NeuroImage: Clinical 20 (2018) 407–414

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Exploring the prediction of emotional valence and pharmacologic effect across fMRI studies of antidepressants

Daniel S. Barron[a,b,*], Mehraveh Salehi[c,d], Michael Browning[e,f], Catherine J. Harmer[e,f], R. Todd Constable[g,h,i], Eugene Duff[j,k]

[a] Yale University School of Medicine, New Haven, CT, USA
[b] Yale University Department of Psychiatry, New Haven, CT, USA
[c] Department of Electrical Engineering, Yale University, New Haven, CT, USA
[d] Yale Institute for Network Science, Yale University, New Haven, CT, USA
[e] Oxford University Department of Psychiatry, Oxford, United Kingdom
[f] Oxford Health NHS Trust, Oxford, UK
[g] Interdepartmental Neuroscience Program, Yale University School of Medicine, New Haven, CT, USA
[h] Department of Radiology and Biomedical Imaging, Yale University School of Medicine, New Haven, CT, USA
[i] Department of Neurosurgery, Yale University School of Medicine, New Haven, CT, USA
[j] Functional Magnetic Resonance Imaging of the Brain Centre, Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
[k] Department of Paediatrics, University of Oxford, UK

ARTICLE INFO

Keywords:
Antidepressant
Emotional valence
Machine learning
Drug development
Predictive analysis
Task-based fMRI

ABSTRACT

Background: Clinically approved antidepressants modulate the brain's emotional valence circuits, suggesting that the response of these circuits could serve as a biomarker for screening candidate antidepressant drugs. However, it is necessary that these modulations can be reliably detected. Here, we apply a cross-validated predictive model to classify emotional valence and pharmacologic effect across eleven task-based fMRI datasets (n = 306) exploring the effect of antidepressant administration on emotional face processing.

Methods: We created subject-level contrast of parameter estimates of the emotional faces task and used the Shen whole-brain parcellation scheme to define 268 subject-level features that trained a cross-validated gradient-boosting machine protocol to classify emotional valence (fearful vs happy face visual conditions) and pharmacologic effect (drug vs placebo administration) within and across studies.

Results: We found patterns of brain activity that classify emotional valence with a statistically significant level of accuracy (70% across-all-subjects; range from 50 to 87% across-study). Our classifier failed to consistently discriminate drug from placebo. Subject population (healthy or unhealthy), treatment group (drug or placebo), and drug administration protocol (dose and duration) affected this accuracy with similar populations better predicting one another.

Conclusions: We found limited evidence that antidepressants modulated brain response in a consistent manner, however found a consistent signature for emotional valence. Variable functional patterns across studies suggest that predictive modeling can inform biomarker development in mental health and in pharmacotherapy development. Our results suggest that case-controlled designs and more standardized protocols are required for functional imaging to provide robust biomarkers for drug development.

1. Introduction

Psychiatric drug development is difficult, expensive, and beset by a high failure rate. The slow onset, unclear biological markers, and variable clinical efficacy even of approved psychiatric drugs makes the potential efficacy of candidate drugs difficult to measure and has led many pharmaceutical companies to withdraw from drug development (Insel et al., 2012; Friedman, 2013). Biomarkers that capture how effective drugs modulate the brain's functional anatomy could prioritize candidate compounds for large clinical trials, thus improving the productivity and cost-effectiveness of drug development.

Clinically approved antidepressants modulate the brain's emotional

* Corresponding author at: 300 George Suite #901, New Haven, CT 06510, USA.
E-mail address: daniel.s.barron@yale.edu (D.S. Barron).

https://doi.org/10.1016/j.nicl.2018.08.016
Received 12 April 2018; Received in revised form 18 July 2018; Accepted 9 August 2018
Available online 11 August 2018

valence circuits, suggesting that the response of these cicruits could serve as a biomarker for screening candidate antidepressant drugs. The emotional faces task has been particularly useful in eliciting the emotional valence circuit (Ekman, 2013; Leppänen, 2006). In this task, a subject is instructed to view a human actors' face and determine the gender of or the emotion expressed. Independent studies have shown that emotional valence networks engaged by this task are affected by antidepressant administration (Murphy et al., 2009). The applicability of these studies to screen for potential antidepressant compounds rests on the ability of the emotional faces task to engage a spatially consistent emotional valence network across populations, specifically the aspect of this network that is affected by antidepressant administration. This applicability may be explored by assessing 2 contrasts: an emotional valence contrast (i.e. is there a consistent difference in activity when positive and negative faces are displayed?) and a pharmacologic contrast (is there a consistent difference when antidepressants are compared to placebos?).

A second advantage of the emotional valence contrast described above is that it can be constructed in either a within or between-subject manner. Duff et al. (Duff et al., 2015) have previously successfully developed a cross-validated machine learning protocol which was able to predict pharmacologic class in analgesic studies within pain stimulation tasks. However, the analgesia literature tends to use within subject designs whereas the antidepressant literature uses between subject designs. The emotional valence contrast is therefore useful as a means of directly comparing classifier performance of within vs. between subject contrasts on the same dataset.

