Application of Drug Efficiency Index Metric for Analysis of Post-Traumatic Stress Disorder and Treatment Resistant Depression Gene Expression Profiles

Nicolas Borisov 1, Yaroslav Ilnytskyy 2,3, Boseon Byeon 2,3,4, Olga Kovalchuk 2,3 and Igor Kovalchuk 2,3,*

1 Moscow Institute of Physics and Technology, 9 Institutsky Lane, Dolgoprudny, Moscow 141701, Russia
2 Department of Biological Sciences, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada
3 Pathway Rx., 2 Fortress Rise SW, Calgary, AB T3H 4Z2, Canada
4 Biomedical and Health Informatics, Computer Science Department, State University of New York, 2 S Clinton St, Syracuse, NY 13202, USA
* Correspondence: igor.kovalchuk@uleth.ca

Abstract: Post-traumatic stress disorder (PTSD) is a severe mental illness with grave social, political, economic, and humanitarian implications. To apply the principles of personalized omics-based medicine to this psychiatric problem, we implemented our previously introduced drug efficiency index (DEI) to the PTSD gene expression datasets. Generally, omics-based personalized medicine evaluates individual drug action using two classes of data: (1) gene expression, mutation, and Big Data profiles, and (2) molecular pathway graphs that reflect the protein–protein interaction. In the particular case of the DEI metric, we evaluate the drug action according to the drug’s ability to restore healthy (control) activation levels of molecular pathways. We have curated five PTSD and one TRD (treatment-resistant depression) cohorts of next-generation sequencing (NGS) and microarray hybridization (MH) gene expression profiles, which, in total, comprise 791 samples, including 379 cases and 413 controls. To check the applicability of our DEI metrics, we have performed three differential studies with gene expression and pathway activation data: (1) case samples vs. control samples, (2) case samples after treatment or/and observation vs. before treatment, and (3) samples from patients positively responding to the treatment vs. those responding negatively or non-responding patients. We found that the DEI values that use the signaling pathway impact activation (SPIA) metric were better than those that used the Oncobox pathway activation level (Oncobox PAL) approach. However, SPIA, Oncobox PAL, and DEI evaluations were reliable only if there were differential genes between case and control, or treated and untreated, samples.

Keywords: post-traumatic stress disorder; treatment-resistant depression; next-generation sequencing data; gene expression; drug efficiency index; differentially expressed genes

1. Introduction

Post-traumatic stress disorder (PTSD) is a pathological mental health condition triggered by a terrifying event—either experiencing it or witnessing it. Symptoms may include flashbacks, nightmares, severe anxiety, and uncontrollable thoughts about the event [1,2]. Given the drastic, humiliating, and oppressing conditions that provoke PTSD, it poses a grave and ever-actual social, political, and humanitarian threat worldwide [3,4]. Most people who go through traumatic events may have temporary difficulty adjusting and coping, but with time and good self-care, they usually get better. If the symptoms worsen, last for more than a month or longer, and interfere with day-to-day functioning, the patient may have PTSD [5,6].

The epidemiology of PTSD seems controversial around the world; this is likely due to cultural differences and non-identical healthcare standards in different nations. For example, about 3.5% of adults in the US have PTSD in a given year. Additionally, 9% of US
people develop it at some point in their life; however, in much of the rest of the world, rates during a given year are between 0.5% and 1% [3]. Higher rates may occur in regions of armed conflict. Consequently, there is debate over the rates of PTSD found in populations. However, despite changes in diagnosis and the criteria used to define PTSD between 1997 and 2013, the epidemiological rates have not changed significantly [7,8]. Women are more vulnerable to PTSD than men [9,10].

Obtaining effective treatment after PTSD symptoms develop can be critical to improving the functional state of a patient. There are different types of treatments available to PTSD patients. The approaches that mitigate PTSD include behavioral and cognitive-behavioral therapies, such as prolonged exposure therapy [11], cognitive processing therapy, eye movement desensitization and reprocessing (EMDR) [12,13], brief eclectic psychotherapy (BEP), narrative exposure therapy (NET), written exposure therapy [14], social supports [15,16], interpersonal psychotherapy [17], prolonged exposure therapy, and relaxation therapy [18,19]. Additionally, doctors recommend moderate exercise to distract from disturbing emotions, build self-esteem, and increase feelings of being in control again, as well as to improve physical health [20].

