OPEN ACCESS

Original research

Brain activity measured by functional brain imaging predicts breathlessness improvement during pulmonary rehabilitation

Sarah L Finnegan ,[1] Michael Browning,[2,3] Eugene Duff,[4] Catherine J Harmer,[2,3] Andrea Reinecke ,[2,3] Najib M Rahman,[5,6,7] Kyle T S Pattinson [1,8]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/thorax-2022-218754).

[1]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK
[2]Oxford Health NHS Foundation Trust, Warneford Hospital, NHS, Oxford, UK
[3]Department of Psychiatry, University of Oxford, Oxford, UK
[4]Department of Paediatrics, University of Oxford, Oxford, UK
[5]Oxford Respiratory Trials Unit, University of Oxford, Oxford, UK
[6]Oxford Chinese Academy of Medicine Institute, University of Oxford, Oxford, UK
[7]Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK
[8]Nuffield Division of Anaesthetics, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

Correspondence to
Dr Sarah L Finnegan, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX1 2JD, UK; sarah.finnegan@ndcn.ox.ac.uk

Received 26 January 2022
Accepted 20 November 2022

To cite: Finnegan SL, Browning M, Duff E, *et al*. *Thorax* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/thoraxjnl-2022-218754

ABSTRACT

Background Chronic breathlessness in chronic obstructive pulmonary disease (COPD) is effectively treated with pulmonary rehabilitation. However, baseline patient characteristics predicting improvements in breathlessness are unknown. This knowledge may provide better understanding of the mechanisms engaged in treating breathlessness and help to individualise therapy. Increasing evidence supports the role of expectation (ie, placebo and nocebo effects) in breathlessness perception. In this study, we tested functional brain imaging markers of breathlessness expectation as predictors of therapeutic response to pulmonary rehabilitation, and asked whether D-cycloserine, a brain-active drug known to influence expectation mechanisms, modulated any predictive model.

Methods Data from 71 participants with mild-to-moderate COPD recruited to a randomised double-blind controlled experimental medicine study of D-cycloserine given during pulmonary rehabilitation were analysed (ID: NCT01985750). Baseline variables, including brain-activity, self-report questionnaires responses, clinical measures of respiratory function and drug allocation were used to train machine-learning models to predict the outcome, a minimally clinically relevant change in the Dyspnoea-12 score.

Results Only models that included brain imaging markers of breathlessness-expectation successfully predicted improvements in Dyspnoea-12 score (sensitivity 0.88, specificity 0.77). D-cycloserine was independently associated with breathlessness improvement. Models that included only questionnaires and clinical measures did not predict outcome (sensitivity 0.68, specificity 0.2).

Conclusions Brain activity to breathlessness related cues is a strong predictor of clinical improvement in breathlessness over pulmonary rehabilitation. This implies that expectation is key in breathlessness perception. Manipulation of the brain's expectation pathways (either pharmacological or non-pharmacological) therefore merits further testing in the treatment of chronic breathlessness.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Pulmonary rehabilitation is an effective treatment for many, but not all people with chronic obstructive pulmonary disease (COPD who suffer from chronic breathlessness in COPD. Baseline patient characteristics predicting improvements in breathlessness are unknown.

WHAT THIS STUDY ADDS

⇒ This is the first study to identify a model capable of predicting changes in breathlessness over pulmonary rehabilitation at the individual patient level. The study shows that prerehabilitation breathlessness expectation related brain activity is a strong predictor of clinical improvement in breathlessness over pulmonary rehabilitation.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Manipulation of the brain's expectation pathways (either pharmacological or non-pharmacological) merits further testing in the treatment of chronic breathlessness.

INTRODUCTION

Chronic breathlessness is a key feature of chronic obstructive pulmonary disease (COPD) with symptoms often persisting despite maximal medical therapy. Pulmonary rehabilitation is the best treatment for chronic breathlessness in COPD[1] but the response is variable. Thirty percent of people who complete pulmonary rehabilitation derive no clinical benefit.[2] Despite considerable research, we still do not know which patient characteristics predict beneficial response to pulmonary rehabilitation.[2–5] The ability to predict outcome has a number of potential benefits. These include improving our understanding of underlying mechanisms, identifying targets for personalised medicine which may allow more accurate allocation of scarce healthcare resources.

Breathlessness severity is often poorly explained by objective clinical measures.[6] This has prompted research into identifying brain perceptual mechanisms that may explain this discordance. A body of work has recently identified that brain processes relating to expectation (akin to placebo and nocebo effects) have an important role in contributing to breathlessness severity. Whether brain-derived metrics may help predict outcome from pulmonary rehabilitation is unknown, and prediction models until now have not included measures of expectation.

Between-subject variability in therapeutic response is increasingly recognised as a confounder in clinical

trials. A personalised medicine approach aims to identify subgroups of patients that respond to a specific therapy. In psychiatry, brain-derived metrics using functional neuroimaging have taken similar approaches to identifying subtypes of depression that may respond to bespoke therapies.[7] Such techniques rely on biomarkers, which may consist of predictive combinations of biochemical, genetic, demographic, physiological or cognitive measures. In the context of treating breathlessness, predictive biomarkers could pave the way for novel pharmacological and non-pharmacological treatments. These may either work as stand-alone therapies, or by enhancing other therapies, such as pulmonary rehabilitation.

In this study, we aimed to predict improvements in breathlessness during pulmonary rehabilitation by analysing baseline data from a longitudinal experimental medicine study of D-cycloserine on breathlessness during pulmonary rehabilitation. We selected D-cycloserine, which is a partial agonist at the NMDA receptor in the brain, for its action on neural plasticity and influence on brain expectation mechanisms associated with cognitive behavioural therapies.[8–10] Brain-based pharmacological adjuncts may be one opportunity to boost the effects of pulmonary rehabilitation. We hypothesised baseline brain activity in response to breathlessness-related expectation would predict improvement in breathlessness over pulmonary rehabilitation, and that if D-cycloserine indeed had an effect on expectation then it would emerge as a significant factor in the prediction model. Given that moderators of treatment success of pharmacological agents such as D-cycloserine remain unclear, this information will help build a better picture of the brain-behaviour changes that may underly response to pulmonary rehabilitation and therefore clarify its potential value as a therapeutic target.

METHODS AND MATERIALS

A brief overview of materials and methods is presented here with full details included within online supplemental materials. The study and statistical analysis plan were preregistered on bioXIV (https://osf.io/bfqds/). This was an analysis of data from a longitudinal experimental medicine study of patients with COPD over a course of pulmonary rehabilitation. Parts of the study were first published in a characterisation of baseline patient clusters[11] and subsequently in the investigation of the effect of D-cycloserine on brain activity (preprint https://doi.org/10.1101/2021.06.24.21259306). The analysis conducted for this study is novel, not previously reported and is the first use of predictive analysis using this dataset.

Participants

Seventy-one participants (18 female, median age 71 years (46–85 years)) (online supplemental table 1) were recruited immediately prior to enrolment in a National Health Service-prescribed course of pulmonary rehabilitation. Full demographic details are included within online supplemental materials and are published separately (preprint is available at https://doi.org/10.1101/2021.06.24.21259306).

Study protocol

Data for this analysis were acquired at baseline assessment held at the start of a pulmonary rehabilitation course, and following completion of the pulmonary rehabilitation at 6–8 weeks. At each study visit, identical measures were collected. Following the first visit, participants were randomised in a double-blind procedure to receive either 250 mg oral D-cycloserine or a matched placebo. Participants received a single dose on four occasions 30 min prior to the onset of the first four pulmonary rehabilitation sessions.

Table 1 List of measures included within each of the three models (indicated by 'X')

Included data	Brain only model	Brain and non-imaging measures model	Non-imaging measure model
Drug ID	x	x	x
Responder or non-responder label	x	x	x
Brain activity			
Amygdala	x	x	
Hippocampus	x	x	
Anterior insula	x	x	
Anterior cingulate	x	x	
Posterior insula	x	x	
Putamen	x	x	
Superior marginal gyrus	x	x	
Superior frontal gyrus	x	x	
Precuneus	x	x	
Medial prefrontal cortex	x	x	
Caudate	x	x	
Posterior cingulate	x	x	
Angular gyrus	x	x	
Precentral gyrus	x	x	
Middle frontal gyrus	x	x	
Questionnaires			
D12		x	x
CES-D		x	x
TRAIT		x	x
FSS		x	x
SGRQ		x	x
MRC		x	x
BCS		x	x
Vigilance		x	x
Physiology			
FEV1/FVC		x	x
MSWT—HR change		x	x
MSWT—SPO2 change		x	x
MSWT—distance		x	x
MSWT—BORG change		x	x
BMI		x	x
Pack-years		x	x
Age		x	x
Sex		x	x

Drug ID labels corresponded to whether the participant received D-cycloserine or placebo.

BCS, breathlessness catastrophising scale; BMI, body mass index; BORG, rating of perceived exertion; CES-D, Centre for Epidemiologic Studies Depression Scale; D12, Dyspnoea-12; FEV1, forced expiratory volume; FSS, Fatigue Severity Scale; FVC, forced vital capacity; HR, heart rate; MRC, Medical Research Council; MSWT, Modified Shuttle Walk Test; SGRQ, St George's Respiratory Questionnaire; SPO2, oxygen saturation; TRAIT, Trait Anxiety Inventory.

Self-report questionnaires

All questionnaires (table 1) were scored according to respective manuals: Dyspnoea-12 (D12) Questionnaire,[12] Centre for Epidemiologic Studies Depression Scale,[13] Trait Anxiety Inventory,[14] Fatigue Severity Scale,[15] St George's Respiratory Questionnaire,[16] Medical Research Council (MRC) breathlessness

scale,[17] Breathlessness catastrophising scale, adapted from the Catastrophic Thinking Scale in Asthma,[18] Breathlessness vigilance, adapted from the Pain Awareness and Vigilance Scale Breathlessness Awareness and Vigilance Scale.[19]

Physiological measures

Spirometry and two Modified Shuttle Walk Tests (MSWT) were collected using standard protocols.[20 21] Participant height and weight were recorded at each visit. Arterial oxygen saturations were collected at rest and following the MSWT.

