

## Deep learning model of fMRI connectivity predicts PTSD symptom trajectories in recent trauma survivors



Shelly Sheynin<sup>a</sup>, Lior Wolf<sup>a,\*</sup>, Ziv Ben-Zion<sup>b,c</sup>, Jony Sheynin<sup>e</sup>, Shira Reznik<sup>b</sup>, Jackob Nimrod Keynan<sup>b,d</sup>, Roee Admon<sup>f,g</sup>, Arieh Shalev<sup>h</sup>, Talma Hendler<sup>b,c,i,j,1</sup>, Israel Liberzon<sup>e,1</sup>

<sup>a</sup> School of Computer Science, Tel Aviv University, Tel-Aviv, Israel

<sup>b</sup> Sagol Brain Institute Tel-Aviv, Wohl Institute for Advanced Imaging, Tel Aviv Sourasky Medical Center, Tel-Aviv, Israel

<sup>c</sup> Sagol School of Neuroscience, Tel-Aviv University, Tel Aviv, Israel

<sup>d</sup> Department of Psychiatry and Behavioral Science, Stanford University School of Medicine, Stanford, USA

<sup>e</sup> Department of Psychiatry and Behavioral Science, Texas A&M University Health Science Center, TX, USA

<sup>f</sup> School of Psychological Sciences, University of Haifa, Haifa, Israel

<sup>g</sup> The Integrated Brain and Behavior Research Center (IBBRC), University of Haifa, Haifa, Israel

<sup>h</sup> Department of Psychiatry, New York University Grossman School of Medicine, New York, NY, USA

<sup>i</sup> School of Psychological Sciences, Faculty of Social Sciences, Tel-Aviv University, Tel-Aviv, Israel

<sup>j</sup> Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

### ARTICLE INFO

#### Keywords:

fMRI  
Deep learning  
Attention mechanism  
End-to-end neural network  
PTSD symptom clusters

### ABSTRACT

Early intervention following exposure to a traumatic life event could change the clinical path from the development of post traumatic stress disorder (PTSD) to recovery, hence the interest in early detection and underlying biological mechanisms involved in the development of post traumatic sequelae. We introduce a novel end-to-end neural network that employs resting-state and task-based functional MRI (fMRI) datasets, obtained one month after trauma exposure, to predict PTSD symptoms at one-, six- and fourteen-months after the exposure. fMRI data, as well as PTSD status and symptoms, were collected from adults at risk for PTSD development, after admission to emergency room following a traumatic event. Our computational method utilized a per-region encoder to extract brain regions embedding, which were subsequently updated by applying the algorithmic technique of pairwise attention. The affinities obtained between each pair of regions were combined to create a pairwise co-activation map used to perform multi-label classification. The results demonstrate that the novel method's performance in predicting PTSD symptoms, in a prospective manner, outperforms previous analytical techniques reported in the fMRI literature, all trained on the same dataset. We further show a high predictive ability for predicting PTSD symptom clusters and PTSD persistence. To the best of our knowledge, this is the first deep learning method applied on fMRI data with respect to prospective clinical outcomes, to predict PTSD status, severity and symptom clusters. Future work could further delineate the mechanisms that underlie such a prediction, and potentially improve single patient characterization.

### 1. Introduction

Post-traumatic stress disorder (PTSD) is a common psychiatric disorder with significant clinical and public health impact, due to its high prevalence, chronicity, associated functional impairment and frequent comorbidities (Kessler, 2000; Shalev et al., 2017). While most trauma-exposed survivors who develop initial PTSD symptoms exhibit rapid remission (about 56%) or show delayed/partial remission (about 27%) over fourteen months, a subset of about 17% do not remit and suffer from chronic PTSD (Galatzer-Levy et al., 2013). The development of

PTSD in a subgroup of survivors, and the tenacity of the protracted disorder, suggest a long-lasting trauma induced neuro-behavioral alteration (Pitman et al., 2012). Longitudinal studies examining multi-modal dimensions of the response to trauma (e.g., symptoms, cognitive functions, brain structure and functioning) are optimally suited to detect the underlying neuro-behavioral moderators of non remitting PTSD (Ben-Zion et al., 2019a; 2018; Pitman et al., 2012; Shalev et al., 2017).

Converging neuroimaging studies have suggested abnormalities in brain regions involved in emotional processing in individuals with PTSD (Etkin and Wager, 2007; Shin and Liberzon, 2010). These include emotional reactivity and salience processing abnormalities in the amygdala, insula, and dorsal anterior cingulate cortex (ACC), and emotion-regulation and contextual processing abnormalities in medial and lateral

\* Corresponding author.

E-mail address: [liorwolf@gmail.com](mailto:liorwolf@gmail.com) (L. Wolf).

<sup>1</sup> Equal contributors.

prefrontal cortices, and ventral ACC (Etkin et al., 2015; Pitman et al., 2012; Rabinak et al., 2011; Shalev et al., 2017; Sripada et al., 2012a; 2012b).

Nevertheless, these findings come mainly from cross-sectional studies comparing a group of PTSD patients to a group of trauma-exposed individuals who did not develop PTSD. Moreover, the dynamics in clinical symptoms observed through the first year after trauma points to associated changes over time in underlying neural processing (Bryant et al., 2015; Galatzer-Levy et al., 2013; Shalev et al., 2019). To date, in studies attempting classification of PTSD, prospective data driven investigations have been scarce, and mainly relied on hypotheses regarding known brain abnormalities in PTSD. Previous large scale studies have mostly utilized resting state fMRI (rs-fMRI) or structural MRI (for example; (Brown et al., 2014; Rabinak et al., 2011)), yet studies of whole brain data-driven analysis both at rest and during different tasks, are still needed (Ben-Zion et al., 2020; DiGangi et al., 2016; King et al., 2016; Wang et al., 2020). Investigating brain connectivity during task in addition to rest could enhance the validity of finding with respect to real life mental processing.

In addition, to date, work on PTSD classification has largely overlooked the dynamics of evolving PTSD symptom trajectories and considered single time-point outcomes. Some longitudinal studies (Kendrick et al., 2018; Mason et al., 2002; Qi et al., 2018; Richmond et al., 2011; Scher et al., 2017; Shalev et al., 2019) yielded both group and individual level PTSD risk prediction based on early behavioral and psychological measures, yet, the question of underlying neurobehavioral mechanisms remains open. Not surprisingly, meta-analyses (Brewin et al., 2000; Ozer et al., 2003) and systematic reviews (Brewin, 2005; Brewin et al., 2000; Heron-Delaney et al., 2013; Ozer et al., 2003) have focused on group-level risk indicators without a clear path to clinical implementation at the individual level (Heron-Delaney et al., 2013). All of these gaps, warrant the current prospective investigation of underlying brain mechanisms using our novel end-to-end neural network on resting state and task fMRI data obtained closely following trauma exposure.

Another challenge in examining the trajectory of PTSD following trauma is the heterogeneity of clinical symptoms. According to DSM-5, PTSD consists of four symptom clusters: Intrusion (Criteria B), Avoidance (Criteria C), Negative alterations in mood and cognition (Criteria D), and Hyperarousal (Criteria E); raising the possibility that different symptom clusters may reflect different mechanisms/processes involved. It is therefore of the utmost importance to further consider, not only PTSD diagnosis but also the specific symptom cluster or constellation. However, the majority of studies finding qualitatively distinct profiles among diagnosed individuals have not been utilized to predict the four symptom clusters of PTSD. In this work, we fill this gap by training our deep learning model to identify distinct, symptom-based clusters of PTSD according to DSM-5 criteria.

Machine learning (ML) approaches are increasingly utilized to overcome the problem of characterization, prediction, and treatment selection for individuals suffering from a variety of psychiatric disorders (Schultebraucks and Galatzer-Levy, 2019). For example, Galatzer-Levy et al. (2017) employed a support vector machine (SVM) for prediction of PTSD trajectories of trauma survivors and Rosellini et al. (2018) used an ensemble of multiple classifiers, including random forests (Breiman, 2001), SVM, and regularized regression, to develop a risk score for earthquake survivors. Recently, a growing number of studies have applied ML methodology on neuroimaging data to predict and characterize a variety of psychiatric disorders (Bleich-Cohen et al., 2014; Haller et al., 2014; Koutsouleris and Kambeitz, 2016; Liu et al., 2012; Mikolas et al., 2016; Mourao-Miranda et al., 2012; van der Ploeg et al., 2016; Rive et al., 2016; Zeng et al., 2012), including PTSD (Galatzer-Levy et al., 2017; Gong et al., 2014; Gradus et al., 2017; Jin et al., 2017; Karstoft et al., 2015; Liu et al., 2015; Omurca and Ekinci, 2015; Saxe et al., 2017; Schultebraucks and Galatzer-Levy, 2019). To date, however, no studies have applied a single ML method on both resting-state and task-based neuroimaging data to classify PTSD diagno-

sis simultaneously at three different time-points, using brain data only from the first time-point.

With the advent of deep learning (DL) methods, a variety of neural network approaches have been applied for fMRI classification, including autoencoders (Patel et al., 2016), Recurrent Neural Networks (Dakka et al., 2017) and Convolutional Neural Networks (CNNs) (Bengs et al., 2019; Khosla et al., 2018; Mao et al., 2019; Riaz et al., 2018; Zou et al., 2017). Several recent reports (Bengs et al., 2019; Mao et al., 2019) treated the fMRI data as a 4D volume and used a 4D CNN network to learn spatial and temporal features simultaneously. Connectivity fingerprints, created as multi-channel tensors recording the coupling of each voxel to distinct target Regions of Interest (ROIs), were used as input to a CNN (Khosla et al., 2018). Li et al. (2019) used an inductive graph neural network (GNN) on top of elaborate summary vectors to obtain ASD (Autism Spectrum Disorders) classification. Other methods used summary statistics per region of interest to perform classification. For example, Riaz et al. (2018) used the pooled activations across multiple regions to compute features and obtain pairwise similarities in activations, which are used for classification. However, none of these deep learning methods have been applied in the field of PTSD to date.

In this work, we used clinical data and neural indices, collected shortly after trauma to predict the course of PTSD severity, including distinct clusters of PTSD-related symptom/variables, during the first year after trauma exposure (Ben-Zion et al., 2019b). We additionally applied survival analysis within our prospective design, to predict the persistence of PTSD. Finally, statistical tests were applied to identify pairwise correlations in our PTSD prediction model which discriminate between PTSD patients and controls. The neural data included fMRI scans collected within one-month after trauma during: resting state, emotional reactivity task (Hariri et al., 2000), and Safe or Risky Domino Choice (SRDC) task (Ben-Zion et al., 2021; Kahn et al., 2002). The clinical diagnosis was obtained using the Clinician-Administered PTSD Scale (CAPS) instrument (Weathers et al., 2018), at one-, six- and fourteen-months post-trauma.

Our novel analytical methodology utilized connectivity (correlation) maps extracted from pairs of brain regions, in a whole-brain analysis. These connectivities were obtained using a novel end-to-end neural network by first computing embedding for every brain region and then further applying an attention mechanism (Vaswani et al., 2017), which allowed to focus on the most informative features of the input. Using the attention mechanism, we aggregated together the embeddings that demonstrated the highest similarities. These combined embeddings were subsequently used to produce whole-brain functional connectivity maps for our classifier.

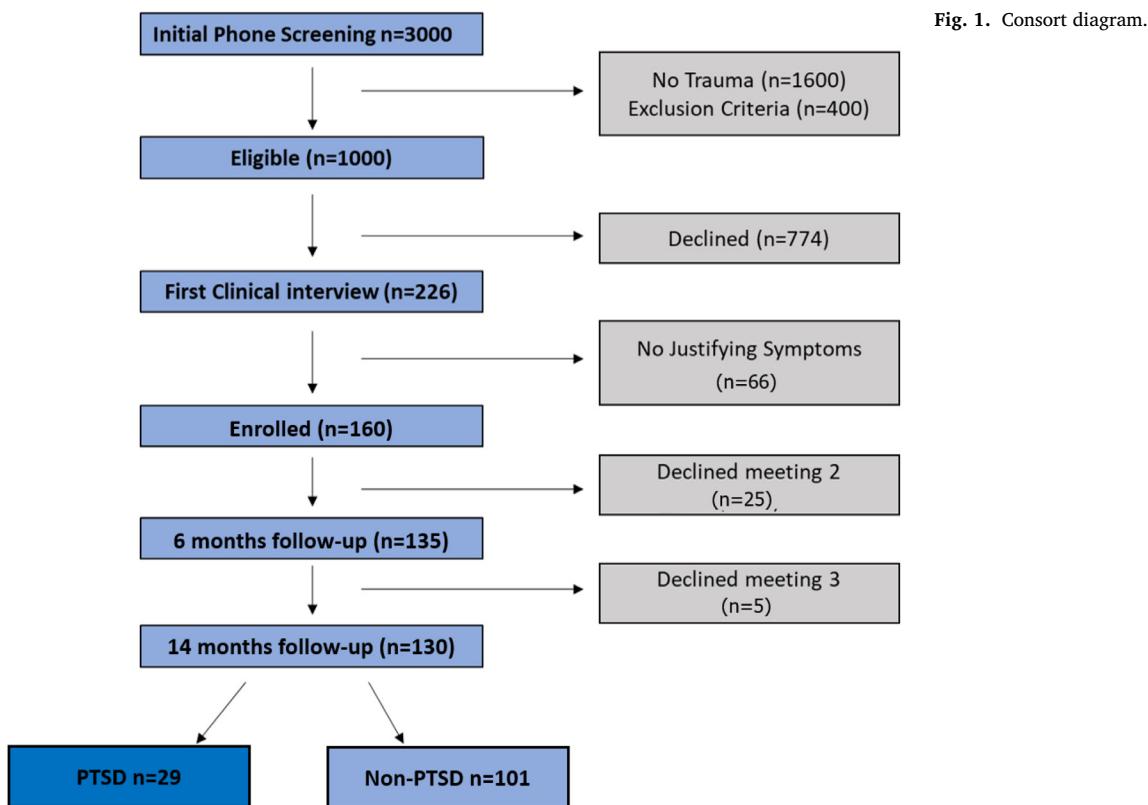
To benchmark our predictive model, we compared the model with previously used ML and DL approaches by training all the techniques from scratch on the same dataset, showing a significant improvement on all the tasks, both in classification accuracy, AUC (Area Under the Curve) and average precision.