Many of the approaches, unfortunately, are not effective. A 2019 Cochrane review, for example, did not find any evidence to support the use of a standard intervention offered to everyone, and multiple-session interventions may result in worse outcomes than no intervention for some individuals [21]. Additionally, the World Health Organization disapproves of the application of benzodiazepines and antidepressants during acute stress (with symptoms lasting shorter than one month) [22]. Some evidence supports the use of hydrocortisone for the reduction of PTSD symptoms in adults, while there is limited or no evidence supporting the use of propranolol, escitalopram, temazepam, or gabapentin [15,23,24].

Considering pharmacotherapy, four CYP2D6 inhibitors (sertraline, fluoxetine, paroxetine, and venlafaxine) have a small to modest benefit over placebo [4]. Selective serotonin reuptake inhibitors (SSRIs) and serotonin–norepinephrine reuptake inhibitors (SNRIs) may have some benefits for PTSD symptoms [4,25,26]. Tricyclic antidepressants are equally effective but are less well tolerated [27]. Benzodiazepines are not the best solution for PTSD due to a lack of evidence of benefits and the risk of worsening PTSD symptoms [28]. Prazosin, an alpha-1 adrenergic antagonist, was used by war veterans to reduce nightmares, although there was some variability in symptom reduction and dosage used in this population [29–31]. Glucocorticoids are beneficial in the short term to protect against neurodegeneration triggered by PTSD. However, their further application may cause neurodegeneration [31].

Cannabis is another possible therapy option for PTSD patients. Some authors, however, do not recommend cannabis for the treatment of PTSD. They argue there is no sufficient evidence for the efficacy of cannabinoids [32,33] or smoked cannabis [34]. The cannabinoids, however, are sometimes given to PTSD patients [32,35–43]. Additionally, the application of cannabis or derived products is widespread among US veterans with PTSD [44]. An increasing number of states permit and have legalized the use of medical cannabis for the treatment of PTSD [34,45–50]. The low efficacy of cannabis and cannabinoids for PTSD or other conditions could be because cannabis extracts are given randomly, without knowing beforehand whether the combination of active ingredients would be effective.

Recently, we have proposed a novel approach to the assessment of the cannabis drug efficiency index (CDEI) for individual patients. We calculated the CDEI using expression-based signaling pathway activation levels [51]. The corresponding software drug efficiency index (DEI) metric was presented and validated on the prediction of cannabis drug efficiency for the treatment of inflammation in 3D epithelial tissues, including intestine, oral, and skin tissues [51]. The changes in the expression of genes involved in various signaling/metabolic/cytoskeleton pathways often help in the evaluation of personalized drug action on individual patients, particularly for the treatment of cancer [52–56]; dia-
betes [57,58]; inflammation [59,60], including COVID-19 cases [61,62]; hemostasis disorders [63]; rheumatoid arthritis [64]; and for delaying aging [65,66].

We can evaluate pathway activation levels using various mathematical algorithms, such as the Oncobox pathway activation level (PAL) [67], topology analysis of pathway phenotype association (TAPPA) [68], topology-based score (TBScore) [69], pathway-express (PE) [70], signaling pathway impact analysis (SPIA) [71], in silico pathway activation network decomposition analysis (iPANDA) [65], etc. Such a pathway activation level calculation utilizes expression/mutation Big Data profiles. The protein–protein interaction graphs are used from several databases such as Biocarta [72], KEGG [73], HumanCyc [74], Qiagen (https://www.qiagen.com/us/shop/genes-and-pathways/pathway-central; accessed on 15 December 2022), NCI [75], Reactome [76], PathBank [77], etc.

In the current work, we have attempted to use and modify our gene expression analysis software for the DEI metric [51] to analyze the gene expression data in PTSD patients. We found that the DEI metric is meaningful if there are enough differentially expressed genes between the case and control samples and the pre-treatment and post-treatment samples. Otherwise, the DEI metric is not reliable, which correlates with high \( p \)-values for the differential Student’s \( t \)-test.

2. Materials and Methods

2.1. PTSD Gene Expression Datasets

We processed six gene expression datasets for PTSD and treatment-resistant depression (TRD) patients before and after treatment or observation, as well as corresponding gene expression profiles for the control/healthy samples (See Table 1). For each dataset, except dataset GSE185855, there was no specific treatment for PTSD given. In other words, the main expectation was that time cures stress. The untreated cases were cases before the observation period, whereas the treated cases were those studied after treatment.

The microarray gene expression data were normalized using the quantile normalization (QN) method [78], and the next-generation sequencing (NGS) profiles were normalized using the Differentially Expression in Sequencing 2 (DESeq2) approach [79].