MRI measures

Image acquisition

MRI of the brain was carried out using a Siemens 3T MAGNETOM Trio. A T1-weighted (MPRAGE) structural scan (voxel size: 1×1 × 1 mm) was collected and used for registration purposes. A T2*-weighted, gradient echo planar image (EPI) scan sequence (TR, 3000 ms; TE 30 ms; voxel size: 3×3×3 mm) was used to collect functional imaging data during the word cue task.

Word cue task

Given sufficient fearful breathlessness exposures, the suggestion alone of the situation can be sufficient to drive a top-down neural cascade and produce breathlessness in the absence of afferent inputs. We drew on this link to probe the neural responses of breathlessness-related expectation by examining the activity of brain regions responding to breathlessness-related word cues.[22 23] Brain activity was correlated with corresponding visual analogue ratings of breathlessness and breathlessness-anxiety collected during scanning. During the fMRI scanning, participants were presented with a word cue, for example, 'climbing stairs' in white text on a black background for 7 s. Participants were then asked, 'how breathless would this make you feel' (wB) and 'how anxious would this make you feel' (wA). To each question participants responded within a 7 s window using a button box and Visual Analogue Scale. The response marker always initially appeared at the centre of the scale, with the anchors 'not at all' and 'very much' at either end. Scan duration was 7 min and 33 s.

ANALYSIS

Regions of interest

Fifteen regions of interest were selected a priori (figure 1), encompassing regions associated with sensory and affective processing of breathlessness as well as body, symptom perception and emotional salience.[22 24 25] Regions were defined by standard anatomical atlas maps (Harvard-Oxford Atlas and Destrieux' cortical atlas), thresholded at 40% probability and binarised.

Brain imaging analysis

Image processing was carried out using the Oxford Centre for Functional MRI of Brain Software Library (FMRIB, Oxford, UK; FSL V.5.0.8; https://ww.fmrib.ox.ac.uk/fsl/), MATLAB R2018b (Mathworks, Natick, Massachusetts, USA) and associated custom scripts. Functional MRI processing was performed using FEAT (FMRI Expert Analysis Tool, within the FSL package).

Preprocessing and single subject models

Data were preprocessed according to standard protocols which included motion correction and physiological noise removal, before being entered into single subject general linear models. These models captured brain activity during the periods in which the breathlessness-related word cues were presented allowing us to examine expectation-related processes (online supplemental figure 3). Further details regarding preprocessing and model specifics can be found within online supplemental materials. For each participant, the mean signal in response to the breathlessness-related word cues was extracted for each brain region (figure 1). This gave each participant 15 brain-derived scores to enter into the predictive models.

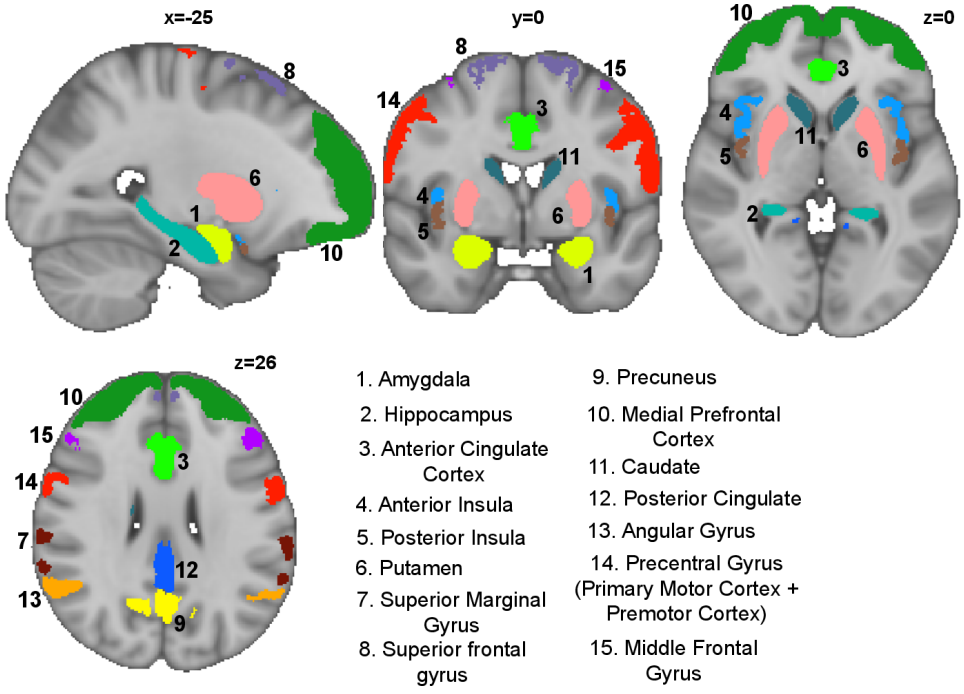


Figure 1 Region of interest map showing 15 brain areas.

Table 2 Logistic regression coefficients for predictive power of the computationally derived brain-behaviour model (model prediction labels), and baseline D12 on pulmonary rehabilitation outcome, measured as a change in D12 score above the minimal clinical important difference

	Coefficient	P value
Intercept	3.571	0.007
Model prediction labels	−2.809	<0.001
Baseline D12	0.071	0.191

Significance is expressed as false discovery rate (FDR)—corrected p values.
D12, Dyspnoea-12.

Definition of response to pulmonary rehabilitation

Responsiveness to pulmonary rehabilitation was defined as a change in D12 score, a well-validated clinical measure of breathlessness, of three or more points, consistent with the minimal clinically important difference .[12] To examine whether baseline D12 score differed significantly between responders and non-responders to pulmonary rehabilitation we conducted a comparison of average baseline D12 score. In addition, a single logistic regression model (table 2) was applied to the data using MATLAB's mnrfit function to examine whether baseline D12 explained pulmonary rehabilitation outcome over and above the best model prediction. Significance was set as false discovery rate corrected $p<0.05$.

Predictive models

Physiological measures

Spirometry and two MSWT were collected using standard protocols.[20 21] Participant height and weight were recorded at each visit. Sex was self-reported. Arterial oxygen saturations were collected at rest and following the MSWT.

Models were programmed using R V.3.6.1 (2019-07-05). Modelling procedure remained the same for each of the three models (figure 2).

1. All measures were centred and scaled. Checks were performed to determine whether any measures were highly correlated (R>0.8) or linear combinations of each other.
2. To correct for imbalance in the number of responders/non-responder a resampling procedure. Imbalanced classes can affect classifier performance. Random OverSample Examples (ROSE) was carried out. ROSE, an R package, creates an artificially balanced sample using a smoothed bootstrap approach.[26]
3. An elastic net procedure was used to identify the number most relevant features for inclusion into the model. Elastic net procedure was selected for its ability to regularise, improve data sparsity via feature selection and cluster correlated measures together (for more details see online supplemental materials). Features were selected based on ranked coefficients.
4. Model training parameters—C, kappa and sigma were selected based on an internal repeated cross-validation procedure (10-fold cross validation repeated 3 times). In all instances automated tuning parameter selection for the values, with a tune length of 5, was used within R's caret package. Train test data were kept separate across folds, with the algorithm never having access to the entire dataset. The best tuning parameters were selected automatically by R's caret package from across cross validation folds.
5. These parameters were used to train a Support Vector Classifier with radial kernel to predict outcomes in the entire dataset.
6. Model performance was assessed internally using accuracy, sensitivity, specificity and area under the curve. Full confusion matrices are presented along with calibration curves. Model significance was assessed with a one-tailed binomial test of model accuracy compared with the null information rate.

RESULTS

Participants

At baseline the median MRC breathlessness score of the 71 participants was 3 (IQR 1), median FEV1/FVC (forced expiratory volume/ forced vital capacity) was 0.55 (IQR 0.15), median FEV1% predicted was 58 (IQR 21).

Responders and non-responders

A total of 41/71 participants in the primary dataset met the criteria of a change in D12 score of three or more points to be considered a responder[12] (24 D-cycloserine, 17 placebo), and 30 did not (13 D-cycloserine, 17 placebo). No significant interaction between responders and non-responders and drug was identified using χ^2 analysis ((1, N=71) = 1.6, p=0.21). Group changes to D12, St Georges and MSWT are presented in table 3 and online supplemental table 3.

Feature selection: brain imaging only model

The elastic net procedure identified 13 of 15 brain-derived metrics and drug as relevant for model inclusion (table 1). These features were: expectation-related brain activity within amygdala, caudate, prefrontal cortex, hippocampus, superior frontal gyrus, anterior insula, drug, posterior cingulate cortex, putamen, posterior insula, middle frontal gyrus, precuneus, precentral gyrus and angular gyrus.

Feature selection: brain and non-imaging measure model

The elastic net procedure identified 12 of 15 brain-derived metrics, 13 of 20 non-imaging measures including drug as relevant for model inclusion (table 1). These features were brain activity within: superior frontal gyrus, hippocampus, angular gyrus, superior marginal gyrus, amygdala, prefrontal cortex, precuneus, anterior cingulate cortex, anterior insula, middle frontal gyrus, posterior insula and putamen. Behavioural features identified as relevant for model inclusion were D12, anxiety, depression, MRC, the three St George's domains (active, impact, symptoms), MWST BORG, heart rate and SpO_2 change, fatigue, age and body mass index.

Feature selection: non-imaging measures model

Of the 20 questionnaire and physiological features available, only D12 survived the feature selection process (table 1).

Model results: internal validation

Three models with variables selected by the elastic net procedure were assessed for their ability to discriminate responders from non-responders (table 4).