To our knowledge, this is the first report of deep learning method used to predict PTSD diagnosis and symptom clusters (based on DSM-5), in a prospective design, using neural data obtained one-month following trauma. While we considered rs-fMRI to be the most common modality used in this type of analysis (Sripada et al., 2012a; Yuan et al., 2018), we reproduced the high discriminatory performance of our method by also employing task-based fMRI. Thus, the predictive power of our method as well as its superiority over the algorithmic baselines, was demonstrated multiple times.

## 2. Methods

### 2.1. Participants

One hundred seventy-one adults (87 women, mean age 34.22 years, range 18–65 years) who were admitted to a general hospitals emergency department after a traumatic event underwent clinical assessments and



fMRI scans, at one-, six- and fourteen-months following trauma (T1, T2, and T3, respectively). Eleven participants were excluded from the analysis at T1, due to partial or low-quality data, resulting in a final dataset of  $n=160$  individuals with valid fMRI and clinical data, obtained one-month after trauma (T1) (see "Enrolled" in Fig. 1). Participants PTSD status and severity were determined by the CAPS (Blake et al., 1995; Weathers et al., 2018), a structured clinical interview corresponding to DSM-based PTSD criteria as determined by the dimensions of frequency, intensity, and severity of symptoms. The results shown by Stein et al. (2014) argue for a broad definition of PTSD defined by any one of the different systems to capture all clinically significant cases of PTSD in future studies. Thus, based on the recommendations to use a "broad diagnostic approach", individuals were included if they met PTSD diagnosis by either: (i) DSM-IV, (ii) DSM-5, or (iii) CAPS-IV total score of 40 or greater.

Out of the  $n=160$  enrolled at T1,  $n=135$  and  $n=130$  completed clinical follow-up assessments at T2 and T3 (see Fig. 1). Twenty-five participants were lost to follow-up between T1 to T2, and an additional five participants were lost to follow-up between T2 and T3. While loss to follow-up in longitudinal studies is a well-known, often unresolved challenge, in this study it was relatively low. Importantly, no significant differences in age, gender, or initial symptom severity (CAPS total scores) were found between participants who completed the study and those who dropped out between T1 to T2 or between T2 to T3 ( $p>0.05$  for all comparisons). Another challenge is imbalance label distribution, i.e., class imbalance: 72% of the subjects met criteria for PTSD diagnosis in T1, 29% in T2 and 23% in T3. We address the imbalance in Section 2.5.2. In Table 6 we present: (a) The number of participants, at each time-point, including their PTSD status in the test and training set (b) The number of participants, in each time-point, that meet the criteria for each symptom cluster in the test and training set as described in Section 2.5.5. It is important to note that our model consists of training data only from the first time-point, and labels (PTSD/symptom cluster diagnosis) from all the three time-points.

**Fig. 1.** Consort diagram.

## 2.2. fMRI tasks

The data were collected and analyzed separately for patients performing three different fMRI paradigms: (i) Resting-state fMRI (rs-fMRI) scan. A 10 min rs-fMRI scan in which participants were instructed to keep their eyes open, focusing on a fixation cross. (ii) Emotional Reactivity Task (Hariri et al., 2000) Here, participants were asked to select the face/shape that matched target face/shape, as accurately and quickly as possible. The task included four blocks of shapes, and four blocks of emotional faces (angry, fearful, surprised and neutral faces). (iii) Safe or Risky Domino Choice task (SRDC) (Kahn et al., 2002) Participants played an interactive competitive Domino game for fourteen minutes, in which they were instructed to take risks in order to win. This task was previously validated and performed in both healthy and clinical populations (Admon et al., 2013; Assaf et al., 2009; Ben-Zion et al., 2021; Goenon et al., 2012; Hyatt et al., 2012; Kahn et al., 2002; Thaler et al., 2019). It involves decision-making (goal-conflict behavior), execution (risky vs. safe choice), anticipation (emotional regulation), and response to an outcome (punishment, non-punishment, reward, non-reward) and contains both visual and auditory cues. In this work, we examined the functional connectivity during the tasks, without considering the different conditions (i.e. without employing any hypothesis on the analysis).

## 2.3. fMRI acquisitions

For all participants, whole-brain functional and anatomical images were acquired using a 3.0 Tesla Siemens MRI system (MAGNETOM Prisma, Germany with a 20-channel head coil at our lab in Tel-Aviv Sourasky Medical Center. Functional images were acquired in an interleaved order (anterior to posterior), using a T2\*-weighted gradient-echo planar imaging pulse sequence (TR/TE=2000/28ms, flip angle= 90°, voxel size  $2.2 \times 2.2 \times 2.2$ mm, FOV= $220 \times 220$ mm, slice thickness=3mm, 36 slices per volume). A T1-weighted three-dimensional anatomical image was also collected, using a magnetization prepared rapid gradient echo (MPRAGE) sequence

**Table 1**

The number of participants in each time-point for a typical train/test split. The assignment between train and test is random at the patient level, and each patient may or may not participate in later on time-points, leading to variability between splits.

	Training			Test			Training			Test			
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	
Total	130	108	107	30	27	23	Total	130	108	107	30	27	23
PTSD	92	30	23	23	9	6	B	105	55	35	23	13	9
							C	97	47	32	22	14	8
							D	90	48	20	25	9	8
							E	96	55	43	22	15	9
(a)							(b)						

(TR/TE=2400/2.29ms, flip angle = 8°, voxel size 0.7×0.7×0.7 mm, FOV = 224 × 224 mm), enabling optimal localization of the functional effects.

#### 2.4. fMRI pre-processing

Pre-processing was performed using FMRIPREP 1.4 (Esteban et al., 2018), and further statistical analysis was performed using SPM12 (Friston, 2003). This process includes: (i) Slice time correction (ii) Head motion correction by six-parameter rigid body spatial transformations, (iii) A 4th-degree interpolation for detection and correction head motions (iv) Co-registration of the functional to corresponding structural maps using the normalized mutual information (NMI) objective, (v) Parcellation using the probabilistic Harvard-Oxford cortical and subcortical structural atlases (including 48 cortical and 21 subcortical areas, in both hemispheres; total of  $N = 117$  regions per participant), and (vi) Extraction of time courses for each participant individually. For each subject, activation levels of 117 brain areas were derived across task conditions. This time course included 300,450,195 samples (TR=2sec) for rs-fMRI, SRDC task and emotional reactivity task, respectively.

#### 2.5. Analytics

Our analysis is conducted along four different axes. The first axis is predicting PTSD diagnosis at three different time-points. Here, the same model was trained on rs-fMRI and two different fMRI tasks. The second axis studies symptom based clusters of PTSD. We trained our model on rs-fMRI to identify the distinct classes of PTSD. In clinical use, it is beneficial to predict remission from PTSD. To this end, the third axis is a survival analysis within our prospective design, which predicts the chronicity of PTSD from T1 to T3. In particular, whether patients with PTSD at T1 still meet the same diagnostic criteria at T3. Finally, for the fourth axis, we applied statistical tests to identify pairwise correlations in our PTSD predicting model (the first axis) that discriminate between PTSD patients and controls.

All computational experiments employed a cross-validation scheme, in which five random splits of the dataset to train and test were created (number of participants in each time-point in training and test is reported in Table 6). In each iteration of the re-sampling procedure, 80% of the data were used as train and 20% as test. To benchmark our predictive model, we compared the model with recent ML and DL approaches by training all the baselines from scratch on our dataset, showing a significant improvement on all the fMRI tasks, both in classification accuracy, AUC and average precision. To demonstrate the contribution of the various components of our method and the different hyper-parameters in the network, we have conducted an ablation analysis in which we considered variants of the model and alternatives to several components of the network.

##### 2.5.1. Network architecture

The training dataset  $\mathcal{X} = \{(x_1, y_1), \dots, (x_{n'}, y_{n'})\}$  consists of  $n'$  participants, where  $x_i$  represents time-series signals and  $y_i$  is a matching label

in  $\{0, 1\}^l$  where  $l = 3$  for PTSD diagnosis prediction (see Section 2.5.4),  $l = 20$  for PTSD symptom clusters prediction (see Section 2.5.5) and  $l = 1$  for survival analysis (see Section 2.5.6). As mentioned in Section 2.4, each fMRI scan is parcellated into  $N = 117$  brain regions, each associated with a 1D sequence of  $T = 300$  data points for rs-fMRI,  $T = 195$  for emotional reactivity task, and  $T = 450$  for SRDC task. A single sample  $x_i = (s_1, s_2, \dots, s_N)$  thus contains the information of multiple brain regions, where each brain region  $i$  is associated with a 1D sequence  $s_i \in \mathbb{R}^T$  obtained by aggregating the activations of this region. Our novel attention approach includes six stages presented in Fig. 2 and can be described by the following set of equations, which are applied to all regions  $i \in [N]$ :

1. Region embedding Given time-series signal  $s_i$ , the feature extraction network  $E$  computes the region embedding  $e_i \in \mathbb{R}^d$  of the brain region  $s_i$ . Unless otherwise specified, we employ an embedding dimension of  $d = 32$ :

$$e_i = E(s_i) \quad (1)$$

The feature extractor network  $E$  includes six 1D convolutional layers, each is followed by ReLU non-linearity. For all the convolutional layers, the kernel size is 3 and the padding is 1. The first convolutional layer includes 64 filters, the second 128 filters, the third and the fourth 256 filters, and the two last convolution layers include 512 filters. In addition to ReLU non-linearity, all the convolutional layers, except of the third and the fifth, are followed by Max pooling with kernel size 2 and stride of 2. The last layer is linear layer that outputs  $d$  features for every brain region.

2. Pairwise connectivity map The similarity network  $F$  computes pairwise correlation  $k_{ij} \in [0, 1]$ , represents the connection strength between brain regions embeddings  $e_i, e_j$ :

$$k_{i,j} = F([e_i, e_j]) \quad (2)$$

where  $[a, b]$  is the concatenation of vectors  $a$  and  $b$ . We apply this network to all the different pairs of brain region embeddings, resulting in the connectivity matrix  $K = [k_{ij}] \in \mathbb{R}^{N \times N}$ . The similarity network  $F$  is a MLP with three linear layers followed by a sigmoid.

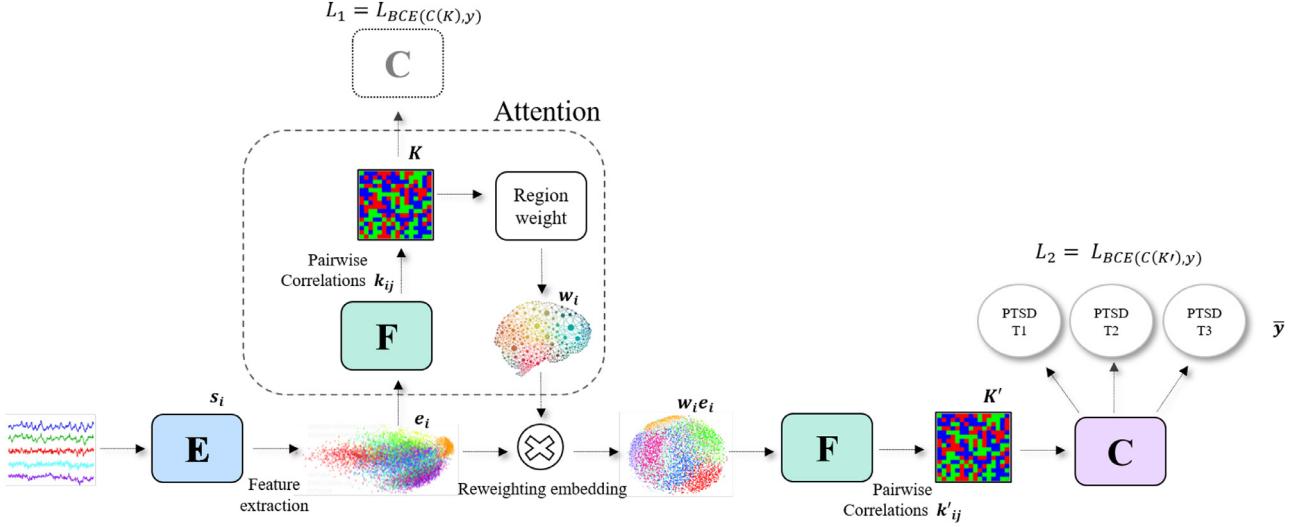
3. Attention mechanism The attention weight  $w_i$  is computed by the relative importance of each outgoing interaction, as reflected by the pairwise correlations. Thus, the region attention weight is obtained for each region  $i$ , by aggregating the scores  $k_{ij}$  across all other regions  $j$ :

$$w_i = \beta + (1 - \beta) \sum_{j \neq i} k_{ij} \quad (3)$$

where  $k_{ij}$  is viewed as an attention value and  $\beta$  is a weighting parameter.