<table>
<thead>
<tr>
<th>Reference</th>
<th>GSE ID; clinical description</th>
<th>Profiling platform</th>
<th>Treatment</th>
<th>Untreated cases</th>
<th>Untreated controls</th>
<th>Treated cases</th>
<th>Treated controls</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[80]</td>
<td>GSE860</td>
<td>Affymetrix U95A Array</td>
<td>Not specified</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9068</td>
</tr>
<tr>
<td>[81]</td>
<td>GSE64814-GPL6244</td>
<td>Affymetrix Gene 1.0 ST Array</td>
<td>Not specified</td>
<td>24</td>
<td>48</td>
<td>24</td>
<td>48</td>
<td>12,631</td>
</tr>
<tr>
<td>[81]</td>
<td>GSE64814-GPL11154</td>
<td>Illumina HiSeq 2000</td>
<td>Not specified</td>
<td>47</td>
<td>94</td>
<td>47</td>
<td>94</td>
<td>10,184</td>
</tr>
<tr>
<td>[82]</td>
<td>GSE81761 no improvement; non-responders</td>
<td>Affymetrix U133 Plus 2.0 Array</td>
<td>*</td>
<td>39</td>
<td>27</td>
<td>5</td>
<td>6</td>
<td>22,878</td>
</tr>
<tr>
<td>[82]</td>
<td>GSE81761 improvement; responders</td>
<td>Affymetrix U133 Plus 2.0 Array</td>
<td>*</td>
<td>39</td>
<td>27</td>
<td>6</td>
<td>6</td>
<td>22,878</td>
</tr>
<tr>
<td>[30,83]</td>
<td>GSE97356</td>
<td>Illumina HiSeq 2000</td>
<td>Not specified</td>
<td>81</td>
<td>201</td>
<td>8</td>
<td>21</td>
<td>25,830</td>
</tr>
<tr>
<td>[84]</td>
<td>GSE185855 TRD: non-responders</td>
<td>Illumina HiSeq 2000</td>
<td>Ketamine</td>
<td>8</td>
<td>21</td>
<td>8</td>
<td>13</td>
<td>55,178</td>
</tr>
<tr>
<td>[84]</td>
<td>GSE185855 TRD: responders</td>
<td>Illumina HiSeq 2000</td>
<td>Ketamine</td>
<td>17</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>55,178</td>
</tr>
</tbody>
</table>

* Treatment: Evidence-based sleep treatment; cognitive behavioral therapy for insomnia; automatic positive airway pressure therapy.

2.2. Pathway Activation Level Assessment

Previously, we introduced two computational methods for the activation levels of cell signaling pathways. One approach, called Oncobox PAL [67], calculates for each pathological case sample the following quantities:
1. Log-fold-change, \(^{LFC}_n\), i.e., \(\log_2\) ratio of gene \(n\) mRNA concentrations in the test sample and in the control pool (median value in the control group)

2. Pathway \(p\) activation level, \(^{PAL}_p = \sum_n \text{NII}_{np} \cdot \text{ARR}_{np} \cdot \frac{^{LFC}_n}{\sum_n |\text{ARR}_n|}\). Here, the node involvement index, \(\text{NII}_{np}\), is the Boolean flag for gene product \(n\) concerning the pathway \(p\). The discrete value \(\text{ARR}_{np}\) (activator/repressor role) is determined for a gene \(n\) in the pathway \(p\) as follows:

\(\text{ARR}_{np} = -1;\) gene product \(n\) is a repressor of pathway \(p\)

\(-0.5;\) gene product \(n\) is more a repressor than an activator of pathway \(p\)

\(0;\) unclear repressor or activator role in pathway \(p\)

\(0.5;\) gene product \(n\) is more an activator than a repressor of pathway \(p\)

\(1;\) gene product \(n\) is an activator of pathway \(p\).

Note that we calculate LFC and PAL values using the available corresponding control samples. The algorithm for ARR calculations uses pathway topology graphs [85,86]. We checked the adequacy of Oncobox PAL, a method for pathway activation assessment, using the low-level realistic kinetic models for signal propagation along the pathways [86,87]. The Oncobox database for signaling/metabolic/cytoskeleton pathways encounters 51,672 uniformly processed graphs. The full interactome graph contains 361,654 reactions and 64,095 molecular nodes [88] borrowed from the following storehouses: Biocarta [72], KEGG [73], HumanCyc [74], Qiagen (https://www.qiagen.com/us/shop/genes-and-pathways/pathway-central; accessed on 15 Decembe 2022), NCI [75], Reactome [76], and PathBank [77].