The combination of brain and behaviour metrics produced the best classification performance (accuracy—0.83 (95% CI 0.75 to 0.90); sensitivity—0.88; specificity—0.77; p<0.001) and was well calibrated (table 4, (online supplemental figure 6). Weighted variable importance was found to be similar across features, as demonstrated by the thickness of the lines in figure 3. The brain only model was

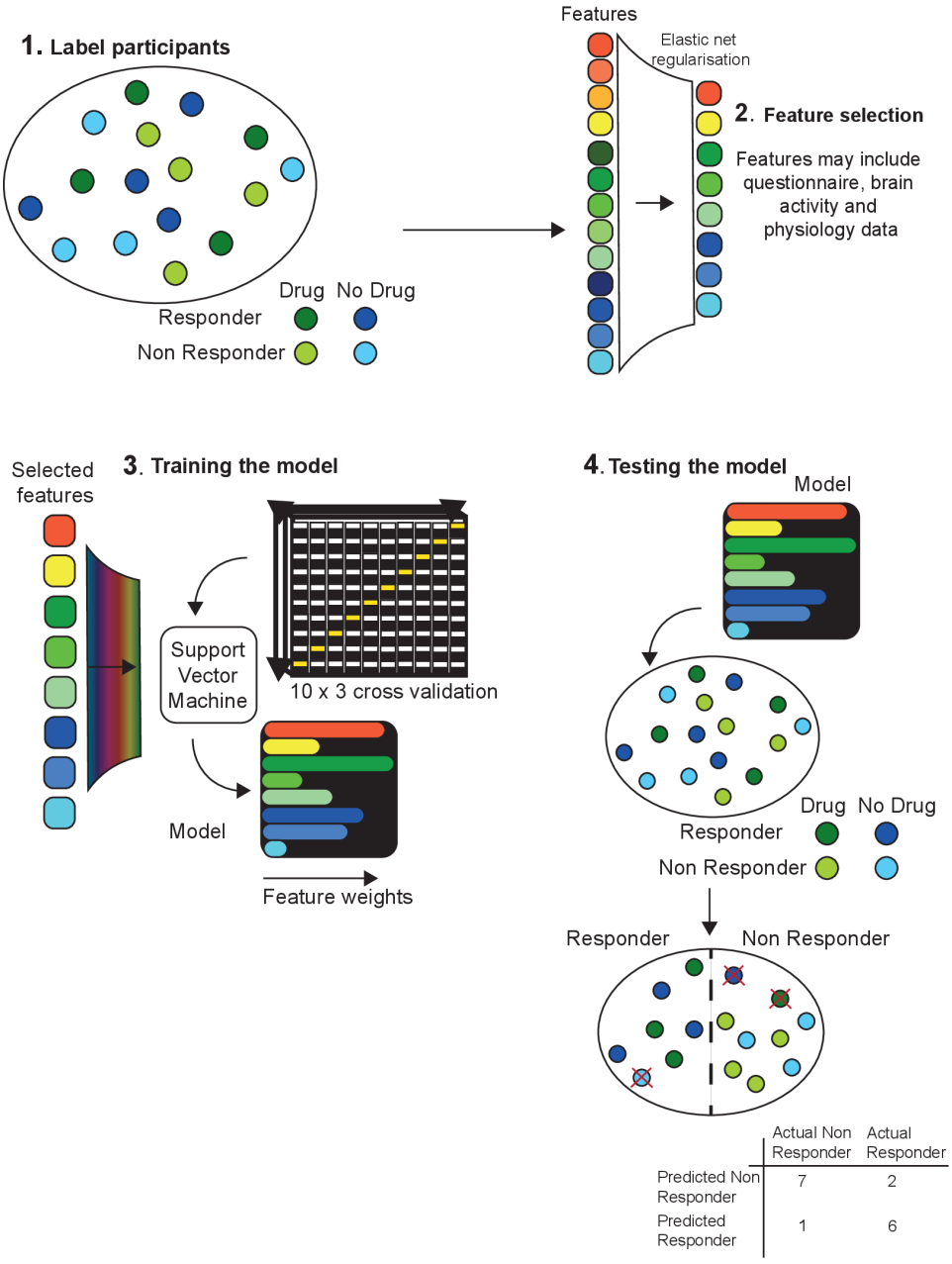


Figure 2 A schematic of the modelling procedure adapted from an illustration by Chekroud and colleagues.[37] (1) Participants received two labels, the first corresponding to drug or placebo and the second 'responder' or 'non-responder' to treatment. (2) An elastic net procedure was used to rank and select the top features. (3) Selected features were used to develop model training parameters in a repeated-cross validation procedure in which the algorithm never has access to the entire dataset. (4) The Trained SVM was then provided with the entire dataset to classify. In addition to model statistics, full confusion matrices were output to assess sensitivity and specificity.

able to correctly categorise participants with statistically significant likelihood (accuracy 0.70 (95% CI 0.58 to 0.81) but demonstrated poor goodness of fit (online supplemental figure 6).

DISCUSSION

Key findings

Using supervised machine learning, this study successfully identified markers that predict clinically relevant improvements in breathlessness over a course of pulmonary rehabilitation. The best model combined brain-imaging markers of breathlessness-expectation, self-report questionnaires and physiology measures, and demonstrated high sensitivity and specificity. Whether or not a participant received D-cycloserine was a significant feature in this model. Our findings demonstrate the first predictive model of change in breathlessness across pulmonary rehabilitation and, for the first time, the clinical relevance of expectation-related brain activity as a therapeutic target in the treatment of breathlessness.

To date, no study has produced a model capable of predicting an individual's change in breathlessness over pulmonary rehabilitation from baseline traits.[2–4] Although previous studies have shown correlations between baseline variables and outcomes,[5] none have attempted to predict outcomes at an individual level. This study is therefore the first to directly predict an individual's change in breathlessness over pulmonary rehabilitation. This was achieved using sensitive brain imaging techniques in order to capture personalised responses to breathlessness expectation which has, until recently remained relatively unexplored.

Table 3 Average scores on questionnaire and behavioural measures before and after pulmonary rehabilitation

	Comparison of scores before and after pulmonary rehabilitation			
	Before pulmonary rehabilitation	After pulmonary rehabilitation	Uncorrected p values	Corrected p values
D12	10.8 (10.5)	6.5 (7.2)	<0.001	<0.001
MSWT distance (m)	341 (260)	379 (300)	<0.001	0.001
St George—Active	62.5 (27.6)	57.0 (30.9)	<0.001	0.001
St George—Impact	31.1 (20.2)	25.7 (24.4)	<0.001	<0.001
St George—Symptom	61.4 (25.0)	54.9 (28.3)	<0.001*	<0.001*

Variance is expressed as IQR.
Significance is reported as exploratory uncorrected p values and as family wise error (*p<0.05) corrected values.
D12, Dyspnoea-12; MSWT, Modified Shuttle Walk Test.

Expectation has been linked with symptom severity across conditions including breathlessness and pain,[27 28] and is well recognised to underly the placebo and nocebo effects. An example of the nocebo effect in breathlessness is provided by a study of healthy volunteers in which, using a conditioning paradigm, a harmless odour was initially paired with induced breathlessness. Subsequently, the odour alone was shown to drive brain activity in the periaqueductal grey and anterior cingulate cortex leading to breathlessness despite the absence of afferent respiratory input.[28] In Abdallah *et al*,[29] expectation-related brain activity was associated with poorer responses to opioids in breathlessness, potentially explaining why clinical trials of opioids in the management of breathlessness have been unsuccessful.[30 31]

Fear and anxiety are key components of expectation, which recent research suggests may play a key role in the mechanisms and maintenance of breathlessness.[22 32 33] Despite this, expectation-related effects have not previously been considered in prediction studies of pulmonary rehabilitation outcome. Our previous work showed a clear correlation between expectation-related brain activity in areas that include the anterior insula, anterior cingulate cortex and prefrontal cortex, and improvements in breathlessness over pulmonary rehabilitation.[22] However, while these studies suggest baseline cognitive state may be a therapeutically relevant target, importantly, the methods employed so far did not attempt to predict the response of an individual. Taken together, converging lines of evidence point towards expectation-related processes as a clear potential therapeutic target.

In this study, we focused on brain activity changes within a set of pre-selected regions of interest associated with breathlessness-expectation and body and symptom perception.[22 24 25] In the original trial, we hypothesised that D-cycloserine would augment the therapeutic effects of pulmonary rehabilitation across this network, via its effects on neural plasticity and promotion of expectation-related learning.[10 34]

Using data-driven techniques, 13 of the 15 brain-derived metrics (and drug) were identified as relevant for model inclusion. Selected brain areas spanned the components of relevant body and symptom perception and emotional salience networks. The resulting brain-only model, while statistically significant (p=0.02) and possessing good sensitivity (0.93), did not distinguish responders from non-responders with sufficient specificity (0.40).

By, enriching the brain-only models with questionnaire and physiology measures improved performance considerably. In this enriched model, 12 brain-derived metrics and 13 non-imaging-derived metrics, which included self-report questionnaire measures, physiology and drug, were identified as relevant for model inclusion. Measures of accuracy (0.83), sensitivity (0.88) and specificity (0.77) all suggest this model was able to significantly (p<0.001) predict pulmonary rehabilitation outcome.

Within the non-imaging measure only model, D12 alone was selected by the elastic net and was not found to be significantly (p=0.09, sensitivity=0.68, specificity=0.20) predictive of pulmonary rehabilitation outcome. No other of the 13 non-imaging-derived metrics available was found to contribute to the model. That only D12 was selected suggests that the remaining measures, which were important predictors of rehabilitation outcome in the enriched model, interact strongly with brain activity. These results highlight the value of approaching breathlessness from a multimodal data perspective.

Table 4 Model statistics for brain imaging only, brain and non-imaging measure models and non-imaging measures only model

		Brain only full model		Brain and non-imaging measures full model		Non-imaging measures full model	
Accuracy		0.70		0.83		0.66	
95% CI		58 to 81		0.72 to 0.90		0.54 to 0.77	
Sensitivity		0.93		0.88		0.68	
Specificity		0.40		0.77		0.20	
AUC		0.79		0.87		0.70	
P value		0.02*		<0.001**		0.09	
		Reference		Reference		Reference	
Confusion matrices		Yes	No	Yes	No	Yes	No
Prediction	Yes	38	18	36	7	28	11
	No	3	12	5	23	13	19

All models contained drug ID as an additional term.
P value is expressed as the result of a one-tailed binomial test of model accuracy compared with the null information rate.
*p<0.05
AUC, area under the curve.