4. Reweighting embedding The weights are used to produce a new set of brain region embeddings. The updated embedding  $e'_i = w_i e_i$  leverages the interactions of the region with other regions, i.e., reweighted embedding  $w_i e_i$  is employed instead of the embedding  $e_i$ .

5. Updated connectivity map Computing a second matrix of similarity scores  $K' = [k'_{ij}] \in \mathbb{R}^{N \times N}$ , between the updated embeddings  $e'_i =$



**Fig. 2.** Illustration of the inference and train stages in PTSD diagnosis prediction: (i) At inference, given the time-series signal, the feature extractor  $E$  outputs an embedding for every brain region. The network  $F$  receives as input two concatenating embeddings and computes their connectivity score. Applying  $F$  to all the pairs of brain regions results in the connectivity matrix  $K$ . The attention mechanism computes region weight by aggregating the scores  $k_{ij}$ . The embedded brain regions features are updated by reweighting of the embedding  $e_i$ . A second connectivity matrix  $K'$  of pairwise correlations  $k'_{ij}$  is computed by  $F$ . Finally, prediction is performed by the classifier  $C$ . (ii) At training, the classification network  $C$  is used twice: once for  $K'$  and once for  $K$ .

$w_i e_i$  and  $e'_j = w_j e_j$  using the same network  $F$ :

$$k'_{ij} = F([w_i e_i, w_j e_j]) \quad (4)$$

where  $k'_{ij} \in [0, 1]$ . Similar forms of aggregations can be found in many attention schemes (Schwartz et al., 2019) (see Section 3 for further analysis).

**6. Classification** Performing prediction by the classifier network  $C$ , based on the connectivity matrix  $K'$ . The probability threshold for classification is set to 0.5.

$$\bar{y} = C(k') \quad \text{where} \quad \bar{y} \in \{0, 1\}^l \quad (5)$$

The classifier  $C$  has three linear layers, the first two are followed by ReLU and Dropout and the last one outputs  $l$  logits.

### 2.5.2. Training objective

Training employs the binary cross entropy with logits multi-label loss (BCE), which treats each of the  $l$  labels as a separate binary classification problem. As presented in Fig. 2, It is used twice: once on the connectivity matrix  $K$ , using classifier  $C$  (Eq. (5)), and once on the updated matrix  $K'$ . As mentioned in 2.5.1,  $y, C(k) \in \{0, 1\}^l$  where  $l = 3$  in PTSD diagnosis prediction,  $l = 20$  in PTSD symptom clusters prediction and  $l = 1$  in survival analysis.

$$\mathcal{L}_1 = -\frac{1}{|\mathcal{X}|} \sum_{(x,y) \in \mathcal{X}} [y \log \sigma(C(k)) + (1-y) \log(1 - \sigma(y, C(k)))], \quad (6)$$

$$\mathcal{L}_2 = -\frac{1}{|\mathcal{X}|} \sum_{(x,y) \in \mathcal{X}} [y \log \sigma(C(k')) + (1-y) \log(1 - \sigma(y, C(k')))] \quad (7)$$

$$\mathcal{L} = \lambda \cdot \mathcal{L}_1 + (1-\lambda) \cdot \mathcal{L}_2, \quad (8)$$

where  $K$  and  $K'$  are functions of  $x$  (Eq. (2),(4)),  $\sigma$  is sigmoid function, and  $\lambda$  is a tradeoff parameter. To overcome in training the imbalance in label distribution described in Section 2.1, we used weight balancing which alters the weight that each training sample carries when computing the loss. For each time-point, the weight was proportional to the ratio between the number of subjects with PTSD and the total number of subjects in this time-point.

### 2.5.3. Training hyperparameters

We used the Adam optimizer with  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ , and an initial learning rate of 0.00001. The learning rate is decreased by 10 every 30 epochs. In all experiments, the number of epochs is 100, we used a batch size of 20, loss weight  $\lambda = 0.6$  and attention value  $\beta = 0.9$ .

### 2.5.4. Predicting PTSD diagnosis

Our primary analysis concerned the prediction of PTSD diagnosis at three different time-points. This problem is treated as a three-label classification and the network architecture described in Section 2.5.1 is used with  $l = 3$ . The training dataset  $\mathcal{X} = \{(x_1, y_1), \dots (x_{n'}, y_{n'})\}$  consists of  $n'$  participants, where  $x_i$  represents time-series signals and  $y_i$  is a matching label in  $\{0, 1\}^3$ . Here, we treat each of the three time-points as a separate binary classification problem. Although we consider rs-fMRI to be the most suitable modality, to evaluate our model due to its ubiquitous nature and since it is not stimuli-dependent, we additionally trained the proposed architecture on different fMRI tasks, as mentioned in Section 2.2.

### 2.5.5. Identifying unique PTSD symptom clusters

According to DSM-5, PTSD includes four symptom clusters: Intrusion (criteria B) consisting of five symptoms; Avoidance (criteria C)-two symptoms; Negative alterations in mood and cognition (criteria D)- seven symptoms; and Hyperarousal (criteria E)- six symptoms. The classification is based on PTSD diagnosis (by DSM-5): meeting criteria B if there is at least one of the five symptoms; criteria C- if at least one of the two symptoms, criteria D- if at least two of the seven symptoms; and criteria E- if at least two of the six symptoms.

To identify distinct, symptom-based clusters of PTSD, we trained our model, each time separately, on rs-fMRI dataset. The network architecture described in Section 2.5.1 is used independently for each time-point  $T1, T2, T3$  with  $l = 20$ . The training dataset  $\mathcal{X} = \{(x_1, y_1), \dots (x_{n'}, y_{n'})\}$  consists of  $n'$  participants, where  $x_i$  represents time-series signals and  $y_i$  is a matching label in  $\{0, 1\}^{20}$ , for 20 symptoms in each time-point. In other words, for each time-point T1,T2 and T3, the label corresponding to each participant is a 20 dimensional binary vector  $\{0, 1\}^{20}$ , such that each coordinate represents one PTSD symptom. The first five coordinates correspond to the five symptoms of criteria B, the next two to the symptoms of criteria C and so on. The probability threshold for each

symptom is set to 0.5, and the classification to each cluster is based on DSM-5 clinical diagnosis. The results are reported in [Section 3.2](#).

### 2.5.6. Survival analysis

In clinical use, it is beneficial to predict remission from PTSD. Starting with a trauma-exposed group with high rates of PTSD at T1 (about 70%), we observe recovery over time (from T1 to T2, and from T2 to T3) resulting in a final group with mostly recovered healthy individuals and only a subset of PTSD patients (about 23%). We therefore aim to explore the chronicity of the PTSD symptoms from T1 to T3.

To this end, we trained our model to predict whether patients of each symptom cluster would still meet the same diagnostic criteria at T3. In this case, the network architecture described in [Section 2.5.1](#) is used with  $l = 1$ . The training dataset  $\mathcal{X} = \{(x_1, y_1), \dots (x_{n'}, y_{n'})\}$  consists of  $n'$  participants, where  $x_i$  represents time-series signals and  $y_i$  is a matching label in  $\{0, 1\}$  indicating whether the patient still meet the same diagnosis at T3.

### 2.5.7. Identifying discriminative connectivities

To identify discriminative connectivities, we compared the pairwise connectivities ( $k'_{ij}$ ) between participants with and without a diagnosis of PTSD, at each one of the three time-points. Independent t-tests were used to detect connectivities that significantly differed between the two groups, with a statistical threshold of alpha=0.05 and FDR-correction for multiple-comparisons (Benjamini and Hochberg, 1995). The results are reported in [Section 3.4](#).

## 2.6. Ablation analysis

To demonstrate the advantage of learning a single representation for multiple time-points, we trained an alternative method that examines each time-point separately (a separate binary model is trained for each time-point). We performed an ablation analysis that demonstrates the contribution of the various components of our method, by employing multiple method variants of it. First, a model in which the network  $E$  is the identity transformation was tested, i.e. we replaced the 32 dimension embedding with the time-series signal input. The second variant we employed was a model in which  $F$  was replaced by a Pearson correlation. Last, to examine the contribution of attention mechanism, we considered several alternatives to our reweighting approach. (i) **A model based on the first connectivity map** This model includes only networks  $E$ ,  $F$  and classifier  $C$  applied on matrix  $K$ . It is similar to deep fMRI (Riaz et al., 2018), however it differs in the architecture of the networks. (ii) **Graph convolutional neural network method as aggregation function** This model employs a graph convolutional neural network (Niepert et al., 2016) (GCN) instead of our proposed reweighting in [Eq. \(3\)](#). Based on a message passing neural network, the GCN method learns new representations by aggregating feature vectors of the neighboring regions. We tested several architectures and report the best one.

## 3. Results

### 3.1. PTSD diagnosis prediction

We report the mean and the standard deviation of the accuracy for each of the three time-points and prediction tasks as well as the AUC and average precision values from the prediction scores. The same experiments are also repeated when minimizing the mean squared error (MSE) for the CAPS score, instead of the binary PTSD classification. Since there is a high correlation between CAPS-IV and CAPS-5 total scores across all time-points ( $r>0.95$ ), we report only CAPS-IV total scores in this section. CAPS-IV had been extensively used in neuroimaging studies of PTSD to date, and we report it here in order to keep continuity.

We consider two recently reported approaches: The deep fMRI (Riaz et al., 2018) network and a slightly modified version of the graph

neural network approach of Li et al. (2019) (best effort reimplementation). Since Random Forest (RF) is a popular method in the literature for fMRI classification, we employed RF with 100 estimators (several options have been tested and the one with the best performance is reported), where the input was the concatenation of the data from all regions at all time-points.

The accuracy results of our model and baselines are reported in [Tables 2–4](#) for rs-fMRI, emotional reactivity task, and SRDC task (Kahn et al., 2002), respectively. Furthermore, the AUC binary classification rates, and average precision rates, for all the tasks and baselines are reported in [Tables 5 and 6](#), respectively.

**Ablation analysis** As described in [Section 2.6](#), we employed multiple variants for our method (see results in [Table 2](#)). First, we trained an alternative method that examines each time-point separately (a separate binary model was trained for each time-point). Although the alternative model yielded accurate rates of 81.0% in T1, 77.0% in T2 and 80.0% in T3, it can be seen from [Table 2](#) that utilizing the prospective design to predict the dynamic PTSD diagnosis leads to a higher accuracy. The second variant, in which network  $E$  is the identity transformation, yielded accuracy rates of 70.5% in T1, 51.0% in T2 and 59.1% in T3, demonstrating the importance of using the feature extraction network  $E$ . Another variant of our model, in which  $F$  is replaced by a Pearson correlation, produced accuracy rates of 73.3% in T1, 51.0% in T2 and 59.1% in T3, demonstrating the importance of using the similarity network  $F$  to compute the pairwise connectivity map.

Furthermore, we considered several alternatives to our reweighting approach (see [Section 2.6](#) for full details): (i) **A model based on the first connectivity map** This model yielded results of 72.0% in T1, 56.6% in T2 and 57.5% in T3, demonstrating the importance of updating the connectivity map  $K$  using an attention mechanism. (ii) **Graph convolutional neural network method as aggregation function** We can observe from [Tables 2, 3, 4, 5](#) the high accuracy and the high AUC values respectively, of the GCN variant on rs-fMRI data and all the tasks. In particular, we can observe high performance in T1 (Accuracy= 82.1%, AUC=0.78 on rs-fMRI, Accuracy=84.1%, AUC=0.77 on emotional reactivity task and Accuracy=79.1%, AUC=0.64 on SRDC task), and reasonably good performance on T3 (Accuracy=72.4%, AUC=0.74 on rs-fMRI, Accuracy=66.5%, AUC=0.72 on emotional reactivity task and Accuracy=67.9%, AUC=0.72 on SRDC task).