Here, we apply a machine-learning classifier to a large set of studies of antidepressant effects on brain responses during an emotional faces tasks. We explore the consistency of the emotional valance effect considered both within and between-subjects and the between-subject pharmacologic effect. Because these studies use protocols with considerable variability in scanners, experimental tasks and patient cohorts, we further aim to explore the effect of protocol variability on signature generalizability. To accomplish this, we exploit a dimensionality reduction step (Yoshida et al., 2017) to reduce voxel-wise data to functionally homogenous parcels defined in an independent dataset by an unsupervised algorithm (Shen et al., 2013). We then apply the gradient boosted machine (GBM) classifier to predict emotional valence (fearful vs happy face presentation) and pharmacologic class (antidepressant versus placebo), to test whether a consistent, cross-study signature may be identified, and to understand which study protocols generate a more generalizable signature.

2. Methods and materials

For each of eleven datasets, subject-level contrast of parameter estimates of the emotional faces task were created and divided into 268 regions using the Shen whole-brain parcellation scheme. Each region was used as a feature within a cross-validated gradient-boosting machine protocol that classified emotional valence and pharmacologic effect within and across studies. Feature weightings were then mapped onto the brain to allow anatomic localization and visualization.

2.1. Datasets

Eleven independent datasets from eight task-based fMRI studies of the effect of antidepressant administration on emotional face processing were available for analysis, representing 306 subjects (See Table 1 for key features of the dataset; NB: the number of subjects per study differs from the original publications, reflecting that some data could not be located for inclusion in our study and that one study (Warren) has recruited more participants since the time of our study). These studies were all performed in the Harmer lab from 2006 to 2015 and made use of healthy subjects (H) without previous history of mental illness and subjects selected based on the presence of symptoms consistent with a disorder (i.e. Major Depressive Disorder) or symptom (i.e. neuroticism or dysphoria). In these studies, the Beck Depression Inventory and the Eysenck Personality Questionnaire, neuroticism dimension were used to assess these symptoms. Although specific aspects of the study varied (e.g. antidepressant dose and duration), all versions investigated group differences in whole-brain BOLD response when subjects viewed happy and fearful faces. In this study, we selected only happy and fearful emotional face presentation, as these were the most consistently used emotions in our available dataset. Individual studies each obtained ethical approval from the local ethics committee.

2.2. MRI processing

Standard preprocessing and mapping analysis were employed using tools from FMRIB's Software Library (FSL) package (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). The FSL FMRI Expert Analysis Tool (FEAT) was used for general linear modeling (GLM) (Jenkinson et al., 2012). Subject-level contrast of parameter estimate (COPE) maps for each contrast (e.g. happy versus fixation) were produced in native patient space. These COPE maps were used in subsequent classification analyses, as described below. See Supplementary Methods for more details and Fig. 1 for an illustration of the analysis pipeline.

2.3. Machine learning method

Cognitive models of depression suggest that patients process negative relative to positive stimuli differently from non patients, and that these cognitive processes are causative in the illness. Therefore a contrast looking at the emotional processing circuit activation to negative vs. positive faces may be able to identify illness specific signatures and how the brain's emotional circuits change in response to treatment. We chose a forced-choice gradient boosting machine (GBM) for classification due to its robustness to outliers and its ability to map features back into anatomical brain space (Friedman, n.d.).

Predictive analyses are prone to overfitting when the number of features far outweighs the number of subjects (Yoshida et al., 2017). Given our available dataset of 306 subjects, we had to reduce the number of features from voxels (~900,000 in 2 mm isotropic space). To this end, we selected the Shen 268-node resting-state fMRI atlas, defined by a group-wise spectral clustering algorithm applied to an independent dataset consisting of 45 subjects (Shen et al., 2013; Finn et al., 2015). We transformed the Shen atlas from MNI-152 space into native patient space using linear and nonlinear FSL transforms (Jenkinson et al., 2012) and used the average COPE values within each parcel to produce 268 features per subject for the classifier.

2.4. We trained 2 overall types of classifiers

1) Emotional Valence Classifier. This analysis determined whether and where a signal for emotional valence was consistent enough to discriminate fear from happy face visual conditions. We assessed the performance of the emotional valence classifier with two different types of feature inputs to determine the impact of inter-subject variability and task variability. The first subtracted fear and happy responses within-subject, to account for average differences in visual responses across subjects (i.e. the classifier compared the FvH COPE contrast image to the HvF COPE contrast image). The second compared fear versus fixation COPE files and happy versus fixation COPEs and accounted for across-study differences in task, without being able to minimize individual subject variability in the visual response. Duff et al. (Duff et al., 2015) were able to minimize inter-subject variability through within-subject contrasts wherein each subject received a placebo and drug condition, thus allowing pharmacologic effect to be isolated from variability due to individual differences and/or task. Because the pharmacologic effect in our studies was necessarily between subjects, we used the valence