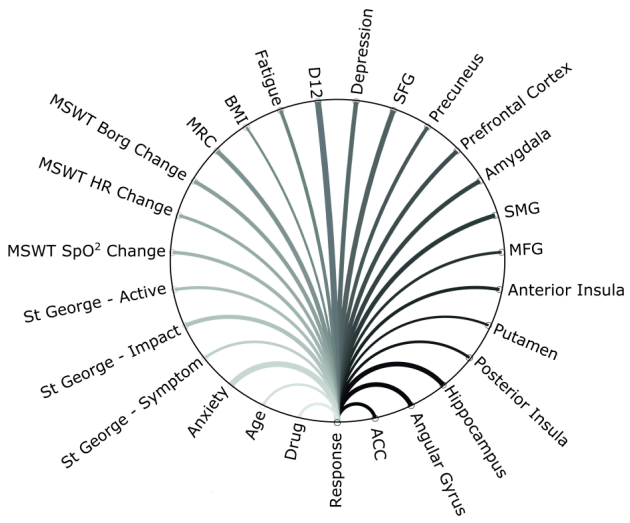


Figure 3 Schematic representation of the best predictive model. Predictive brain imaging and non-imaging measures are shown linked to treatment response by weighted lines, indicating variable importance. ACC, anterior cingulate cortex; BMI, body mass index; D12, Dyspnoea-12; HR, heart rate; MRC, Medical Research Council; MFG, middle frontal gyrus; MSWT, Modified Shuttle Walk Test; SMG, supramarginal gyrus; SpO2, oxygen saturation.

The retained brain activity features implicate a range of brain networks encompassing functions of cognitive control, symptom perception and sensory integration. Activity within these regions has been shown to predict outcome to cognitive behavioural therapy in social anxiety disorder[35] and obsessive–compulsive disorder.[36] In our paradigm, in which patients were shown breathlessness related word-cues, triggering expectation related processes, activity within cognitive control network may indicate the allocation of attentional resources. The retained questionnaires within the brain and non-imaging measure model together highlight another important domain of breathlessness: symptom perception. We suggest that baseline symptom perception may act to directly influence the interpretation of the new experiences of pulmonary rehabilitation as positive or negative. At scale, functional neuroimaging is not practicable as a prerehabilitation screening tool, both in terms of cost and access. We do not therefore propose that patients undergo functional neuroimaging as a diagnostic test. However, the use of functional neuroimaging in the research setting can enhance our understanding of how breathlessness may be targeted. For example, the retained brain activity features implicate the allocation of attentional resources, symptom perception and sensory integration. Thus, functional neuroimaging is able to provide insights into the relationship between brain activity and behaviour that is not possible with other techniques. Building on this information, new behavioural therapies may seek to specifically target these domains in parallel or prior to pulmonary rehabilitation.

D-cycloserine has been shown to augment changes to expectation, boosting the therapeutic effects in trials examining anxiety, post-traumatic stress disorder and other mental health conditions,[8–10] where acute dosages administered prior to exposure-based therapies appear to variously and significantly reduce self-report symptoms of acrophobia, improve clinical symptoms of panic disorder, reduce threat response to fearful faces and associated decision-making reaction times. As a drug which acts on expectation-related brain activity pathways it is therefore not surprising that whether a participant received D-cycloserine or placebo was a retained as a feature in both the brain only model, and to a lesser extent the brain and non-imaging model.

Limitations and future work

The major limitation of this study is the lack validation of the model in an external dataset. While some studies hold out a proportion of the original data to create an external validation dataset, this technique was not possible here due to restrictions of sample size. To address these limitations, we used a cross validation approach to provide an indication of out of sample transferability in which the support vector machine was exposed to multiple iterations of the sub-sampled dataset during model training, and therefore never 'saw' the entire dataset until the test phase. Models with a large number of measures compared with events (responder or non-responder) risk overfitting and demonstrate poor generalisability to novel datasets. Our dataset contained 35 potential features and therefore was at risk of overfitting. To address this issue, we reduced the number of data-dimensions via feature selection, employed cross-validation and used an automated tuning of the regularisation parameter 'C'. However, while these techniques may ameliorate some of the risk of overfitting, a future study with larger sample size, or independently collected datasets, would take the next steps to externally validate the brain-behaviour model and allow assessment of generalisability.

A key feature of support vector machines is that they fit high-dimensional discriminatory planes between multiple measures to predict an outcome. This multivariate approach affords greater sensitivity in distinguishing between non-separable distributions along a single dimension. The additional use of a non-linear kernel also enables us to capture relationships between highly disparate biological features which often demonstrate non-linear profiles and as a result predict changes in breathlessness. Although this technique leads to less intuitive interpretations of feature weightings, the methods used are the first to successfully demonstrate a relationship between breathlessness expectation related brain activity and changes to reported breathlessness over pulmonary rehabilitation. To reconcile these challenges and move towards eventual clinical application, we suggest that this model form the basis for further studies scrutiny via first an external validation dataset and then further interventional studies.

While larger sample sizes are now required to translate these mechanistic models into clinical relevance, the data provides evidence that breathlessness expectation related brain activity at baseline strongly influences how patients respond to treatment in a predictable manner.

CONCLUSIONS

This study offers the first steps towards brain-based predictive biomarkers for pulmonary rehabilitation outcome. We have shown that models including objective brain markers of breathlessness-expectation are able to predict, for the first time, which patients will have clinically important improvements in breathlessness over pulmonary rehabilitation. Such models could provide new insights into the mechanisms by which breathlessness may be targeted, paving the way for targeted behavioural and pharmacological interventions.

Contributors SF: acquisition of data, analysis, interpretation, drafting, editing and approving manuscript. MB: approval of statistical methods, editing and approving manuscript. ED: approval of statistical methods, editing and approving manuscript. CH: data interpretation, editing and approving manuscript. AR: data interpretation, editing and approving manuscript. NMR: study interpretation, editing and approving manuscript. KTSP: study design, interpretation, editing, approving manuscript and guarantor.

Funding This work was supported by the JABBS Foundation and Dunhill Medical Trust. This research was funded in whole, or in part, by the Wellcome Trust 203139/Z/16/Z. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission. CH is supported by the National Institute for Health Research Biomedical Research Centre based at Oxford Health NHS Foundation Trust and The University of Oxford, and by the UK Medical Research Council. KTSP and NMR are supported by the National Institute for Health Research Biomedical Research Centre based at Oxford University Hospitals NHS Foundation Trust and the University of Oxford. AR is funded by a fellowship from MQ: Transforming Mental Health and is supported by the NIHR Oxford Health Biomedical Research Centre.

Competing interests CH has valueless shares in p1vital and serves on their advisory panel. She has received consultancy payments from p1vital, Zogenix, J&J, Pfizer, Servier, Eli-Lilly, Astra Zeneca, Lundbeck. KTSP is named as co-inventors on a provisional UK patent application titled Use of cerebral nitric oxide donors in the assessment of the extent of brain dysfunction following injury. NMR, has received consulting fees from Rocket Medical UK. MB has received travel expenses from Lundbeck for attending conferences, has shares in P1vtial Products and has acted as a consultant for Jansen and for CHDR. Drs Pattinson and Finnegan are named as co-inventors on a provisional U.K. patent titled "Discordant sensory stimulus in VR based exercise" UK Patent office application: 2204698.1 filing date 31/3/2022. The remaining authors have no biomedical financial interests or potential conflicts of interest.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by South Central Oxford REC B (Ref: 118784, Ethics number: 12/SC/0713). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Sarah L Finnegan http://orcid.org/0000-0003-2249-7612
Andrea Reinecke http://orcid.org/0000-0001-5130-7318
Kyle T S Pattinson http://orcid.org/0000-0003-1353-2199