A parameter sensitivity analysis was performed by varying the value of the two parameters  $\beta$  ([Eq. \(3\)](#)) and  $\lambda$  ([Eq. \(8\)](#)) to validate the stability of the method. The results for the first parameter, given in [Fig. 3](#), show that the method is largely insensitive to  $\beta$  for rs-fMRI, and more sensitive in the two other tasks. For  $\lambda = 1$ , the loss  $\mathcal{L}$  becomes  $\mathcal{L}_1$  ([Eq. \(6\)](#)), making use of the first connectivity map  $K$  only, even though the second connectivity map  $K'$  is still computed. However, at inference the classification is done by  $K'$ .

Our method is largely stable to  $\lambda$  for the rs-fMRI data, with the exception of a slight trade-off between the accuracy at T2 and T3 (see [Fig. 4](#)). For the other tasks, increasing  $\lambda$  up to a certain point improves accuracy, but as  $\lambda$  increases beyond that point ( $\approx 0.2$ ), the performance decreases. We note that for the emotional reactivity task and SRDC task, the optimal parameters are not the ones shown in the results tables, and that their performance can be improved (this is the result of selecting the parameters early on in the development process). Specifically, setting  $\lambda = 0.2$  in SRDC task highly improves the performance.

### 3.2. PTSD symptom cluster prediction

The classification accuracy rates as well as AUC values, for PTSD symptom cluster prediction, of our model and alternative Deep fMRI method (Riaz et al., 2018), are reported in [Table 6](#). The number of participants at each time-point, including their division to different symptom clusters is reported in [Table 8](#). In addition, the AUC for each symptom prediction is reported in [Table 6](#). We obtain a high predictive ability for PTSD symptom cluster prediction with an average AUC=

**Table 2**Resting-state prediction. Accuracy of binary classification and MSE of CAPS (mean  $\pm$  SD).

Method	Classification accuracy			CAPS-IV MSE		
	T1	T2	T3	T1	T2	T3
Deep FMRI <a href="#">Riaz et al. (2018)</a>	74.0 $\pm$ 6.6	59.0 $\pm$ 12.0	43.4 $\pm$ 9.0	0.5 $\pm$ 0.2	0.5 $\pm$ 0.2	0.3 $\pm$ 0.2
Li et al. <a href="#">Li et al. (2019)</a>	72.0 $\pm$ 7.0	62.9 $\pm$ 5.0	76.2 $\pm$ 6.0	0.7 $\pm$ 0.3	0.9 $\pm$ 0.7	1.1 $\pm$ 0.8
Random Forest <a href="#">(Li et al., 2019)</a>	68.3 $\pm$ 7.6	63.4 $\pm$ 9.0	73.2 $\pm$ 6.2	0.4 $\pm$ 0.3	0.3 $\pm$ 0.2	0.3 $\pm$ 0.2
Raw data MLP	46.5 $\pm$ 20.0	55.2 $\pm$ 14.0	33.6 $\pm$ 25.0	0.7 $\pm$ 0.5	0.4 $\pm$ 0.2	0.3 $\pm$ 0.2
<b>Ours</b>	<b>88.6 <math>\pm</math> 2.0</b>	<b>80.4 <math>\pm</math> 3.0</b>	<b>84.0 <math>\pm</math> 5.4</b>	<b>0.2 <math>\pm</math> 0.0</b>	<b>0.3 <math>\pm</math> 0.1</b>	<b>0.1 <math>\pm</math> 0.1</b>
Ours + GCN	82.1 $\pm$ 5.0	54.6 $\pm$ 6.7	72.4 $\pm$ 8.6	0.4 $\pm$ 0.1	0.3 $\pm$ 0.1	0.2 $\pm$ 0.3
Ours 3	81.0 $\pm$ 3.0	77.0 $\pm$ 3.0	80.0 $\pm$ 1.0	0.9 $\pm$ 0.2	0.4 $\pm$ 0.1	0.1 $\pm$ 0.0
independent						
Ours no reweighting	72.0 $\pm$ 7.0	56.6 $\pm$ 11.0	57.5 $\pm$ 13.0	0.4 $\pm$ 0.1	0.4 $\pm$ 0.2	0.6 $\pm$ 0.4
Ours Pearson as F	73.3 $\pm$ 7.0	55.3 $\pm$ 9.0	51.0 $\pm$ 10.0	1.1 $\pm$ 0.1	0.4 $\pm$ 0.2	0.6 $\pm$ 0.2
Ours E =identity	70.5 $\pm$ 7.0	51.0 $\pm$ 9.0	59.1 $\pm$ 10.0	1.1 $\pm$ 0.6	1.0 $\pm$ 0.5	0.9 $\pm$ 0.6

**Table 3**

Prediction based on Emotional reactivity task.

Method	Classification accuracy			CAPS-IV MSE		
	T1	T2	T3	T1	T2	T3
Deep FMRI ( <a href="#">Riaz et al., 2018</a> )	79.3 $\pm$ 3.5	37.0 $\pm$ 5.4	32.2 $\pm$ 11.2	0.9 $\pm$ 0.2	1.2 $\pm$ 0.6	0.7 $\pm$ 0.2
Li et al. <a href="#">Li et al. (2019)</a>	79.8 $\pm$ 4.6	64.0 $\pm$ 8.1	70.8 $\pm$ 13.5	1.1 $\pm$ 0.1	1.3 $\pm$ 0.4	0.9 $\pm$ 0.5
Random Forest ( <a href="#">Li et al., 2019</a> )	74.9 $\pm$ 3.7	59.0 $\pm$ 12.0	68.5 $\pm$ 7.1	1.0 $\pm$ 0.1	1.4 $\pm$ 0.4	1.0 $\pm$ 0.1
Raw data MLP	74.9 $\pm$ 6.5	60.9 $\pm$ 5.8	59.7 $\pm$ 5.1	1.3 $\pm$ 0.3	1.5 $\pm$ 0.3	1.0 $\pm$ 0.3
<b>Ours</b>	<b>85.8 <math>\pm</math> 3.6</b>	<b>80.0 <math>\pm</math> 11.2</b>	<b>69.3 <math>\pm</math> 9.1</b>	<b>0.8 <math>\pm</math> 0.2</b>	<b>1.0 <math>\pm</math> 0.0</b>	<b>0.7 <math>\pm</math> 0.1</b>
Ours + GCN	84.1 $\pm$ 2.4	69.8 $\pm$ 7.0	66.5 $\pm$ 9.1	0.9 $\pm$ 0.4	1.4 $\pm$ 0.4	0.9 $\pm$ 0.1
Ours no reweighting	80.0 $\pm$ 2.4	53.4 $\pm$ 7.3	40.4 $\pm$ 2.4	0.9 $\pm$ 0.1	1.3 $\pm$ 0.6	1.0 $\pm$ 0.1
Ours Pearson as F	76.6 $\pm$ 8.0	41.4 $\pm$ 11.8	61.8 $\pm$ 11.3	1.4 $\pm$ 0.2	1.6 $\pm$ 0.4	0.9 $\pm$ 0.2
Ours E =identity	76.6 $\pm$ 2.4	61.5 $\pm$ 4.5	67.4 $\pm$ 8.7	0.9 $\pm$ 0.2	1.4 $\pm$ 0.3	0.9 $\pm$ 0.0

**Table 4**

Prediction based on Safe or Risky Domino Choice (SRDC) task.

Method	Classification accuracy			CAPS-IV MSE		
	T1	T2	T3	T1	T2	T3
Deep FMRI ( <a href="#">Riaz et al., 2018</a> )	74.1 $\pm$ 2.4	30.0 $\pm$ 9.6	22.7 $\pm$ 4.6	0.9 $\pm$ 0.0	0.8 $\pm$ 0.2	0.6 $\pm$ 0.1
Li et al. <a href="#">Li et al. (2019)</a>	70.6 $\pm$ 5.4	64.5 $\pm$ 9.8	67.1 $\pm$ 10.0	0.8 $\pm$ 0.3	0.8 $\pm$ 0.2	0.6 $\pm$ 0.2
Random Forest ( <a href="#">Li et al., 2019</a> )	74.1 $\pm$ 2.4	31.0 $\pm$ 10.9	22.8 $\pm$ 4.5	1.0 $\pm$ 0.1	1.0 $\pm$ 0.2	0.8 $\pm$ 0.1
Raw data MLP	71.5 $\pm$ 3.7	52.2 $\pm$ 3.9	62.4 $\pm$ 9.1	1.1 $\pm$ 0.1	1.2 $\pm$ 0.1	0.8 $\pm$ 0.2
<b>Ours</b>	<b>80.7 <math>\pm</math> 1.8</b>	<b>68.3 <math>\pm</math> 10.0</b>	<b>70.4 <math>\pm</math> 7.5</b>	<b>0.8 <math>\pm</math> 0.1</b>	<b>0.7 <math>\pm</math> 0.2</b>	<b>0.6 <math>\pm</math> 0.1</b>
Ours + GCN	79.1 $\pm$ 6.8	46.4 $\pm$ 11.8	67.9 $\pm$ 13.5	0.9 $\pm$ 0.0	0.9 $\pm$ 0.2	0.8 $\pm$ 0.1
Ours no reweighting	79.2 $\pm$ 1.6	55.3 $\pm$ 10.5	58.6 $\pm$ 11.8	0.9 $\pm$ 0.1	1.2 $\pm$ 0.2	0.8 $\pm$ 0.2
Ours Pearson as F	67.5 $\pm$ 4.6	47.1 $\pm$ 16.8	52.1 $\pm$ 13.6	1.5 $\pm$ 0.6	2.0 $\pm$ 0.6	1.1 $\pm$ 0.3
Ours E =identity	66.7 $\pm$ 8.2	44.2 $\pm$ 11.0	52.6 $\pm$ 5.6	1.0 $\pm$ 0.2	1.1 $\pm$ 0.2	0.8 $\pm$ 0.1

**Table 5**AUC of binary classification (mean  $\pm$  SD).

Method	Resting state			Hariri			Domino		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Deep FMRI ( <a href="#">Riaz et al., 2018</a> )	0.51 $\pm$ 0.01	0.73 $\pm$ 0.08	0.69 $\pm$ 0.06	0.59 $\pm$ 0.11	0.70 $\pm$ 0.03	0.72 $\pm$ 0.04	0.53 $\pm$ 0.0	0.63 $\pm$ 0.05	0.63 $\pm$ 0.05
Li et al. <a href="#">Li et al. (2019)</a>	0.69 $\pm$ 0.04	0.72 $\pm$ 0.03	0.74 $\pm$ 0.02	0.71 $\pm$ 0.07	0.75 $\pm$ 0.06	0.75 $\pm$ 0.02	0.70 $\pm$ 0.06	0.74 $\pm$ 0.02	0.76 $\pm$ 0.01
Random Forest ( <a href="#">Li et al., 2019</a> )	0.62 $\pm$ 0.08	0.70 $\pm$ 0.03	0.74 $\pm$ 0.01	0.57 $\pm$ 0.07	0.70 $\pm$ 0.03	0.72 $\pm$ 0.02	0.53 $\pm$ 0.05	0.69 $\pm$ 0.03	0.72 $\pm$ 0.01
Raw data MLP	0.70 $\pm$ 0.04	0.72 $\pm$ 0.02	0.72 $\pm$ 0.05	0.69 $\pm$ 0.04	0.72 $\pm$ 0.03	0.71 $\pm$ 0.03	0.65 $\pm$ 0.08	0.69 $\pm$ 0.02	0.68 $\pm$ 0.06
<b>Ours</b>	<b>0.80 <math>\pm</math> 0.05</b>	<b>0.73 <math>\pm</math> 0.02</b>	<b>0.76 <math>\pm</math> 0.01</b>	<b>0.80 <math>\pm</math> 0.03</b>	<b>0.80 <math>\pm</math> 0.02</b>	<b>0.76 <math>\pm</math> 0.01</b>	<b>0.76 <math>\pm</math> 0.03</b>	<b>0.76 <math>\pm</math> 0.04</b>	<b>0.76 <math>\pm</math> 0.01</b>
Ours+GCN	0.78 $\pm$ 0.09	0.69 $\pm$ 0.05	0.74 $\pm$ 0.02	0.77 $\pm$ 0.03	0.76 $\pm$ 0.03	0.72 $\pm$ 0.04	0.64 $\pm$ 0.08	0.72 $\pm$ 0.03	0.72 $\pm$ 0.02

