EP, Vennart W, Wise RG, Howard MA, Harris RE, Lee M, Wartolowska K, Wanigasekera V, Wilson FJ, Whitlock M, Tracey I, Woolrich MW, Smith SM. Learning to identify CNS drug action and efficacy using multistudy fMRI data. Sci Transl Med 2015;7:274ra216.
[9] Eloyan A, Muschelli J, Nebel MB, Liu H, Han F, Zhao T, Barber AD, Joel S, Pekar JJ, Mostofsky SH, Caffo B. Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging. Front Syst Neurosci 2012;6:61.
[10] Ganzel BL, Kim P, Glover GH, Temple E. Resilience after 9/11: multimodal neuroimaging evidence for stress-related change in the healthy adult brain. Neuroimage 2008;40:788–95.
[11] Geha PY, Baliki MN, Harden RN, Bauer WR, Parrish TB, Apkarian AV. The brain in chronic CRPS pain: abnormal gray-white matter interactions in emotional and autonomic regions. Neuron 2008;60:570–81.
[12] Gianaros PJ, Jennings JR, Sheu LK, Greer PJ, Kuller LH, Matthews KA. Prospective reports of chronic life stress predict decreased grey matter volume in the hippocampus. Neuroimage 2007;35:795–803.
[13] Kelly PA, Viding E, Wallace GL, Schaer M, De Brito SA, Robustelli B, McCrory EJ. Cortical thickness, surface area, and gyrification abnormalities in children exposed to maltreatment: neural markers of vulnerability? Biol Psychiatry 2013;74:845–52.
[14] Klit H, Finnerup NB, Jensen TS. Central post-stroke pain: clinical characteristics, pathophysiology, and management. Lancet Neurol 2009;8:857–68.
[15] Labus JS, Naliboff B, Kilpatrick L, Liu C, Ashe-McNalley C, Dos Santos IR, Alaverdyan M, Woodworth D, Gupta A, Ellingson BM, Tillisch K, Mayer EA. Pain and Interoception Imaging Network (PAIN): a multimodal, multisite, brain-imaging repository for chronic somatic and visceral pain disorders. Neuroimage 2015. Epub ahead of print DOI: 10.1016/j.neuroimage.2015.04.018.
[16] Labus JS, Van Horn JD, Gupta A, Alaverdyan M, Torgerson C, Ashe-McNalley C, Irimia A, Hong JY, Naliboff B, Tillisch K, Mayer EA. Multivariate morphological brain signatures predict chronic abdominal pain patients from healthy control subjects. PAIN 2015;156:1545–54.
[17] Li L, Wu M, Liao Y, Ouyang L, Du M, Lei D, Chen L, Yao L, Huang X, Gong Q. Grey matter reduction associated with posttraumatic stress disorder and traumatic stress. Neurosci Biobehav Rev 2014;43:163–72.
[18] Luby J, Belden A, Botteron K, Marrus N, Harms MP, Babb C, Nishino T, Barch D. The effects of poverty on childhood brain development: the mediating effect of caregiving and stressful life events. JAMA Pediatr 2013;167:1135–42.
[19] May A. Chronic pain may change the structure of the brain. PAIN 2008; 137:7–15.
[20] McEwen BS, Gianaros PJ. Stress- and allostasis-induced brain plasticity. Annu Rev Med 2011;62:431–45.
[21] Morey RA, Gold AL, LaBar KS, Beall SK, Brown VM, Haswell CC, Nasser JD, Wagner HR, McCarthy G, Mid-Atlantic MW. Amygdala volume changes in posttraumatic stress disorder in a large case-controlled veterans group. Arch Gen Psychiatry 2012;69:1169–78.
[22] Pezawas L, Meyer-Lindenberg A, Drabant EM, Verchinski BA, Munoz KE, Kolachana BS, Egan MF, Mattay VS, Hariri AR, Weinberger DR. 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. Nat Neurosci 2005;8:828–34.
[23] Rao U, Chen LA, Bidesi AS, Shad MU, Thomas MA, Hammen CL. Hippocampal changes associated with early-life adversity and vulnerability to depression. Biol Psychiatry 2010;67:357–64.
[24] Redlich R, Almeida JJ, Grotegerd D, Opel N, Kugel H, Heindel W, Arolt V, Phillips ML, Dannlowski U. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. JAMA Psychiatry 2014;71:1222–30.
[25] Rosa MJ, Seymour B. Decoding the matrix: benefits and limitations of applying machine learning algorithms to pain neuroimaging. PAIN 2014; 155:864–7.
[26] Schwartz N, Temkin P, Jurado S, Lim BK, Heifets BD, Polepalli JS, Malenka RC. Chronic pain. Decreased motivation during chronic pain requires long-term depression in the nucleus accumbens. Science 2014; 345:535–42.
[27] Sekiguchi A, Sugiura M, Taki Y, Kotozaki Y, Nouchi R, Takeuchi H, Araki T, Hanawa S, Nakagawa S, Miyauchi CM, Sakuma A, Kawashima R. Brain structural changes as vulnerability factors and acquired signs of post-earthquake stress. Mol Psychiatry 2013;18:618–23.
[28] Smallwood RF, Laird AR, Ramage AE, Parkinson AL, Lewis J, Clauw DJ, Williams DA, Schmidt-Wilcke T, Farrell MJ, Eickhoff SB, Robin DA. Structural brain anomalies and chronic pain: a quantitative meta-analysis of gray matter volume. J Pain 2013;14:663–75.
[29] Ung H, Brown JE, Johnson KA, Younger J, Hush J, Mackey S. Multivariate classification of structural MRI data detects chronic low back pain. Cereb Cortex 2014;24:1037–44.
[30] Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. N Engl J Med 2013;368: 1388–97.
[31] Yap MB, Whittle S, Yucel M, Sheeber L, Pantelis C, Simmons JG, Allen NB. Interaction of parenting experiences and brain structure in the prediction of depressive symptoms in adolescents. Arch Gen Psychiatry 2008;65:1377–85.
[32] Yarnitsky D. Conditioned pain modulation (the diffuse noxious inhibitory control-like effect): its relevance for acute and chronic pain states. Curr Opin Anaesthesiol 2010;23:611–15.
[33] Zhang S, Seymour B. Technology for chronic pain. Curr Biol 2014;24: R930–935.