REFERENCES

1 Reijnders T, Schuler M, Wittmann M, *et al*. The impact of disease-specific fears on outcome measures of pulmonary rehabilitation in patients with COPD. *Respir Med* 2019;146:87–95.
2 Garrod R, Marshall J, Barley E, *et al*. Predictors of success and failure in pulmonary rehabilitation. *Eur Respir J* 2006;27:788–94.
3 Selzler A-M, Simmonds L, Rodgers WM, *et al*. Pulmonary rehabilitation in chronic obstructive pulmonary disease: predictors of program completion and success. *COPD* 2012;9:538–45.
4 Vagaggini B, Costa F, Antonelli S, *et al*. Clinical predictors of the efficacy of a pulmonary rehabilitation programme in patients with COPD. *Respir Med* 2009;103:1224–30.
5 Boutou AK, Tanner RJ, Lord VM, *et al*. An evaluation of factors associated with completion and benefit from pulmonary rehabilitation in COPD. *BMJ Open Respir Res* 2014;1:e000051.
6 Stefan MS, Priya A, Martin B, *et al*. How well do patients and providers agree on the severity of dyspnea? *J Hosp Med* 2016;11:701–7.
7 Chen C-H, Ridler K, Suckling J, *et al*. Brain imaging correlates of depressive symptom severity and predictors of symptom improvement after antidepressant treatment. *Biol Psychiatry* 2007;62:407–14.
8 Ressler KJ, Rothbaum BO, Tannenbaum L, *et al*. Cognitive enhancers as adjuncts to psychotherapy: use of D-cycloserine in phobic individuals to facilitate extinction of fear. *Arch Gen Psychiatry* 2004;61:1136–44.
9 Hofmann SG, Smits JAJ, Rosenfield D, *et al*. D-Cycloserine as an augmentation strategy with cognitive-behavioral therapy for social anxiety disorder. *Am J Psychiatry* 2013;170:751–8.
10 Reinecke A, Nickless A, Browning M, *et al*. Neurocognitive processes in D-cycloserine augmented single-session exposure therapy for anxiety: a randomized placebo-controlled trial. *Behav Res Ther* 2020;129:103607.
11 Finnegan SL, Harrison OK, Harmer CJ, *et al*. Breathlessness in COPD: linking symptom clusters with brain activity. *Eur Respir J* 2021;58. doi:10.1183/13993003.04099-2020. [Epub ahead of print: 18 11 2021].
12 Yorke J, Moosavi SH, Shuldham C, *et al*. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010;65:21–6.
13 Radloff LS. The CES-D scale. *Appl Psychol Meas* 1977;1:385–401.
14 Spielberger CD. State-Trait Anxiety Inventory. In: *The Corsini encyclopedia of psychology*, 2010.
15 Krupp LB, LaRocca NG, Muir-Nash J, *et al*. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989;46:1121–3.
16 Jones PW, Quirk FH, Baveystock CM, *et al*. A self-complete measure of health status for chronic airflow limitation. The St. George's respiratory questionnaire. *Am Rev Respir Dis* 1992;145:1321–7.
17 Bestall JC, Paul EA, Garrod R, *et al*. Usefulness of the medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581–6.
18 De Peuter S, Lemaigre V, Van Diest I, *et al*. Illness-specific catastrophic thinking and overperception in asthma. *Health Psychology* 2008;27:93–9.
19 McCracken LM, Vowles KE, Eccleston C. Acceptance-based treatment for persons with complex, long standing chronic pain: a preliminary analysis of treatment outcome in comparison to a waiting phase. *Behav Res Ther* 2005;43:1335–46.
20 Bradley J, Howard J, Wallace E, *et al*. Validity of a modified shuttle test in adult cystic fibrosis. *Thorax* 1999;54:437–9.
21 Levy ML, Quanjer PH, Booker R, *et al*. Diagnostic spirometry in primary care: Proposed standards for general practice compliant with American Thoracic Society and European Respiratory Society recommendations: a General Practice Airways Group (GPIAG)1 document, in association with the Association for Respiratory Technology & Physiology (ARTP)2 and Education for Health3 1 www.gpiag.org 2 www.artp.org 3 www.educationforhealth.org.uk. *Prim Care Respir J* 2009;18:130–47.
22 Herigstad M, Faull OK, Hayen A, *et al*. Treating breathlessness *via* the brain: changes in brain activity over a course of pulmonary rehabilitation. *Eur Respir J* 2017;50. doi:10.1183/13993003.01029-2017. [Epub ahead of print: 12 09 2017].
23 Herigstad M, Hayen A, Reinecke A, *et al*. Development of a dyspnoea word cue set for studies of emotional processing in COPD. *Respir Physiol Neurobiol* 2016;223:37–42.
24 Esser RW, Stoeckel MC, Kirsten A, *et al*. Brain activation during perception and anticipation of dyspnea in chronic obstructive pulmonary disease. *Front Physiol* 2017;8:617.
25 Faull OK, Hayen A, Pattinson KTS. Breathlessness and the body: neuroimaging clues for the inferential leap. *Cortex* 2017;95:211–21.
26 Lunardon N, Menardi G, Torelli N. Rose: a package for binary imbalanced learning. *R J* 2014;6:79–92.
27 Janssens T, Ritz T. Perceived triggers of asthma: key to symptom perception and management. *Clin Exp Allergy* 2013;43:1000–8.
28 Vlemincx E, Sprenger C, Büchel C. Expectation and dyspnoea: the neurobiological basis of respiratory nocebo effects. *Eur Respir J* 2021;58. doi:10.1183/13993003.03008-2020. [Epub ahead of print: 23 09 2021].
29 Abdallah SJ, Faull OK, Wanigasekera V, *et al*. Opioids for breathlessness: psychological and neural factors influencing response variability. *Eur Respir J* 2019;54:1900275.
30 Ferreira DH, Louw S, McCloud P, *et al*. Controlled-Release oxycodone vs. placebo in the treatment of chronic Breathlessness-A multisite randomized placebo controlled trial. *J Pain Symptom Manage* 2020;59:581–9.
31 Currow D, Louw S, McCloud P, *et al*. Regular, sustained-release morphine for chronic breathlessness: a multicentre, double-blind, randomised, placebo-controlled trial. *Thorax* 2020;75:50–6.
32 Janssens T, De Peuter S, Stans L, *et al*. Dyspnea perception in COPD: association between anxiety, dyspnea-related fear, and dyspnea in a pulmonary rehabilitation program. *Chest* 2011;140:618–25.
33 Marlow LL, Faull OK, Finnegan SL, *et al*. Breathlessness and the brain: the role of expectation. *Curr Opin Support Palliat Care* 2019;13:200–10.
34 Monahan JB, Handelmann GE, Hood WF, *et al*. D-Cycloserine, a positive modulator of the N-methyl-D-aspartate receptor, enhances performance of learning tasks in rats. *Pharmacol Biochem Behav* 1989;34:649–53.
35 Klumpp H, Fitzgerald DA, Piejko K, *et al*. Prefrontal control and predictors of cognitive behavioral therapy response in social anxiety disorder. *Soc Cogn Affect Neurosci* 2016;11:630–40.
36 Falconer E, Allen A, Felmingham KL, *et al*. Inhibitory neural activity predicts response to cognitive-behavioral therapy for posttraumatic stress disorder. *J Clin Psychiatry* 2013;74:895–901.
37 Chekroud AM, Zotti RJ, Shehzad Z, *et al*. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry* 2016;3:243–50.

Supplementary Material

Brain activity measured by functional brain imaging predicts breathlessness improvement during pulmonary rehabilitation

Sarah L. Finnegan (sarah.finnegan@ndcn.ox.ac.uk)

Michael Browning

Eugene Duff

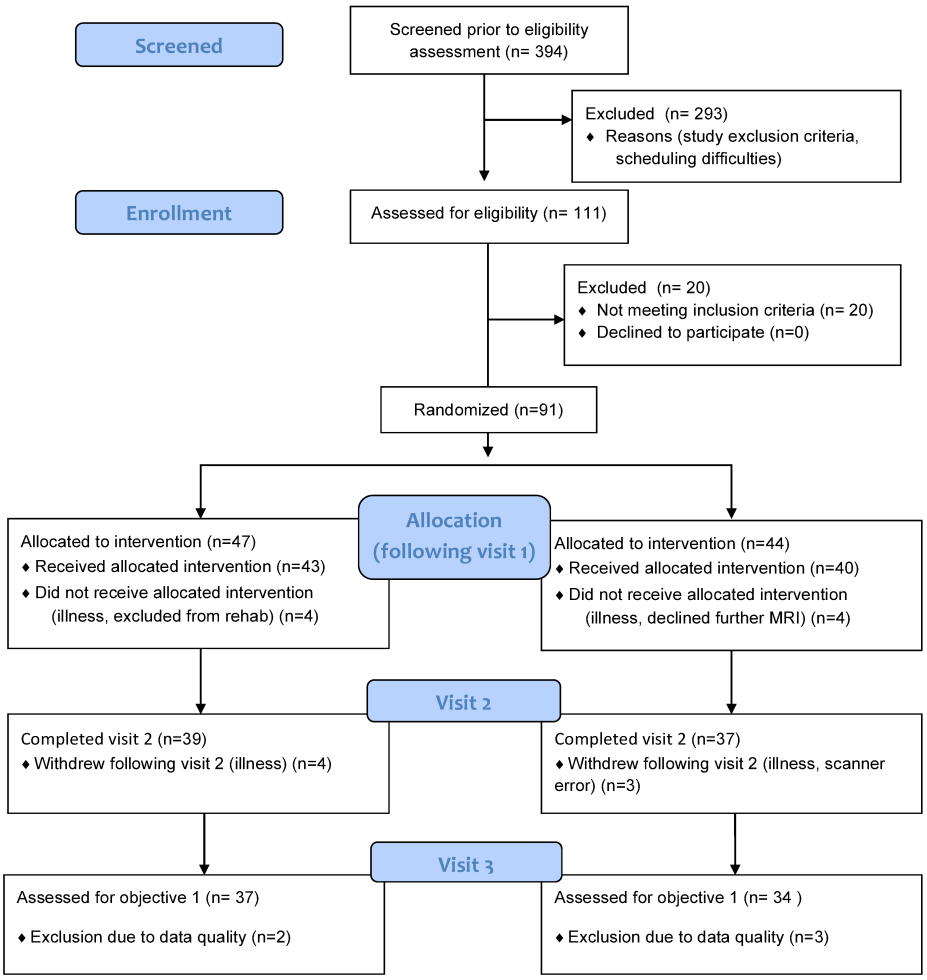
Catherine J. Harmer

Andrea Reinecke

Najib M. Rahman

Kyle T.S. Pattinson (kyle.pattinson@nda.ox.ac.uk)

Online Data Supplement



Supplementary Figure 1. *Consort diagram illustrating stages of participant recruitment from initial screening through to completion at visit 3.*

91 participants (30 female, median age 70 years; range 46-85 years) with COPD were recruited to this study. 72 participants completed all three visits. 1 participant's brain imaging data was lost due to data collection error.

Inclusion criteria

Study inclusion criteria were: a diagnosis of COPD and admittance to pulmonary rehabilitation. Exclusion criteria were: inadequate understanding of verbal and written

English, significant cardiac, psychiatric (including depression under tertiary care) or metabolic disease (including insulin-controlled diabetes), stroke, contraindications to either D-cycloserine (including alcoholism) or magnetic resonance imaging (MRI), epilepsy, claustrophobia, regular therapy with opioid analgesics or home oxygen therapy.