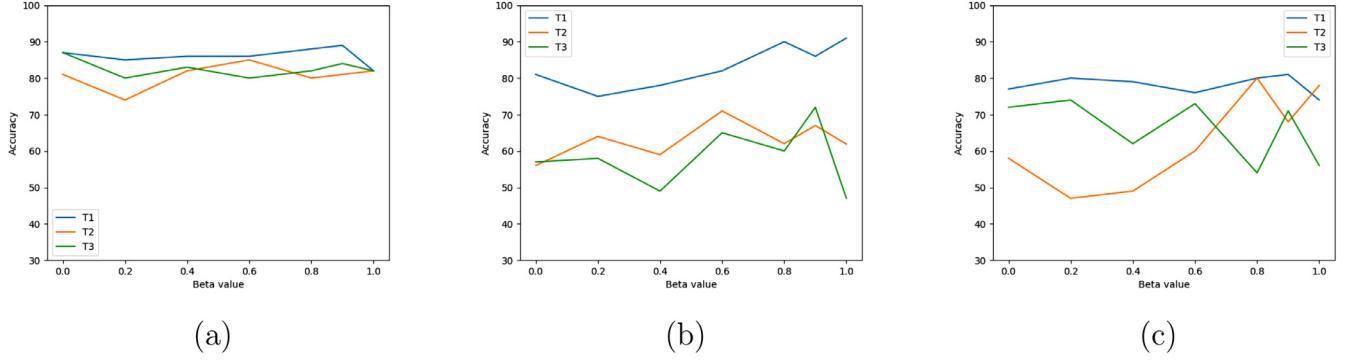
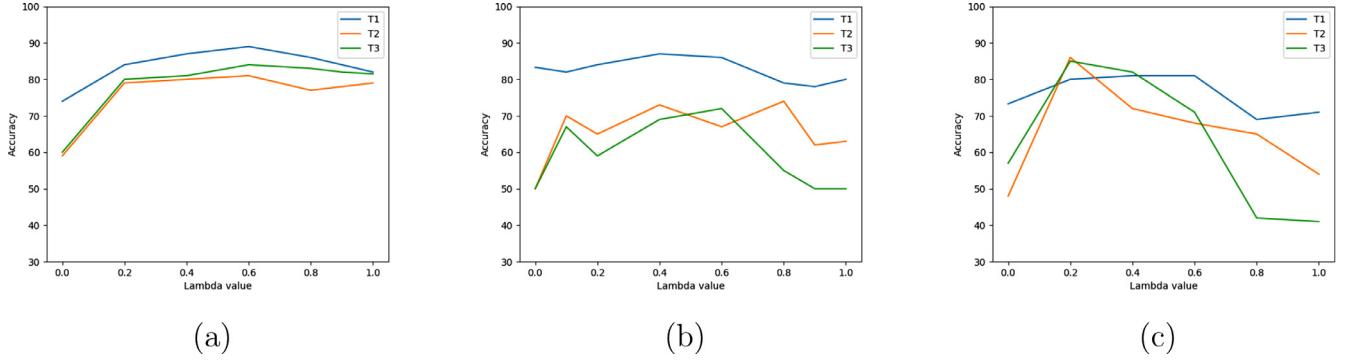
0.81, average Accuracy=81.2% for class B; average AUC= 0.8, average Accuracy=73.0% for class C; average AUC= 0.78, average accuracy of 75.8% for class D. Finally, we obtain average AUC= 0.78 and average Accuracy=71.0% for class E (averaging on the three time-points).

In order to visualize the embedding space, we computed for each participant a dominant cluster of symptoms. This was determined by com-

puting a cluster score for each cluster and participant (i.e. percentage of symptoms the participant meets out of all the symptoms in the cluster), normalizing the scores by considering the participant's percentile among all participants and then choosing, if relevant, the most significant cluster for the participant. In other words, if there is no dominant cluster for the participant, we do not visualize the point and skip to the

**Table 6**Average precision of binary classification (mean  $\pm$  SD).

Method	Resting state			Hariri			Domino		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Deep FMRI (Riaz et al., 2018)	0.50 $\pm$ 0.00	0.64 $\pm$ 0.01	0.65 $\pm$ 0.01	0.50 $\pm$ 0.05	0.65 $\pm$ 0.09	0.65 $\pm$ 0.01	0.50 $\pm$ 0.00	0.64 $\pm$ 0.07	0.63 $\pm$ 0.02
Li et al. Li et al. (2019)	0.71 $\pm$ 0.03	0.68 $\pm$ 0.04	0.68 $\pm$ 0.03	0.60 $\pm$ 0.08	0.66 $\pm$ 0.04	0.66 $\pm$ 0.04	0.61 $\pm$ 0.04	0.67 $\pm$ 0.04	0.71 $\pm$ 0.07
Random Forest (Li et al., 2019)	0.53 $\pm$ 0.03	0.65 $\pm$ 0.02	0.66 $\pm$ 0.02	0.52 $\pm$ 0.01	0.65 $\pm$ 0.01	0.66 $\pm$ 0.00	0.49 $\pm$ 0.01	0.65 $\pm$ 0.02	0.65 $\pm$ 0.01
Raw data MLP	0.58 $\pm$ 0.04	0.65 $\pm$ 0.04	0.65 $\pm$ 0.03	0.55 $\pm$ 0.03	0.66 $\pm$ 0.05	0.68 $\pm$ 0.05	0.55 $\pm$ 0.02	0.66 $\pm$ 0.05	0.66 $\pm$ 0.03
<b>Ours</b>	<b>0.77<math>\pm</math>0.03</b>	<b>0.73<math>\pm</math>0.00</b>	<b>0.76<math>\pm</math>0.04</b>	<b>0.71<math>\pm</math>0.03</b>	<b>0.76<math>\pm</math>0.05</b>	<b>0.77<math>\pm</math>0.01</b>	<b>0.71<math>\pm</math>0.02</b>	<b>0.67<math>\pm</math>0.02</b>	<b>0.79<math>\pm</math>0.05</b>
Ours+GCN	0.73 $\pm$ 0.09	0.65 $\pm$ 0.02	0.70 $\pm$ 0.05	0.70 $\pm$ 0.06	0.63 $\pm$ 0.03	0.67 $\pm$ 0.06	0.66 $\pm$ 0.03	0.66 $\pm$ 0.02	0.68 $\pm$ 0.03

**Fig. 3.** Sensitivity to parameter  $\beta$  of Eq. (3) (a) Resting-state task (b) Emotional reactivity task (c) SRDC task.**Fig. 4.** Sensitivity to the loss tradeoff  $\lambda$  in Eq. (8). (a) Resting-state task (b) Emotional reactivity task (c) SRDC task

next participant. For full pseudo-code see Algo. 1. A visualization of the learned features, before the classification to a specific cluster is shown in Fig. 5. In particular, the features of each participant obtained from the first linear layer in the classifier  $C$ , that takes the connectivity matrix  $K'$  as input, flatten it into a single vector and projects it to feature vector of size 100. Each 100 dimensional feature point is embedded to two-dimensional point using t-SNE method (Maaten and Hinton, 2008), and the color of the visualized point is determined using the participant's dominant cluster of symptoms.

### 3.3. Survival analysis

We used the survival analysis in our prospective design to predict persistence of PTSD. Our model was trained for predicting whether participants still meet the same diagnostic criteria at T3. The fraction of non-remitting participants at T3 was 28/115, i.e., out of 115 participants that met PTSD criteria at T1, 28 met the same criteria at T3.

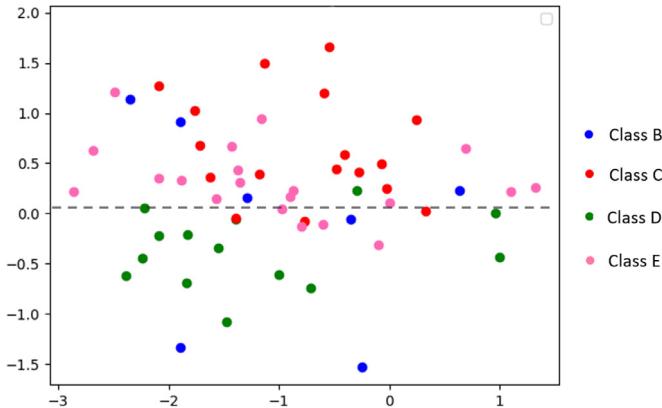
The input to the subsequently trained predictive model was rs-fMRI data of participants, obtained one-month after trauma (T1), and the same cross-validation splits were employed as described in Section 3. A high predictive ability for chronicity of PTSD was obtained, with an  $AUC=0.84 \pm 0.02$  and  $Accuracy = 81.33\% \pm 5.37$ .

### 3.4. Identifying discriminative connectivities

All significant findings of the  $t$ -test identifying the discriminative power of the pairwise similarities ( $k'_{ij}$ ) between PTSD and non-PTSD participants are given in the supplementary material, including both our results and those of the alternative models. Pairwise association analysis revealed several significant functional connectivity patterns, in line with previous PTSD neuroimaging literature, which are discussed below for each time-point.

#### 3.4.1. Neural measures at one-month post-trauma (T1) Clinical status at one-month post-trauma (T1)

*resting-state fMRI* Fifty four significant rs-fMRI connectivities were found. Out of these, all but one includes the amygdala, a key brain region critically involved in emotional processing, both in the healthy population and in psychopathology (Davis and Whalen, 2001; Sripada et al., 2012a). Subset of amygdala connectivities are consistent with PTSD neuroimaging literature (see Fig. 6 ), including amygdala-functional connectivity with (i) bilateral hippocampus, (ii) bilateral medial frontal cortices, (iii) right parahippocampal gyrus,(iv) left insular cortex, (v) bilateral precuneus, (vi) bilateral cingulate gyrus. Moreover, amygdala-parahippocampal gyrus connectivity at T1 also predicted PTSD diagnosis a year later (T3). *Emotional reactivity task* Only three out of the



**Fig. 5.** a t-SNE embedding of the last feature map before the classification to PTSD cluster. The color of each point represents the dominant cluster of the participant. There are 10,13,22,20 subjects with dominant symptoms of class B,C,D,E respectively.

#### Algorithm 1: Symptom cluster Dominance calculation

```

Data: numSymptoms array with the predicted number of
      symptoms for each (cluster,subject);
numTotalSymptoms array with the total number of symptoms each
      symptom cluster consists (Specifically,
      numTotalSymptoms = [5, 2, 7, 6]);
1 for c in clusters do
2   for s in subjects do
3     | scores(c,s) = numSymptoms(c,s)/numTotalSymptoms(c);
4 for c in clusters do
5   let sortedScoresC ← sort(scores(c));
6   for s in subjects do
7     | normalizedScores(c,s)=max( find_indices(scores(c,s) >
      sortedScoresC )/numSubjects;
8 for s in subjects do
9   subjectScores ← normalizedScores(:,s);
10  maxscore1, maxscore2 ← first and second largest score in
      subjectScores;
11  if (maxscore1 - maxscore2 > 0.1) then
12    | visualize dominantCluster with score maxscore1 for subject
      s;
13  else
14    | skip subject s;
```

49 significant connectivities that were found involved the amygdala with (i) right superior temporal gyrus, (ii) left middle temporal gyrus, and (iii) left frontal pole. Nevertheless, as this is a visual processing task involving faces and shapes, we observed various connectivities at T1 known to be involved in visual processing (Hariri et al., 2000; Hendl et al., 2003). For example, connectivities of the inferior temporal gyrus and connectivities of the occipital and temporo-occipital gyri. *SRDC task* Two out of 99 significant connectivities that were found involved the amygdala, (i) right amygdala with right lateral occipital cortex (ii) left amygdala with right lateral occipital cortex. The rest of the connectivities are related to the nature of this dynamic task, as was shown in previous work (Admon et al., 2013; Ben-Zion et al., 2021; Thaler et al., 2019). For example, visual processing involving occipital and inferior temporal areas, auditory processing involving heschls gyrus (Dierks et al., 1999) and frontal areas involve in decision making (Rushworth et al., 2011), anticipation and responses to risk and reward.

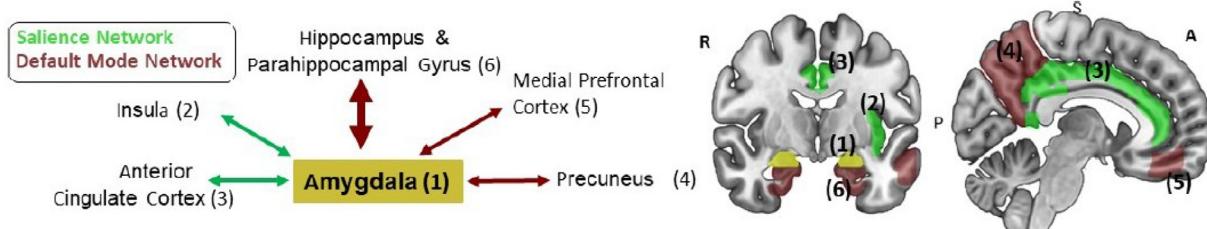
#### 3.4.2. Neural measures at one-month post-trauma (T1) Clinical status at six-months post-trauma (T2)

*Resting-state fMRI (rs-fMRI)* Four rs-fMRI significant connectivities were found. None of them is consistent with the neuroimaging literature of PTSD. *Emotion reactivity task*

As in T1, there are connectivities of the Inferior Temporal gyrus and temporo-occipital gyrus, which are related to the visual processing of this task (Miyashita, 1993). *SRDC Task* The majority of the 135 significant connectivities that were found are highly relevant to the task (Admon et al., 2013; Ben-Zion et al., 2021). Among these, 23 significant connectivities of the accumbens (known to be involved in the processing and analyzing rewarding stimuli) with different relevant regions, such as the hippocampus and parahippocampal (part of the Default Mode-Network) and 33 significant connectivities of the posterior cingulate cortex (related to intrinsic control networks and also part of the Default Mode-Network).

#### 3.4.3. Neural measures at one-month post-trauma (T1) Clinical status at fourteen-months post-trauma (T3)

*Resting-state fMRI (rs-fMRI)* Seven significant connectivities were found, including one which is specifically relevant to PTSD: right amygdala with right parahippocampal gyrus. *Emotion reactivity task* Out of 84 significant connectivities were found, only one involves the amygdala, namely amygdala-brain stem connectivity. Similar to T1 and T2, we observe the dominance of neural connectivities related to visual processing: Almost 50% of the connectivities involve the occipital (visual) cortex and 17 connectivities involve the fusiform cortex/gyrus, known to be the  $\alpha$ face areag (Grill-Spector et al., 2004; Kanwisher et al., 1997). *SRDC task* Out of the 115 significant connectivities that were found, seven connectivities involve the amygdala. Interestingly, ten connectivities involve the insula and nine connectivities involve the anterior cingulate cortex, both are part of the Salience Network (Bonnelle et al.,



**Fig. 6.** Selected rs-fMRI connectivity patterns which significantly discriminated between participants with and without PTSD. This figure presents the connectivity patterns of the amygdala which were significantly associated with PTSD diagnosis at T1. In particular, amygdala-parahippocampal connectivity at T1 also significantly discriminated between individuals with and without PTSD at T3 (marked with a bold arrow). Overall, these regions are part of two established large-scale brain networks: Salience Network (regions 1,2,3 in green) and Default-Mode Network (regions 4,5 and 6 in brown). Numbers 1–5 in the scheme on the left match to numbers 1–5 presented on the brain on the right.

2012). Moreover, hippocampus and parahippocampal gyrus are another dominant hubs for connectivities at this time-point (Raichle, 2015).

#### 4. Discussion

Our novel classification method, across all fMRI scans and three time-points, demonstrated substantial improvement in accuracy, compared to all other methods recently suggested for fMRI analysis. We compared the model with previous ML and DL approaches by training all the baselines on the same dataset, showing a significant improvement on all the tasks, both in classification accuracy, AUC and average precision.

The first and primary axis of our work concerned the prediction of PTSD diagnosis at three different time-points. We considered rs-fMRI to be the most suitable modality as it does not have different conditions (in contrast to fMRI tasks). Nevertheless, we also trained the proposed architecture on different fMRI tasks, which provided further valid information on brain processing related to emotional reactivity, risky behavior, and sensitivity to punishments and rewards, as well as additional analyses pointing to the performance of our method in comparison to other approaches. Indeed, training the model on two different fMRI tasks, in addition to rs-fMRI data, may provide a more comprehensive understanding of neural activity in individuals with PTSD.

Here, we utilized the within-subject prospective design to predict a dynamic PTSD diagnosis for each patient, using neural measures obtained one-month after trauma and clinical status obtained one-, six- and fourteen months following the trauma. As demonstrated in Tab. 2, utilizing such a prospective design to predict the dynamic PTSD diagnosis leads to higher accuracy rates than the alternative model which examines each time-point separately. To demonstrate the contribution of the various components of our method and the different hyper-parameters in the network, we conducted an ablation analysis in which we considered several variants of the model and alternatives to the attention mechanism as well as to the functional connectivity network (see Section 3.1). Our model showed high performance in predicting PTSD diagnosis at three time-points, using fMRI scans only from the first time-point, when training both on rs-fMRI dataset (see Table 2) and on emotional reactivity task (see Table 3). Although the performance was lower on the SRDC task (see Table 4), it can be noted that our model still outperformed all other approaches.

Beyond the clinical dichotomous PTSD diagnosis (yes or no) and general symptom severity (as reflected by CAPS total scores), we assessed different PTSD symptom clusters in order to account for the known heterogeneity in symptom manifestation. This characterization could guide a more personalized clinical approach, for example by targeting the most dominant symptom cluster for each individual. The majority of studies found only qualitatively distinct profiles among diagnosed participants, and were conducted using the DSM-IV-TR symptom criteria or combined PTSD symptoms with additional items (e.g., adverse childhood experiences) (McLafferty et al., 2019). To date, Campbell et al. (2020) were the only study to examine latent classes with the DSM-5 criteria for PTSD, using linear and logistic regressions to identify demographic, trauma-related, and psychiatric characteristics associated with membership in each class.

In the second axis of our work, we filled this gap by training our model to identify distinct, symptom-based clusters of PTSD according to the most updated DSM-5 criteria. We obtained high predictive ability with an average AUC of around 0.8 for the different PTSD symptom clusters (across the three time-points) (see Tables 7 and 8). Furthermore, the classification accuracy and AUC in T1 was lower than in T2 and T3. This might be explained by the fact that 72% of the participants met PTSD diagnosis in T1, which according to DSM-5 means that they met the criteria for all PTSD clusters. However, in T2 and T3 there were more participants with unique symptom-based clusters without full PTSD diagnosis. As demonstrated by the visualization of the learned features before classifying to specific symptom cluster (see Fig. 5), participants in the center of the figure typically met the criteria for several differ-

ent symptom clusters (although we visualized them only in one color representing the most dominant symptom criteria). Interestingly, there was a clear separation between participants with dominant symptoms of negative alterations in mood and cognition (green points, cluster D), and those with dominant avoidance symptoms (red points, cluster C). This suggests that symptom clusters C and D (avoidance, and negative alterations in mood and cognition, respectively) are highly separated, (i.e. both obtain unique features compared to the other clusters). This provides initial evidence, with possibly important clinical insights, regarding the "real-world" validity of DSM-5 symptom clusters. In other words, the results support the transition from three clusters in DSM-IV to four clusters in DSM-5, and specifically the division into two distinct symptom clusters - avoidance, and negative alterations in mood and cognition (Kilpatrick et al., 2013; Pai et al., 2017).

The third axis of our work was the survival analysis. In clinical use, it is beneficial to predict remission or persistence of initial PTSD symptoms. Survival analysis studies the time to a dichotomous event (in this case, persistence of early PTSD symptoms), enabling group comparisons, accounting for censoring and inclusion of time-dependent covariates. Previous survival methods were conducted in PTSD populations mostly using single sampling and only demographic and clinical variables. For example, Müller et al. (2018) implemented a survival analytic approach from a lifetime perspective, to determine whether the expected duration of time until PTSD remission was related to a specific type of trauma, and to identify mediators of PTSD remission or persistence. In this work, we used the survival analysis in a prospective design to predict persistence of PTSD (from one-month to fourteen-months post-trauma), and obtained a high predictive ability for chronicity with an AUC of 0.84 and Accuracy of 81.3% for the general PTSD diagnosis.

Finally, for the fourth axis of our analysis, we applied statistical tests to identify pairwise correlations in the PTSD prediction model. The reported resting state discriminative connectivities are consistent with commonly found abnormal resting state functional connectivity patterns in PTSD involving amygdala functional connections (see Fig. 6) (Sripada et al., 2012a). Interestingly, several common resting-state functional connectivity patterns were found, including the amygdala's connectivity with core regions from the Salience Network and the Default Mode Network, which are known to be involved in emotional reactivity and emotional regulation, respectively. (Etkin et al., 2015). In contrast to rs-fMRI, less discriminative connectivities with the amygdala were found in the emotion reactivity task and SRDC task. This could be explained by the fact that while the rs-fMRI is one long scan, the two fMRI tasks involve dynamically changing conditions over time. In this work, in light of our data-driven approach, we examined the functional connectivity during the tasks without considering the different conditions. Still, task-based fMRI reveals process-related abnormalities such as fusiform connectivity with hippocampus/parahippocampus in the emotional reactivity task (related to emotional face processing) and nucleus accumbens connectivity with insula and parahippocampus in the SRDC task, possibly entails risk and reward processing (related to goal-directed behavior). This type of finding might gain robustness by adding hypothesis-driven analysis to the current data-driven approach. For example, examination of brain regions in specific functional contrasts (e.g., amygdala's activity in the contrast of emotional faces vs. shapes) (Ben-Zion et al., 2020). It is worth noting that our resting state connectivity patterns were less significant at T2. Nevertheless, as was shown in previous work (Ben-Zion et al., 2019a; 2021), it might be explained by the dynamic clinical manifestations during the first critical year after trauma (Hepp et al., 2008), in which an intermediary point of six-months might be too "noisy" to isolate chronic PTSD symptom cluster. This is also supported by similar results in previous work examining structural abnormalities in this cohort (Ben-Zion et al., 2019a), as well as the lowered prediction results at Tables 2–4 for T2.

Although our findings are promising and the cohort is relatively large for an fMRI study of clinical population, this work has several limitations that arise from the challenges of recruiting participants on a larger scale.

**Table 7**  
PTSD cluster predictions. Accuracy of Binary classification and AUC (mean  $\pm$  SD).

		Classification accuracy				AUC			
		class B	class C	class D	class E	class B	class C	class D	class E
Ours	T1	75.5 $\pm$ 7.8	69.5 $\pm$ 11.4	72.7 $\pm$ 11.0	69.8 $\pm$ 6.7	0.79 $\pm$ 0.08	0.84 $\pm$ 0.06	0.79 $\pm$ 0.06	0.82 $\pm$ 0.07
	T2	86.6 $\pm$ 7.3	76.4 $\pm$ 4.2	78.5 $\pm$ 2.1	72.4 $\pm$ 1.4	0.79 $\pm$ 0.00	0.77 $\pm$ 0.05	0.75 $\pm$ 0.02	0.76 $\pm$ 0.01
	T3	81.4 $\pm$ 1.4	73.2 $\pm$ 3.5	76.1 $\pm$ 2.6	70.8 $\pm$ 3.6	0.85 $\pm$ 0.02	0.79 $\pm$ 0.03	0.79 $\pm$ 0.01	0.76 $\pm$ 0.01
Deep fMRI	T1	59.1 $\pm$ 7.1	54.2 $\pm$ 1.3	59.1 $\pm$ 1.6	59.0 $\pm$ 2.5	0.54 $\pm$ 0.03	0.61 $\pm$ 0.06	0.55 $\pm$ 0.09	0.60 $\pm$ 0.06
	T2	71.0 $\pm$ 1.0	68.0 $\pm$ 2.1	66.0 $\pm$ 6.0	60.2 $\pm$ 0.2	0.75 $\pm$ 0.03	0.74 $\pm$ 0.04	0.72 $\pm$ 0.02	0.74 $\pm$ 0.02
	T3	64.4 $\pm$ 8.1	57.8 $\pm$ 8.8	59.0 $\pm$ 11.1	62.8 $\pm$ 6.3	0.67 $\pm$ 0.01	0.75 $\pm$ 0.01	0.68 $\pm$ 0.01	0.73 $\pm$ 0.02

**Table 8**  
Mean AUC  $\pm$  std for each PTSD symptom prediction.

	Symptom	T1	T2	T3
class B	1	0.82 $\pm$ 0.10	0.77 $\pm$ 0.10	0.76 $\pm$ 0.13
	2	0.80 $\pm$ 0.05	0.74 $\pm$ 0.03	0.73 $\pm$ 0.10
	3	0.62 $\pm$ 0.02	0.88 $\pm$ 0.02	0.86 $\pm$ 0.01
	4	0.81 $\pm$ 0.07	0.69 $\pm$ 0.03	0.76 $\pm$ 0.01
class C	5	0.73 $\pm$ 0.15	0.74 $\pm$ 0.05	0.83 $\pm$ 0.06
	6	0.75 $\pm$ 0.07	0.72 $\pm$ 0.06	0.80 $\pm$ 0.09
	7	0.85 $\pm$ 0.07	0.73 $\pm$ 0.08	0.73 $\pm$ 0.05
	8	0.72 $\pm$ 0.10	0.78 $\pm$ 0.03	0.73 $\pm$ 0.03
class D	9	0.67 $\pm$ 0.05	0.80 $\pm$ 0.03	0.80 $\pm$ 0.07
	10	0.71 $\pm$ 0.02	0.86 $\pm$ 0.04	0.84 $\pm$ 0.05
	11	0.74 $\pm$ 0.09	0.73 $\pm$ 0.05	0.81 $\pm$ 0.01
	12	0.72 $\pm$ 0.12	0.70 $\pm$ 0.04	0.72 $\pm$ 0.04
class E	13	0.71 $\pm$ 0.10	0.73 $\pm$ 0.03	0.76 $\pm$ 0.04
	14	0.72 $\pm$ 0.07	0.77 $\pm$ 0.07	0.86 $\pm$ 0.03
	15	0.78 $\pm$ 0.06	0.74 $\pm$ 0.03	0.75 $\pm$ 0.06
	16	0.79 $\pm$ 0.05	0.85 $\pm$ 0.09	
	17	0.78 $\pm$ 0.10	0.77 $\pm$ 0.10	0.69 $\pm$ 0.04
	18	0.78 $\pm$ 0.09	0.73 $\pm$ 0.02	0.75 $\pm$ 0.08
	19	0.73 $\pm$ 0.12	0.75 $\pm$ 0.05	0.67 $\pm$ 0.03
	20	0.66 $\pm$ 0.04	0.65 $\pm$ 0.03	0.70 $\pm$ 0.04
Average:		0.75 $\pm$ 0.06	0.76 $\pm$ 0.05	0.77 $\pm$ 0.05

First, there is class imbalance as 115/160, 39/135, 29/130 participants met PTSD diagnostic criteria at T1, T2 and T3, respectively. That is, most individuals recovered from initial symptoms after six- or fourteen-months from the traumatic incident. To overcome this challenge, we used weight-balancing during the training of the model. For each time-point, the weight was proportional to the ratio between the number of subjects with PTSD and the total number of subjects in this time-point. We also reported balanced error measures, such as the AUC. Second, the number of participants with PTSD decreased between the time-points and at T3 there were only 29 participants with PTSD. Splitting the data to 80% training and 20% test resulted in 23 participants with PTSD in the training set, and 6 in the test set. While this might have reduced the statistical power of the results, we repeated the experiment five times (in a cross validation scheme) in order to somewhat mitigate this issue. Lastly, in this work we used fMRI scans at T1 in order to predict PTSD symptoms at T2 and T3, thus precluding mechanistic insights regarding the clinical outcome. Future work should consider longitudinal neuroimaging and neuropsychological measures from all three time-points to enhance explainability of the clinical prediction. This could possibly guide therapeutic approaches that are more mechanistic in nature.