Supplementary Table 1. *Demographic information from the 71 participants who completed all study visits. Variance is expression either in terms of standard deviation (SD) or interquartile range (IQR) depending on the normality of the underlying data distribution.* ***BMI*** *= Body Mass Index,* ***MRC*** *= Medical Research Council.* ***SpO_2%*** *= Peripheral Oxygen saturation, expressed as a percentage. Also listed with prevalence in brackets are recorded comorbidities ordered by frequency*

	Pre-rehabilitation	Post-rehabilitation
Age (median years / range)	**71 / (46-85)**	
Smoking pack-years (years / IQR)	**30.1 ± 28.5**	
BMI (kg.m-2 ± SD)	**26.9 ± 6.0**	**27.4 ± 4.6**
MRC breathlessness scale (IQR)	**3 (1)**	**2 (1)**
Resting SpO_2% (IQR)	**95 ± 3.75**	**94 ± 3.75**
Resting heart rate (beats.min-1 ± SD)	**80.8 ± 14.1**	**78.3 ± 13.0**
FEV1/FVC (IQR)	**0.55 (0.15)**	**0.57 (0.28)**
Duration of breathlessness (years / IQR)	**10 (15.6)**	
Total exacerbations (number / IQR)	**0 (2.3)**	
Comorbidities (frequency)		
Asthma (25)	Reflux and heart burn (22)	Hypertension (24)
Swelling of both ankles (19)	Surgery to the chest (13)	Depression (13)
Diabetes (9)	Heart attack (9)	Bronchiectasis (7)
Osteoporosis (6)	Arrhythmia (6)	Inflammatory bowel disease (5)
Peptic ulcer (5)	Heart failure (2)	Neuromuscular weakness (2)
Tuberculosis (1)		

Supplementary Table 2. *Demographic information from the 71 participants who completed the three study visits expressed as the group total group and D-cycloserine and placebo groups. Variance is expression either in terms of standard deviation (SD) or interquartile range (IQR).* ***BMI*** *= Body Mass Index,* ***MRC*** *= Medical Research Council clinical measure of breathlessness.* ***SpO_2%*** *= Peripheral Oxygen saturation, expressed as a percentage.*

Visit 1 (N=71)	Total	D-cycloserine	Placebo
Age (median years/range)	**71 / (46-85)**	**71 / (47-81)**	**71.5 / (46-85)**
Smoking pack-years (IQR)	**30 / (28.5)**	**34 / (25.6)**	**30 / (30.0)**
BMI kg.m-2 ± SD	**26.9 ± 6.0**	**27.3 ± 6.5**	**26.9 ± 5.7**
MRC (IQR)	**3 (1)**	**3 (1)**	**2.5 (1)**
Resting SpO_2% (IQR)	**95 / (3.8)**	**95 / (3.3)**	**94.5 / (3.0)**
Resting heart rate beats.min-1 ± SD	**80.8 ± 14.1**	**80.8 ± 13.4**	**80.8 ± 15.0**
FEV1/FVC (IQR)	**0.55 / (0.15)**	**0.53 / (0.17)**	**0.56 / (0.13)**

Study Drug

Study drugs were purchased from Ipswich Hospital Pharmacy Manufacturing Unit, Heath Road, Ipswich IP4 5PD, Tel: 01473 703603.

Randomisation Procedure

Once the participant gave written consent to the trial and completed the MRI scan, a member of the team submitted a randomisation form, entering eligibility criteria and minimisation factors. Allocation to active or placebo capsules was carried out by Sealed Envelope Randomisation Services (Sealed Envelope Ltd, Concorde House, Grenville Place, London NW7 3SA). The randomisation number was then provided to

the Oxford Respiratory Trials Unit who dispensed the drug/placebo. Minimisation factors were as follows:

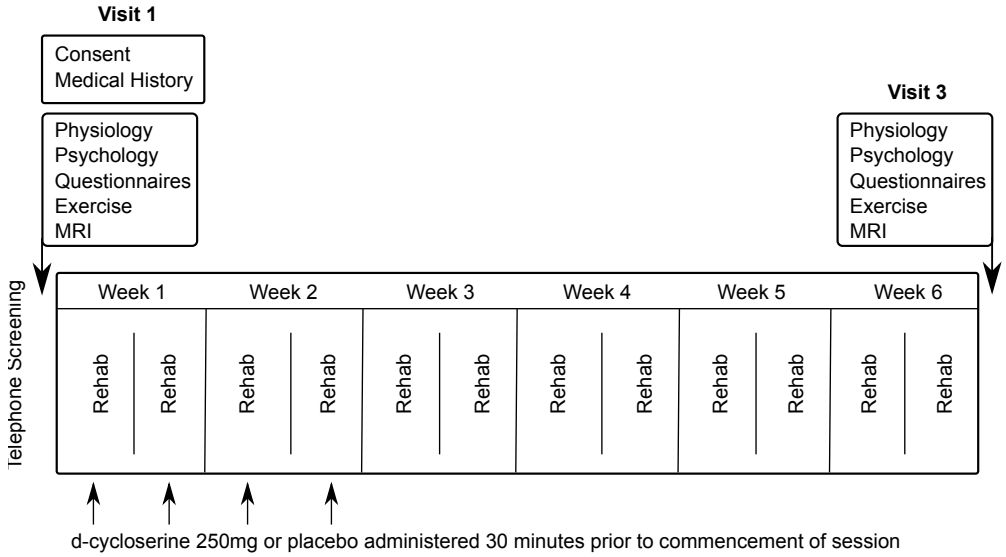
1. Centre
2. MRC grade
3. Diabetes
4. Antidepressant
5. Age at which the participant completed full time education
6. Previous rehabilitation

Randomisation codes were held by Sealed Envelope until study completion, after which at the first stage of unblinding an independent researcher provided study researchers with a coded binarised system for analysis. Researchers remained blinded to group identity until analysis was completed. No side effects were reported.

Study Visit Protocol

Following telephone screening participants were invited to attend their first research session (baseline) prior to starting pulmonary rehabilitation. Pulmonary rehabilitation courses were run by either Oxford Health NHS Foundation Trust, West Berkshire NHS Foundation Trust, or Milton Keynes University Hospitals NHS Trust. Following the successful completion of the first study session participants were then randomised to receive either the study drug or placebo. Participants then attended their first four sessions of pulmonary rehabilitation, 30 minutes prior to each of these sessions they received their assigned study drug or placebo tablet. A second study visit took place following the fourth pulmonary rehabilitation session but before the 6th session. Participants then completed the remainder of their pulmonary rehabilitation course before attending a third study session (Supplementary Figure 2) that occurred in the two weeks after termination of pulmonary rehabilitation. For the purposes of this study,

data were used from visits occurring before (visit one) and following the completion of pulmonary rehabilitation (visit three).



Supplementary Figure 2. *A schematic demonstrating order of visits, rehabilitation sessions and tablet administration throughout the study period. Participants took part in one study visit prior to their first pulmonary rehabilitation session. Study drug/placebo were administered on four occasions over the first four rehabilitation sessions. Participants continued with their pulmonary rehabilitation course for a further four weeks before returning for a final visit.*

Behavioural Measures

Questionnaire Measures

Dyspnoea-12 (D12) Questionnaire: This is a 12-item questionnaire designed to measure the severity of breathlessness and has been validated for use in patients with respiratory disease [1].

Centre for Epidemiologic Studies Depression Scale (CES-D): Depressive symptoms are commonly observed in patients with respiratory disease. This brief questionnaire consists of 20 items investigates the symptoms of depression across a number of factors [2].

State-Trait Anxiety Inventory (STAIT-T): This questionnaire assesses participant's general level of anxiety in particular scenarios via 20 questions asking "how anxious you generally feel" [3].

Fatigue Severity Scale: This 9-point questionnaire quantifies patient fatigue, which is well documented in its association with COPD [4].

St George's Respiratory Questionnaire (SGRQ): There are 50 questions in this questionnaire, which has been developed and validated for use in COPD and asthma. The questions measure the impact of overall health, daily life and well-being [5].

Medical Research Council (MRC) breathlessness scale: The MRC scale quantifies perceived difficulty due to respiratory restrictions on a scale of 1 to 5 [6].

Mobility Inventory (MI): This questionnaire collects data regarding the extent to which a participant avoids certain situations, either alone or accompanied (21-items in each category) [7].

Breathlessness Catastrophising Scale: Adapted from Catastrophic Thinking Scale in Asthma. This 13-point questionnaire was modified for this study by substituting the word "asthma" for "breathlessness" in order to measure catastrophic thinking [8] [9].

Breathlessness Awareness and Vigilance Scale Pain: Adapted from Pain Awareness and Vigilance Scale. This questionnaire was modified by substituting the word "breathlessness" for the word "pain". The 16-point scale measures how much a participant focuses their attention onto their breathlessness [10] [9].

Physiological Measures

A trained respiratory nurse collected spirometry measures of FEV_1 and FVC using Association for Respiratory Technology and Physiology standards [11]. Participants performed two modified incremental shuttle walk tests (MSWT) [12], and heart rate and oxygen saturations (SpO_2) were measured immediately before the MSWT and subsequently every minute until 10 minutes post-exercise (or until participants returned to their baseline state) using a fingertip pulse oximeter (Go_2; Nonin Medical Inc). Before

and after the MWST participants also rated their breathlessness on a modified Borg scale [13]. In a MWST participants must walk between and around two cones, placed 10m apart in time to a set of auditory beeps played from a laptop. Initially the speed of beep repetition is slow, but the participant must increase their walking speed each minute in order to reach the cone before the next beep. Participants continue to walk (or run) until they are too breathless to continue, at which point the total distance walked is recorded.

MRI Acquisition

Prior to each MRI session participants were screened for standard MRI contraindications including metal in or about their person, epilepsy and claustrophobia.

Image acquisition:

Hardware: A Tim System (Siemens Healthcare GmbH) 12-channel head coil.

T1 sequence parameters: TR, 2040ms; TE, 4.68ms; voxel size, 1 x 1 x 1 mm; FOV, 200mm; flip angle, 8°; inversion time, 900ms; bandwidth 130 Hz/Px).

T2*-weighted (functional) sequence parameters: TR, 3000ms; TE 30ms; voxel size 3 x 3 x 3 mm; FOV, 192mm; flip angle 87°; echo spacing 0.49ms.

Functional scan durations: word-task - 215 volumes, 7 minutes and 33 seconds duration.

Field map scans of the B_0 field were obtained to aid the distortion correction of the functional scans: TR, 488ms; TE1, 5.19ms; TE2, 7.65ms; flip angle 60°; voxel size, 3.5 x 3.5 x 3.5 mm.

Word Task

This task was developed and published by Herigstad and colleagues in 2016 for use in the COPD population [14]. Word cues were developed in three key stages; firstly in

collaboration with respiratory practitioners, academics and physiotherapists, a set of 30 word cues associated with breathlessness were created. Next, these cues were provided to patients with COPD alongside a VAS rating scale, allowing patients to rate how breathless and anxious the situations identified by the cues would make them feel. Following adjustments based on participant feedback, the word cues were then computerised and tested in a larger population of COPD patients [14]. Further validation was carried out in the fMRI environment and by for clinical sensitivity with comparisons between changes in key questionnaire measures and word-cue rating. Before the first scan session, participants were given the opportunity to practice using the button box with a set of test words.