## 5. Conclusions

Our work demonstrates a computational approach for identifying objective variables linked to clinical PTSD diagnosis across different time-points during the first critical year following trauma exposure, based on single-session examination of functional connectivity shortly after

exposure to trauma. To this end, we applied a novel deep learning model that employs fMRI scans, obtained shortly after trauma, to predict PTSD symptoms at one-, six- and fourteen-months following trauma. Our method demonstrates a significant improvement in performance in comparison to other analytical techniques reported in fMRI literature. The prediction is highly accurate compared to the existing methods for all three time-points and benefits from learning these all at once, using a single model. We further show a high predictive ability for predicting PTSD symptom clusters and persistence of the disorder. If validated, the objective features derived from our computational model may further guide mechanism-driven interventions for PTSD (e.g., neuromodulation techniques). Our code will be promptly shared as an open-source code.

## Ethical approval

The research study meets all ethical regulations as required by NYU Langone Health Institutional Review Board (IRB) and ethics committee in Tel-Aviv Sourasky Medical Center (Reference number 0207/14). All subjects gave written informed consent in accordance with the Declaration of Helsinki. This study is registered at ClinicalTrials.gov (registration number: NCT03756545)

## Credit authorship contribution statement

**Shelly Sheynin:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Lior Wolf:** Conceptualization, Methodology, Writing - original

draft, Writing - review & editing, Supervision. **Ziv Ben-Zion:** Methodology, Resources, Data curation, Writing - original draft, Writing - review & editing. **Jony Sheynin:** Methodology, Resources, Data curation, Writing - review & editing. **Shira Reznik:** Resources, Software. **Jackob Nimrod Keynan:** Resources, Data curation. **Roei Admon:** Resources. **Arieh Shalev:** Conceptualization, Methodology, Funding acquisition, Writing - original draft, Writing - review & editing. **Talma Hendlér:** Conceptualization, Methodology, Resources, Funding acquisition, Writing - original draft, Writing - review & editing. **Israel Liberzon:** Conceptualization, Methodology, Resources, Funding acquisition, Writing - original draft, Writing - review & editing.

## Acknowledgments

The work was supported by award number **R01-MH-103287** from the **National Institute of Mental Health** (NIMH) given to A. Y.S. (PI), I.L. and T.H. (co-Investigators, subcontractors), and had undergone critical review by the NIMH Adult Psychopathology and Disorders of Aging study section. Additional funding was provided by the ISF Israel Precision Medicine Partnership (IPMP) Program (grant 2923/20). This project has also received funding from the European Research Council (ERC) under the European Unions **Horizon 2020** research and innovation programme (grant ERC **CoG 725974**). Lastly, this work was additionally supported by a grant from the U.S. Department of Defense, award number W81XWH-16-C-0198 and also by the ISRAEL SCIENCE FOUNDATION (grant No. 2923/20), within the Israel Precision Medicine Partnership program. We would like to thank the research team at Tel-Aviv Sourasky Medical Center - including Nili Green, Mor Halevi, Sheli Luvton, Yael Shavit, Olga Nevenchannaya, Iris Rashap, Efrat Routledge, and Ophir Leshets for their major contribution in carrying out this research, including subjects recruitment and screening, and performing clinical, cognitive, and neural assessments. We also want to thank Naomi Fine and Michal Achituv for setting up the clinical aspect of the research. Last but not least, we extend our gratitude to all the participants of this study, who completed all the assessments at three different time-points after experiencing a traumatic event.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.neuroimage.2021.118242](https://doi.org/10.1016/j.neuroimage.2021.118242)

## References

- Admon, R., Lubin, G., Rosenblatt, J.D., Stern, O., Kahn, I., Assaf, M., Hendlér, T., 2013. Imbalanced neural responsivity to risk and reward indicates stress vulnerability in humans. *Cereb. Cortex* 23 (1), 28–35.
- Assaf, M., Kahn, I., Pearson, G.D., Johnson, M.R., Yeshurun, Y., Calhoun, V.D., Hendlér, T., 2009. Brain activity dissociates mentalization from motivation during an interpersonal competitive game. *Brain Imaging Behav.* 3 (1), 24.
- Ben-Zion, Z., Artzi, M., Niruy, D., Keynan, N.J., Zeevi, Y., et al., 2019. Neuroanatomical risk factors for post-traumatic stress disorder (PTSD) in recent trauma survivors. *Biol. Psychiatry*.
- Ben-Zion, Z., Fine, N.B., Keynan, N.J., Admon, R., Green, N., Halevi, M., Fonzo, G.A., Achituv, M., et al., 2018. Cognitive flexibility predicts PTSD symptoms: observational and interventional studies. *Front. Psychiatry* 9, 477.
- Ben-Zion, Z., Fine, N.B., Keynan, N.J., Admon, R., Halpern, P., Liberzon, I., Hendlér, T., Shalev, A.Y., 2019. Neurobehavioral moderators of post-traumatic stress disorder (PTSD) trajectories: study protocol of a prospective MRI study of recent trauma survivors. *Eur. J. Psychotraumatol.* 10 (1), 1683941.
- Ben-Zion, Z., Shany, O., Admon, R., Keynan, N.J., Avisdris, N., Balter, S. R., Shalev, A. Y., Liberzon, I., Hendlér, T., 2021. Differential roles of positive and negative valence systems in the development of post-traumatic stress psychopathology. *bioRxiv*.
- Ben-Zion, Z., Zeevi, Y., Keynan, N.J., Admon, R., Kozlovski, T., Sharon, H., Halpern, P., Liberzon, I., Shalev, A.Y., Benjamini, Y., et al., 2020. Multi-domain potential biomarkers for post-traumatic stress disorder (PTSD) severity in recent trauma survivors. *Transl. Psychiatry* 10 (1), 1–11.
- Bengs, M., Gessert, N., Schlaefke, A., 2019. 4D spatio-temporal deep learning with 4D fMRI data for autism spectrum disorder classification. In: International Conference on Medical Imaging with Deep Learning. London, United Kingdom.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*
- Blake, D.D., Weathers, F.W., et al., 1995. The development of a clinician-administered PTSD scale. *J. Trauma Stress* 8 (1), 75–90.
- Bleich-Cohen, M., Jamshy, S., Sharon, H., Weizman, R., Intrator, N., Poyurovsky, M., Hendlér, T., 2014. Machine learning fmri classifier delineates subgroups of schizophrenia patients. *Schizophr. Res.* 160 (1–3), 196–200.
- Bonnelle, V., Ham, T.E., Leech, R., Kinnunen, K.M., Mehta, M.A., Greenwood, R.J., Sharp, D.J., 2012. Salience network integrity predicts default mode network function after traumatic brain injury. *Proc. Natl. Acad. Sci.* 109 (12), 4690–4695.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32.
- Brewin, C.R., 2005. Risk factor effect sizes in PTSD: what this means for intervention. *J. Trauma Dissociation* 6 (2), 123–130.
- Brewin, C.R., Andrews, B., Valentine, J.D., 2000. Meta-analysis of risk factors for post-traumatic stress disorder in trauma-exposed adults.. *J. Consult. Clin. Psychol.* 68 (5), 748.
- Brown, V.M., LaBar, K.S., Haswell, C.C., Gold, A.L., McCarthy, G., Morey, R.A., 2014. Altered resting-state functional connectivity of basolateral and centromedial amygdala complexes in posttraumatic stress disorder. *Neuropsychopharmacology* 39 (2), 351–359.
- Bryant, R.A., Nickerson, A., Creamer, M., O'Donnell, M., Forbes, D., Galatzer-Levy, I., McFarlane, A.C., Silove, D., 2015. Trajectory of post-traumatic stress following traumatic injury: 6-year follow-up. *Br. J. Psychiatry* 206 (5), 417–423.
- Campbell, S.B., Trachik, B., Goldberg, S., Simpson, T.L., 2020. Identifying PTSD symptom typologies: a latent class analysis. *Psychiatry Res.* 112779.
- Dakka, J., Bashivan, P., Gheiratmand, M., Rish, I., Jha, S., Greiner, R., 2017. Learning neural markers of schizophrenia disorder using recurrent neural networks. *arXiv:1712.00512*.
- Davis, M., Whalen, P.J., 2001. The amygdala: vigilance and emotion. *Mol. Psychiatry* 6 (1), 13–34.
- Dierks, T., Linden, D.E., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., Singer, W., 1999. Activation of Heschls gyrus during auditory hallucinations. *Neuron* 22 (3), 615–621.
- DiGangi, J.A., Tadayyon, A., Fitzgerald, D.A., et al., 2016. Reduced default mode network connectivity following combat trauma. *Neurosci. Lett.* 615, 37–43.
- Esteban, O., et al., 2018. fMRIprep: a robust preprocessing pipeline for functional MRI. *Nat. Methods*.
- Etkin, A., Büchel, C., Gross, J.J., 2015. The neural bases of emotion regulation. *Nat. Rev. Neurosci.* 16 (11), 693–700.
- Etkin, A., Wager, T.D., 2007. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* 164 (10), 1476–1488.
- Friston, K.J., 2003. Statistical Parametric Mapping. Neuroscience databases.
- Galatzer-Levy, I.R., Ankri, Y., Freedman, S., Israeli-Shalev, Y., Roitman, P., Gilad, M., Shalev, A.Y., 2013. Early ptsd symptom trajectories: persistence, recovery, and response to treatment: results from the jerusalem trauma outreach and prevention study (j-tops). *PLoS One* 8 (8), e70084.
- Galatzer-Levy, I.R., Ma, S., Stattnikov, A., et al., 2017. Utilization of machine learning for prediction of post-traumatic stress: a re-examination of cortisol in the prediction and pathways to non-remitting PTSD. *Transl. Psychiatry*.
- Gonen, T., Admon, R., Podlipsky, I., Hendlér, T., 2012. From animal model to human brain networking: dynamic causal modeling of motivational systems. *J. Neurosci.* 32 (21), 7218–7224.
- Gong, Q., Li, L., Du, M., Pettersson-Yeo, W., Crossley, N., Yang, X., Li, J., Huang, X., Mechelli, A., 2014. Quantitative prediction of individual psychopathology in trauma survivors using resting-state fMRI. *Neuropsychopharmacology* 39 (3), 681–687.
- Gradus, J.L., King, M.W., Galatzer-Levy, I., Street, A.E., 2017. Gender differences in machine learning models of trauma and suicidal ideation in veterans of the iraq and afghanistan wars. *J. Trauma Stress* 30 (4), 362–371.
- Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7 (5), 555–562.
- Haller, S., Lovblad, K.-O., Giannakopoulos, P., Van De Ville, D., 2014. Multivariate pattern recognition for diagnosis and prognosis in clinical neuroimaging: state of the art, current challenges and future trends. *Brain Topogr.* 27 (3), 329–337.
- Hariri, A.R., Bookheimer, S.Y., Mazziotta, J.C., 2000. Modulating emotional responses: effects of a neocortical network on the limbic system. *NeuroReport* 11 (1), 43–48.
- Hendlér, T., Rotshtein, P., Yeshurun, Y., Weizmann, T., Kahn, I., Ben-Bashat, D., Malach, R., Bleich, A., 2003. Sensing the invisible: differential sensitivity of visual cortex and amygdala to traumatic context. *NeuroImage* 19 (3), 587–600.
- Hepp, U., Moergeli, H., Buchi, S., Bruchhaus-Steinert, H., Kraemer, B., Sensky, T., Schneider, U., 2008. Post-traumatic stress disorder in serious accidental injury: 3-year follow-up study. *Br. J. Psychiatry* 192 (5), 376–383.
- Heron-Delaney, M., Kenardy, J., Charlton, E., Matsuoka, Y., 2013. A systematic review of predictors of posttraumatic stress disorder (PTSD) for adult road traffic crash survivors. *Injury* 44 (11), 1413–1422.
- Hyatt, C.J., Assaf, M., Muska, C.E., Rosen, R.I., Thomas, A.D., Johnson, M.R., Hylton, J.L., Andrews, M.M., Reynolds, B.A., Krystal, J.H., et al., 2012. Reward-related dorsal striatal activity differences between former and current cocaine dependent individuals during an interactive competitive game. *PLoS One* 7 (5).
- Jin, C., Jia, H., Lanka, P., Rangaprakash, D., Li, L., Liu, T., Hu, X., Deshpande, G., 2017. Dynamic brain connectivity is a better predictor of PTSD than static connectivity. *Hum. Brain Mapp.* 38 (9), 4479–4496.
- Kahn, I., Yeshurun, Y., Rotshtein, P., Fried, I., Ben-Bashat, D., Hendlér, T., 2002. The role of the amygdala in signaling prospective outcome of choice. *Neuron* 33 (6), 983–994.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17 (11), 4302–4311.

- Karstoft, K.-I., Galatzer-Levy, I.R., Statnikov, A., Li, Z., Shalev, A.Y., et al., 2015. Bridging a translational gap: using machine learning to improve the prediction of ptsd. *BMC Psychiatry* 15 (1), 30.
- Kendrick, D., Baker, R., Hill, T., Beckett, K., Coupland, C., Kellezi, B., Joseph, S., Barnes, J., Sleney, J., Christie, N., et al., 2018. Early risk factors for depression, anxiety and post-traumatic distress after hospital admission for unintentional injury: multicentre cohort study. *J. Psychosom. Res.* 112, 15–24.
- Kessler, R.C., 2000. Posttraumatic stress disorder: the burden to the individual and to society.. *J. Clin. Psychiatry*.
- Khosla, M., Jamison, K., Kuceyeski, A., Sabuncu, M.R., 2018. 3D convolutional neural networks for classification of functional connectomes. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*.
- Kilpatrick, D.G., Resnick, H.S., Milanak, M.E., Miller, M.W., Keyes, K.M., Friedman, M.J., 2013. National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. *J. Trauma Stress* 26 (5), 537–547.
- King, A.P., Block, S.R., Sripada, R.K., Rauch, S., Giardino, N., Favorite, T., Angstadt, M., Kessler, D., Welsh, R., Liberzon, I., 2016. Altered default mode network (DMN) resting-state functional connectivity following a mindfulness-based exposure therapy for posttraumatic stress disorder (PTSD) in combat veterans of Afghanistan and Iraq. *Depress. Anxiety* 33 (4), 289–299.
- Koutsouleris, N., Kambeitz, J., 2016. Pattern recognition methods in the prediction of psychosis. In: *Early Detection and Intervention in Psychosis*, 181. Karger Publishers, pp. 95–102.
- Li, X., Dvornek, N.C., Zhou, Y., Zhuang, J., Ventola, P., Duncan, J.S., 2019. Graph neural network for interpreting task-fMRI biomarkers. *MICCAI*.
- Liu, F., Guo, W., Yu, D., Gao, Q., Gao, K., Xue, Z., Du, H., Zhang, J., Tan, C., Liu, Z., et al., 2012. Classification of different therapeutic responses of major depressive disorder with multivariate pattern analysis method based on structural mr scans. *PLoS One* 7 (7), e40968.
- Liu, F., Xie, B., Wang, Y., Guo, W., Fouche, J.-P., Long, Z., Wang, W., Chen, H., Li, M., Duan, X., et al., 2015. Characterization of post-traumatic stress disorder using resting-state fMRI with a multi-level parametric classification approach. *Brain Topogr.* 28 (2), 221–237.
- Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-sne. *J. Mach. Learn. Res.* 9 (Nov), 2579–2605.
- Mao, Z., Su, Y., Xu, G., Wang, X., Huang, Y., Yue, W., et al., 2019. Spatio-temporal deep learning method for ADHD fMRI classification. *Inf. Sci.*
- Mason, S., Wardrobe, J., Turpin, G., Rowlands, A., 2002. The psychological burden of injury: an 18 month prospective cohort study. *Emerg. Med. J.* 19 (5), 400–404.
- McLafferty, M., Ross, J., Waterhouse-Bradley, B., Armour, C., 2019. Childhood adversities and psychopathology among military veterans in the us: the mediating role of social networks. *J. Anxiety Disord.* 65, 47–55.
- Mikolas, P., Melicher, T., Skoch, A., Matejka, M., Slováková, A., Bakstein, E., Hajek, T., Spaniel, F., 2016. Connectivity of the anterior insula differentiates participants with first-episode schizophrenia spectrum disorders from controls: a machine-learning study. *Psychol. Med.* 46 (13), 2695–2704.
- Miyashita, Y., 1993. Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* 16 (1), 245–263.
- Mourao-Miranda, J., Reinders, A., Rocha-Rego, V., Lappin, J., Rondina, J., Morgan, C., Morgan, K.D., Fearon, P., Jones, P.B., Doody, G.A., et al., 2012. Individualized prediction of illness course at the first psychotic episode: a support vector machine MRI study. *Psychol. Med.* 42 (5), 1037–1047.
- Müller, M., Ajdacic-Gross, V., Rodgers, S., Kleim, B., Seifritz, E., Vetter, S., Egger, S.T., Rössler, W., Castelao, E., Preisig, M., et al., 2018. Predictors of remission from ptsd symptoms after sexual and non-sexual trauma in the community: a mediated survival-analytic approach. *Psychiatry Res.* 260, 262–271.
- Niepert, M., Ahmed, M., Kutzkov, K., 2016. Learning convolutional neural networks for graphs. In: *International Conference on Machine Learning*, pp. 2014–2023.
- Omurca, S.I., Ekinici, E., 2015. An alternative evaluation of post traumatic stress disorder with machine learning methods. In: *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*. IEEE, pp. 1–7.
- Ozer, E.J., Best, S.R., Lipsey, T.L., Weiss, D.S., 2003. Predictors of posttraumatic stress disorder and symptoms in adults: a meta-analysis.. *Psychol. Bull.* 129 (1), 52.
- Pai, A., Suris, A.M., North, C.S., 2017. Posttraumatic stress disorder in the dsm-5: controversy, change, and conceptual considerations. *Behav. Sci.* 7 (1), 7.
- Patel, P., Aggarwal, P., Gupta, A., 2016. Classification of schizophrenia versus normal subjects using deep learning. In: *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*. Association for Computing Machinery, New York, NY, USA doi:10.1145/3009977.3010050.
- Pitman, R.K., Rasmussen, A.M., Koenen, K.C., Shin, L.M., Orr, S.P., Gilbertson, M.W., Milad, M.R., Liberzon, I., 2012. Biological studies of post-traumatic stress disorder. *Nat. Rev. Neurosci.* 13 (11), 769–787.
- van der Ploeg, T., Nieboer, D., Steyerberg, E.W., 2016. Modern modeling techniques had limited external validity in predicting mortality from traumatic brain injury. *J. Clin. Epidemiol.* 78, 83–89.
- Qi, W., Ratanathathorn, A., Gevonden, M., Bryant, R., Delahanty, D., Matsuoka, Y., Olff, M., deRoon Cassini, T., Schnyder, U., Seedat, S., et al., 2018. Application of data pooling to longitudinal studies of early post-traumatic stress disorder (PTSD): the international consortium to predict PTSD (ICPP) project. *Eur. J. Psychotraumatol.* 9 (1), 1476442.
- Rabinak, C.A., Angstadt, M., Welsh, R.C., Kennedy, A., Lyubkin, M., Martis, B., Phan, K.L., 2011. Altered amygdala resting-state functional connectivity in post-traumatic stress disorder. *Front. Psychiatry* 2, 62.
- Raichle, M.E., 2015. The brain's default mode network. *Annu. Rev. Neurosci.* 38, 433–447.
- Riaz, A., Asad, M., Arif, S.M.M.R.A., Alonso, E., Dima, D., Corr, P., Slabaugh, G., 2018. Deep fMRI: an end-to-end deep network for classification of fMRI data. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*.
- Richmond, T.S., Ruzeck, J., Ackerson, T., Wiebe, D.J., Winston, F., Kassam-Adams, N., 2011. Predicting the future development of depression or PTSD after injury. *Gen. Hosp. Psychiatry* 33 (4), 327–335.
- Rive, M.M., Redlich, R., Schmaal, L., Marquand, A.F., Dannlowski, U., Grotegerd, D., Veltman, D.J., Schene, A.H., Ruhé, H.G., 2016. Distinguishing medication-free subjects with unipolar disorder from subjects with bipolar disorder: state matters. *Bipolar Disord.* 18 (7), 612–623.
- Rosellini, A.J., Dussaillant, F., Zubizarreta, J.R., et al., 2018. Predicting posttraumatic stress disorder following a natural disaster. *J. Psychiatr. Res.*
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron* 70 (6), 1054–1069.
- Saxe, G.N., Ma, S., Ren, J., Aliferis, C., 2017. Machine learning methods to predict child posttraumatic stress: a proof of concept study. *BMC Psychiatry* 17 (1), 223.
- Scher, C.D., Suvak, M.K., Resick, P.A., 2017. Trauma cognitions are related to symptoms up to 10 years after cognitive behavioral treatment for posttraumatic stress disorder.. *Psychol. Trauma* 9 (6), 750.
- Schultebrucks, K., Galatzer-Levy, I.R., 2019. Machine learning for prediction of posttraumatic stress and resilience following trauma: an overview of basic concepts and recent advances. *J. Trauma Stress* 32 (2), 215–225.
- Schwartz, I., Yu, S., Hazan, T., Schwing, A.G., 2019. Factor graph attention. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2039–2048.
- Shalev, A., Liberzon, I., Marmar, C., 2017. Post-traumatic stress disorder. *N. Engl. J. Med.* 376 (25), 2459–2469.
- Shalev, A.Y., Gevonden, M., Ratanathathorn, A., Laska, E., Van Der Mei, W.F., Qi, W., Lowe, S., Lai, B.S., Bryant, R.A., Delahanty, D., et al., 2019. Estimating the risk of PTSD in recent trauma survivors: results of the international consortium to predict PTSD (ICPP). *World Psychiatry* 18 (1), 77–87.
- Shin, L.M., Liberzon, I., 2010. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology* 35 (1), 169–191.
- Sripada, R.K., King, A.P., Garfinkel, S.N., Wang, X., Sripada, C.S., Welsh, R.C., Liberzon, I., 2012. Altered resting-state amygdala functional connectivity in men with posttraumatic stress disorder. *J. Psychiatry Neurosci.*
- Sripada, R.K., King, A.P., Welsh, R.C., Garfinkel, S.N., Wang, X., Sripada, C.S., Liberzon, I., 2012. Neural dysregulation in posttraumatic stress disorder: evidence for disrupted equilibrium between salience and default mode brain networks. *Psychosom. Med.* 74 (9), 904.
- Stein, D.J., McLaughlin, K.A., Koenen, K.C., Atwoli, L., Friedman, M.J., Hill, E.D., Maercker, A., Petukhova, M., Shahly, V., Van Ommeren, M., et al., 2014. DSM-5 and ICD-11 definitions of posttraumatic stress disorder: investigating narrow and broad approaches. *Depress. Anxiety* 31 (6), 494–505.
- Thaler, A., Gonen, T., Mirelman, A., Helmich, R.C., Gurevich, T., Orr-Urtreger, A., Bloem, B.R., Giladi, N., Hendler, T., et al., 2019. Altered reward-related neural responses in non-manifesting carriers of the parkinson disease related lrrk2 mutation. *Brain Imaging Behav.* 13 (4), 1009–1020.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention is all you need. In: *Advances in Neural Information Processing Systems*, pp. 5998–6008.
- Wang, Z., Zhu, H., et al., 2020. The resting-state functional connectivity of amygdala subregions associated with post-traumatic stress symptom and sleep quality in trauma survivors. *Eur. Arch. Psychiatry Clin. Neurosci.*
- Weathers, F.W., Bovin, M.J., Lee, D.J., Sloan, D.M., Schnurr, P.P., Kaloupek, D.G., Keane, T.M., Marx, B.P., 2018. The clinician-administered PTSD scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans.. *Psychol. Assess.* 30 (3), 383.
- Yuan, H., Phillips, R., Wong, C.K., Zotev, V., Misaki, M., Wurfel, B., Krueger, F., Feldner, M., Bodurka, J., 2018. Tracking resting state connectivity dynamics in veterans with PTSD. *NeuroImage* 19, 260–270.
- Zeng, L.-L., Shen, H., Liu, L., Wang, L., Li, B., Fang, P., Zhou, Z., Li, Y., Hu, D., 2012. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain* 135 (5), 1498–1507.
- Zou, L., Zheng, J., et al., 2017. 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access*.