In this task brain activity was correlated with corresponding visual analogue ratings of anxiety and breathlessness. During the fMRI scanning, participants were presented with a word cue in white text on a black background for 7 seconds. Participants were then asked, “how breathless would this make you feel” (wB) and “how anxious would this make you feel” (wA). To each question participants responded within a 7 second window using a button box and visual analogue scale (VAS). The response marker always initially appeared at the centre of the scale, with the anchors “Not at all” and “Very much” at either end. A control condition, used as a baseline measure of activity in response to the presentation of a visual stimulus was presented 4 times over the course of the scan, consisting of a string of “XXXXXXXXXXXXXXX” with fixed length of 15 characters, and each time was presented for 7 seconds. No rating period followed these control blocks [14].

Imaging Analysis

Functional MRI Preprocessing

Data denoising was carried out as follows: Before the first level analysis, each functional scan was decomposed into maximally independent components using

FMRIB's MELODIC tool (Multivariate Exploratory Linear Optimised Decomposition into Independent Components). "Noise" components were identified by FIX (FMRIB's auto-classification tool, [15, 16]) using the WhII.Standard.RData [17] trained classifier with aggressive clean up option. A Principle Component Analysis (PCA) was run on the FIX identified components to retrain 99% of the variance. Separately, the cardiac and respiratory related physiological signals (recorded via a pulse oximeter and a respiratory bellows) were transformed into a series of regressors, (three cardiac and four respiratory harmonics) as well as an interaction term and a measure of respiratory volume per unit of time (RVT), using FSL's physiological noise modelling tool (PNM). The signal associated with these waveforms (modelled using retrospective image correction (RETROICOR) [18, 19]) was then used to form voxelwise noise regressors.

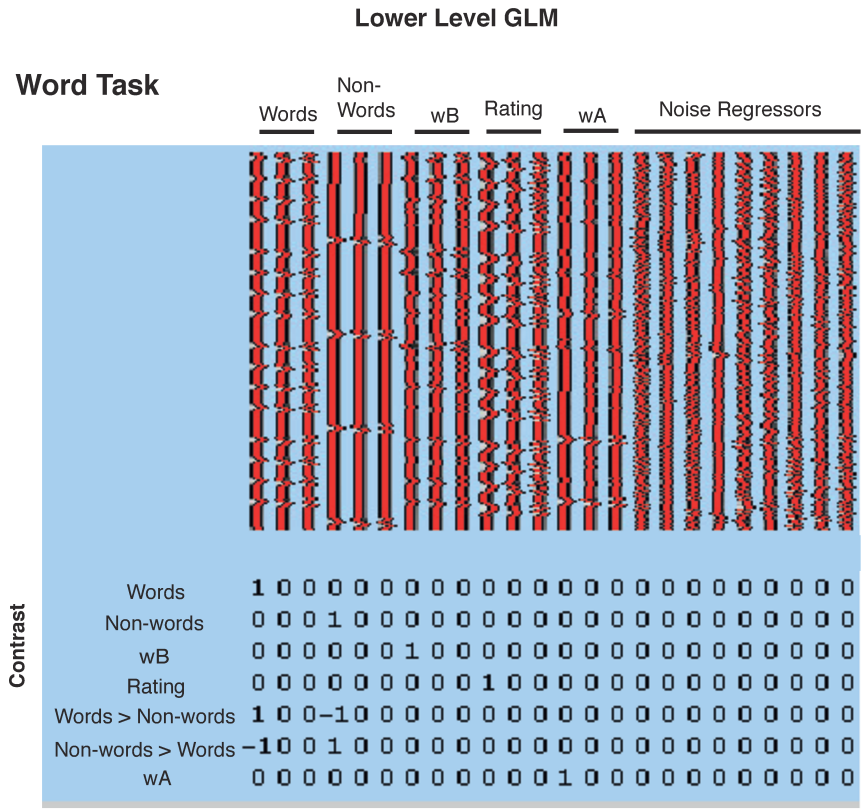
The confounds identified by FSL's FIX and PNM tools, along with sources of noise arising from motion, were then combined into a single model. This single noise model approach builds upon the technique outlined by [20]; and fully detailed by [21]. In these preceding works we employed a step-wise technique whereby physiological noise (identified by PNM) and FIX-identified noise were each removed from the data in separate steps prior to data entry into the lower level model. In the new cleanup pipeline, a single text file containing time-course information relating to FIX identified noise components along with white matter or CSF related noise was included as additional confound EV's within the lower level model, while the PNM-identified noise was entered into the model as a standard voxel-wise confound list. In this updated de-noising pipeline, confounds identified above are added to model at the stage of first-level analysis and thus the functional dataset can be corrected for sources of noise arising from motion, scanner and cerebro-spinal fluid artefacts, cardiac, and respiratory noise in a single step, rather than three.

Functional MRI Analysis

MRI processing was performed using FEAT (FMRI Expert Analysis Tool within the FSL package). The data were corrected for movement using MCFLIRT (Motion correction using FMRIB's Linear Image Registration Tool [22]). Non-brain structures were removed using BET (Brain Extraction Tool [23]). Spatial smoothing was carried out using a full-width-half-maximum Gaussian kernel of 5mm, while high-pass temporal filtering (Gaussian-weighted least squares straight line fitting; 90 s) removed low frequency noise and slow-drift. Distortion correct of EPI data was carried out using a combination of FUGUE (FMRIB's Utility for Geometrically Unwarping EPI's [24, 25] and BBR (Boundary Based Registration; part of the FMR Expert Analysis Tool, FEAT version 6.0 [26]). The data were corrected for physiological noise using FSL's FIX-PNM pipeline. Functional scans were registered in a two-step process to the MNI152 (1x1x1 mm) standard space brain template. Firstly, each subject's EPI was registered to their associated T1-weighted structural image using BBR (6 DOF) with nonlinear field map distortion correction [26]. In the second step the subject's structural image was registered to 1mm standard space via an affine transformation followed by nonlinear registration (using FNIRT: FMRIB's Non-linear Registration Tool [27]).

First Level Processing

Functional MRI: Word-cue task: At the individual subject level, a general linear model (GLM) was created with explanatory variables (EVs) for breathlessness word or non-word presentation, and two de-meaned EVs modeling the reported breathlessness and anxiety response to the word cues (Supplementary Figure 3). Additional explanatory noise variables were included to model the period during which the participant responded using the visual analog scale (VAS).



Supplementary Figure 3 *– An illustration of the generalised linear models (GLM) used lower level analyses for the word task. Abbreviations as follows – wB – breathlessness rating, wA – breathlessness anxiety rating.*

Regions of Interest

Regions of interest defined by the Harvard-Oxford Atlas and Destrieux' cortical atlas are listed here with atlas label identifier in brackets if different from anatomical name.

1. Anterior Insular Cortex (G_insular_short)
2. Posterior Insular Cortex (G_Ins_lg_and_S_cent_ins)
3. Anterior Cingulate Cortex (Cingulate Cortex, anterior division)
4. Amygdala
5. Hippocampus
6. Posterior Cingulate Cortex (Cingulate Cortex, posterior division)

7. Medial Prefrontal Cortex (Frontal Pole)
8. Middle frontal Gyrus
9. Superior Marginal Gyrus (Supramarginal Gyrus)
10. Superior Frontal Gyrus
11. Putamen
12. Precuneus
13. Angular Gyrus
14. Caudate
15. Precentral Gyrus

A 40% probability threshold was applied to the mask of each region. The regions were then registered to each individual before being re-thresholded at 40% probability to avoid interpolation errors and were binarized.

Model specifics and technical definitions

Cross validation – Is a resampling procedure. Data can be split into a number of different training and testing folds (k-folds). This enables the algorithm to learn from the maximum number of new data points.

Support Vector Machines (SVM) – Separates pre-defined classes by establishing an optimal boundary within high-dimensional space called a hyper-plane. In this work we used a linear kernel which draws an assumption on the relationship between activity space and feature space but places no assumption on the distribution that the two are drawn from.

Elastic net regularisation – Model fitting involves a trade-off between bias and variance. Bias is the difference between predicted regression parameters and actual estimator, essentially a measure of accuracy. Variance is the uncertainty of those estimations. Both of these numbers should be low. Regularisation provides a way to reduce the variance while introducing some bias. Elastic net regularisation combines

elements of ridge and lasso regression techniques. Ridge elements - A cost function is applied to the slope of any fit function to reduce sudden changes, helping with the stability of the classifier. Lasso elements – accounts for magnitude of feature contribution to the classifier but struggles where variables are highly correlated. Elastic net includes terms for both slope and magnitude as a mix-ratio which encourages groupings of variables.

Random OverSample Examples (ROSE) – Imbalanced classes can affect classifier performance. ROSE creates an artificially balanced sample using a smoothed bootstrap approach.

Seed setting – Random seeds are used in the generation of models. The random number is set locally each time to the same number to ensure the model sequence is reproducible.

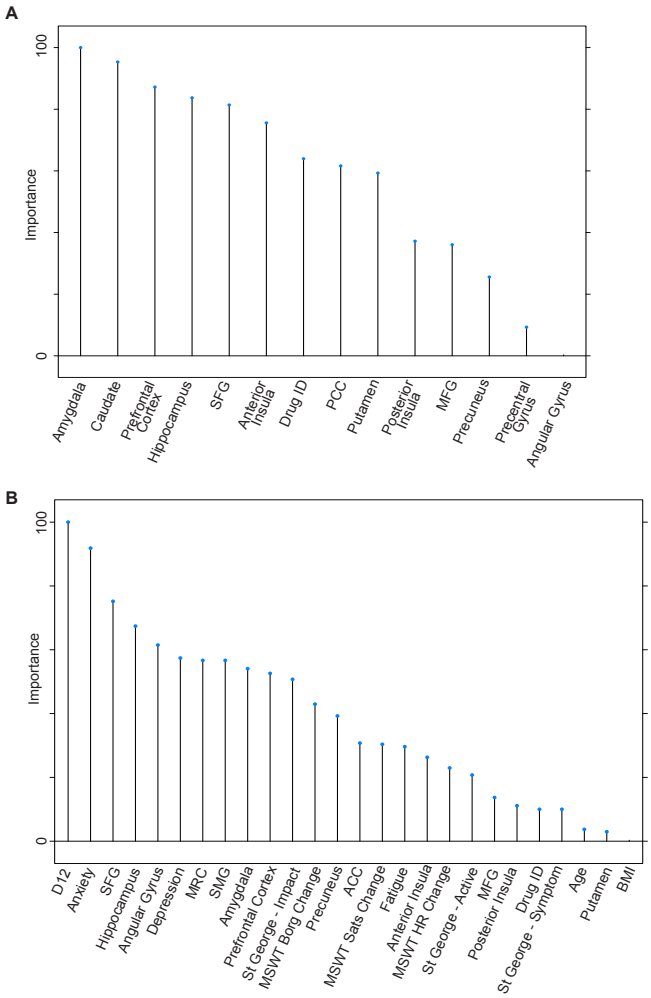
Confusion matrices – A visual representation of the success of a supervised classifier. Where rows represent the classifiers attempt and columns represent the actual class.

		Actual Class	
Predicted Class		Dog	Cat
	Dog	**5**	2
	Cat	3	**3**

Results

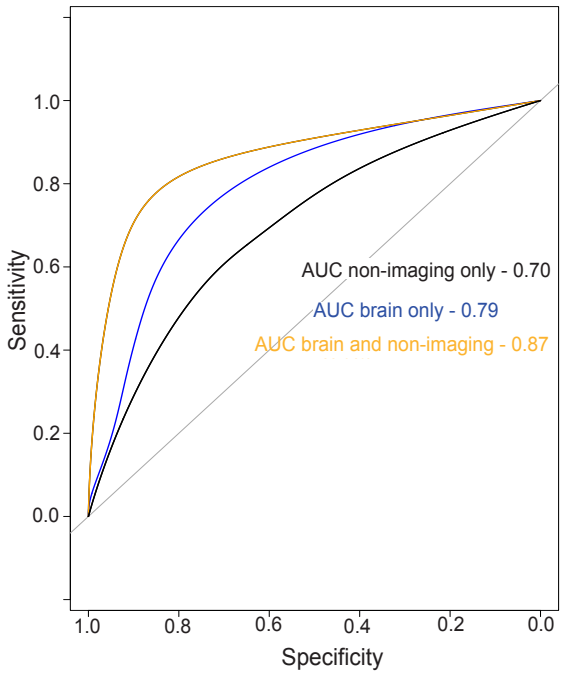
Supplementary Table 3. *A comparison of the number (N (%)) of participants who demonstrated clinically meaningful change in two of the three scores of D12, MSWT and St George scores.*

	Comparison of D12 and St George	Comparison of D12 and MSWT	Comparison of St George and MSWT
Two scores demonstrating clinically significant change	N=29 (40%)	N=16 (23%)	N=18 (25%)
One demonstrating clinically significant change	N=26 (37%)	N=37 (52%)	N=35 (50%)
No demonstration of clinically significant change across either score	N=16 (23%)	N=18 (25%)	N=18 (25%)
Taking a change of 54m of greater as a clinically significant change for the MSWT, a clinically important difference of 4 or greater for the total St George score and minimally clinically meaningful score of 3 for D12			

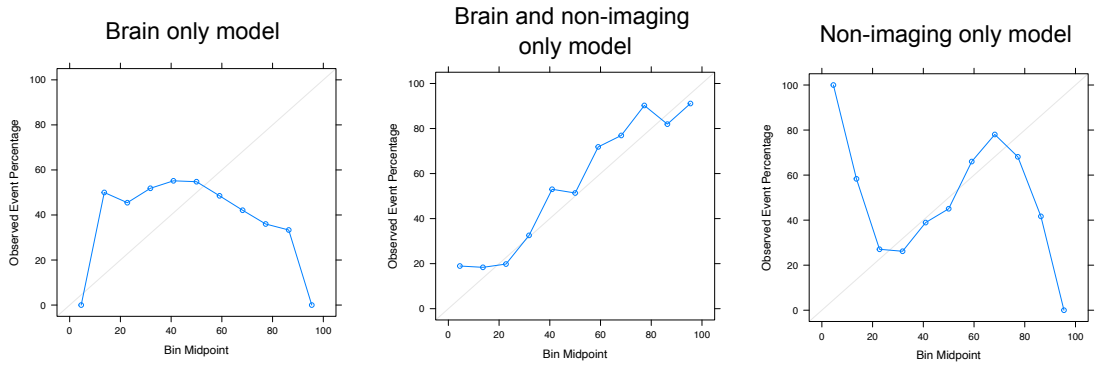


Supplementary Figure 4. *Needle plot of ranked importance for each of the* ***(A)*** *brain derived metrics (full model) and* ***(B)*** *brain derived metrics and non-imaging measures (full model), to*

the classification of responder/non-responders. Abbreviations: SFG – Superior Frontal Gyrus, PCC – Posterior Cingulate Cortex, MFG – Middle Frontal Gyrus, ACC – Anterior Cingulate Cortex, MSWT – Modified Shuttle Walk Test, HR – Heart Rate, BORG – breathlessness scale, BMI – Body Mass Index, ; SpO^2*– Oxygen saturation.*



Supplementary Figure 5. *Receiver Operator Curves for each of the three full models. AUC – area under the curve.*



Supplementary Figure 6. *Calibration Curves for each of the three full models demonstrating goodness of fit.*

1. Yorke, J., et al., *Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12.* 2010. **65**(1): p. 21-26.
2. Radloff, L.S., *The CES-D Scale:A Self-Report Depression Scale for Research in the General Population.* 1977. **1**(3): p. 385-401.
3. Spielberger, C.D., *State-Trait Anxiety Inventory*, in *The Corsini Encyclopedia of Psychology*. 2010.
4. Krupp, L.B., et al., *The fatigue severity scale: Application to patients with multiple sclerosis and systemic lupus erythematosus.* Archives of Neurology, 1989. **46**(10): p. 1121-1123.
5. Jones, P.W., et al., *A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire.* Am Rev Respir Dis, 1992. **145**(6): p. 1321-7.
6. Bestall, J.C., et al., *Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease.* Thorax, 1999. **54**(7): p. 581-586.

7. Chambless, D.L., et al., *The Mobility Inventory for Agoraphobia.* Behav Res Ther, 1985. **23**(1): p. 35-44.
8. De Peuter, S., et al., *Illness-specific catastrophic thinking and overperception in asthma.* Health Psychol, 2008. **27**(1): p. 93-9.
9. Herigstad, M., et al., *Dyspnea-related cues engage the prefrontal cortex: evidence from functional brain imaging in COPD.* Chest, 2015. **148**(4): p. 953-961.
10. McCracken, L.M., K.E. Vowles, and C. Eccleston, *Acceptance-based treatment for persons with complex, long standing chronic pain: a preliminary analysis of treatment outcome in comparison to a waiting phase.* Behav Res Ther, 2005. **43**(10): p. 1335-46.
11. Levy, M.L., et al., *Diagnostic spirometry in primary care: Proposed standards for general practice compliant with American Thoracic Society and European Respiratory Society recommendations: a General Practice Airways Group (GPIAG)1 document, in association with the Association for Respiratory Technology & Physiology (ARTP)2 and Education for Health3 1 www.gpiag.org 2 www.artp.org 3 www.educationforhealth.org.uk.* Prim Care Respir J, 2009. **18**(3): p. 130-47.
12. Bradley, J., et al., *Validity of a modified shuttle test in adult cystic fibrosis.* Thorax, 1999. **54**(5): p. 437-9.
13. Mahler, D.A., et al., *Comparison of clinical dyspnea ratings and psychophysical measurements of respiratory sensation in obstructive airway disease.* Am Rev Respir Dis, 1987. **135**(6): p. 1229-33.
14. Herigstad, M., et al., *Development of a dyspnoea word cue set for studies of emotional processing in COPD.* Respiratory physiology & neurobiology, 2016. **223**: p. 37-42.
15. Griffanti, L., et al., *ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging.* Neuroimage, 2014. **95**: p. 232-47.
16. Salimi-Khorshidi, G., et al., *Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers.* Neuroimage, 2014. **90**: p. 449-68.
17. Filippini, N., et al., *Study protocol: the Whitehall II imaging sub-study.* BMC Psychiatry, 2014. **14**(1): p. 159.
18. Harvey, A.K., et al., *Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise.* J Magn Reson Imaging, 2008. **28**(6): p. 1337-44.
19. Brooks, J.C., et al., *Physiological noise modelling for spinal functional magnetic resonance imaging studies.* Neuroimage, 2008. **39**(2): p. 680-92.
20. Faull, O.K., et al., *Conditioned respiratory threat in the subdivisions of the human periaqueductal gray.* Elife, 2016. **5**.
21. Hayen, A., et al., *Opioid suppression of conditioned anticipatory brain responses to breathlessness.* Neuroimage, 2017. **150**: p. 383-394.
22. Jenkinson, M., et al., *Improved optimization for the robust and accurate linear registration and motion correction of brain images.* Neuroimage, 2002. **17**(2): p. 825-41.
23. Smith, S.M., *Fast robust automated brain extraction.* Hum Brain Mapp, 2002. **17**(3): p. 143-55.

24. Holland, D., J.M. Kuperman, and A.M. Dale, *Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging.* Neuroimage, 2010. **50**(1): p. 175-83.
25. Jenkinson, M., *Fast, automated, N-dimensional phase-unwrapping algorithm.* Magn Reson Med, 2003. **49**(1): p. 193-7.
26. Greve, D.N. and B. Fischl, *Accurate and robust brain image alignment using boundary-based registration.* Neuroimage, 2009. **48**(1): p. 63-72.
27. Andersson, J., *Non-Linear registration, aka spatial normalisation.* FMRIB technical report, 2010. **TR07JA2**.