Another method we have proposed for pathway-based assessment of individual drug reactions is the DEI metric, as we have applied it to cannabis-based drugs. It uses the SPIA method [51,71] for pathway activation levels [51,71]. Among the other methods for gene expression-based assessment of signaling pathway activation, including TAPPA [68], SPIA method [51,71] for pathway activation levels [51,71]. Among the other methods for

We calculate the SPIA values [51,71] for a pathway graph, \(G(V,E)\), where \(V = \{g_1, g_2, \ldots , g_u\}\) is the set of graph nodes (vertices), and \(E = \{(g_i, g_j)\mid g_i \text{ and } g_j \text{ interact}\}\) is the set of graph edges. We define the adjacency matrix \(a\) as follows: \(a_{ij} = 1\), if \(i \neq j\) or \((g_i, g_j) \in E\) and \(a_{ij} = 0\), if \((g_i, g_j) \notin E\). We consider then perturbation factors (PF) for the all genes \(g\) of the pathway \(K\).

\[\text{PF}(g) = \Delta E(g) + \sum_{\gamma \in U_g} \beta_{\gamma g} \frac{PF(\gamma)}{n_{\text{down}}(\gamma)}\]

Here, \(\Delta E(g)\) is the signed log_2-fold change of the gene \(g\) expression level in a given sample compared with the average value for the pool of control samples. The latter term expresses the summation over all the genes \(\gamma\) that belong to the set \(U_g\) of the upstream genes for the gene \(g\). The value of \(n_{\text{down}}(\gamma)\) denotes the number of downstream genes for gene \(\gamma\). The weight factor \(\beta_{\gamma g}\) indicates the interaction type between \(\gamma\) and \(g\): \(\beta_{\gamma g} = 1\) if \(\gamma\) activates \(g\) and \(\beta_{\gamma g} = -1\) when \(\gamma\) inhibits \(g\). We apply the depth-first search method to upstream/downstream genes [51,71].

We use the second term in the formula for the perturbation factor (PF) from the previous paragraph, resulting in the accuracy value, \(\text{Acc}(g) = \text{PF}(g) - \Delta E(g)\). Tarca et al. [71] showed that accuracy vector as expressed as follows:

\[\text{Acc} = B \cdot (I - B)^{-1} \cdot \Delta E,\]
Psychoactives 2023, 2

where

\[ B = \begin{pmatrix}
\beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\
\beta_{21} & \beta_{22} & \ldots & \beta_{2n} \\
\ldots & \ldots & \ldots & \ldots \\
\beta_{n1} & \beta_{n2} & \ldots & \beta_{nn}
\end{pmatrix},
\]

\[ I \text{ is the identity matrix, and}
\]

\[ \Delta E = \begin{pmatrix}
\Delta E(g_1) \\
\Delta E(g_2) \\
\ldots \\
\Delta E(g_n)
\end{pmatrix}.
\]

We calculated the overall score for the pathway perturbation, calculated as \[ \text{SPIA} = \sum_{g} \text{Acc}(g) \] [51,71].

According to the SPIA method [51,71], we calculated the activation levels for 241 signaling pathways. We borrowed the graph topology from the Qiagen pathway database [51,88].

We applied the SPIA calculation with and without the differential expression gene filter. If in a case sample the gene \( g \) had an expression level at least 1.5 times higher or lower than the median expression levels of the control samples, we considered this gene differentially expressed and used it for the SPIA calculations.

2.3. Evaluation of the Individual Drug Action

Having calculated the Oncobox PAL and SPIA values, we then performed the following statistical analysis [51],

1. Obtain the SPIA (PAL) values for each dataset and biological pathway.
2. Calculate the values of the pathway weight (\( w_p \)) factor as follows. For pathways with a positive mean SPIA (PAL) score of the case samples, \( w_p = ((\text{number of case samples with positive SPIA (PAL) score})/(\text{total number of case samples})) \). For pathways with a negative mean SPIA (PAL) score of the case samples, \( w_p = ((\text{number of case samples with negative SPIA (PAL) score})/(\text{total number of case samples})) \).
3. Adjust the mean SPIA (PAL) score of each pathway by the weight factor,

\[ \text{SPIA}_\mu (\text{PAL}_\mu) = \text{mean (SPIA(PAL))} \cdot w_p. \]

4. Perform a Student’s \( t \)-test if the values of SPIA\( _\mu \) (PAL\( _\mu \)) for the pool of case samples are different from 0 (for the pool of control samples, the values of SPIA\( _\mu \) (PAL\( _\mu \)) are clearly equal to 0). During the Student’s \( t \)-test, the following case classes are taken into account: the untreated case (\( U \)), i.e., the pathological state before drug application, should be far from the control (\( C \)).

4.1 The treated case (\( T \)), i.e., the pathological state after drug application, should be close to the control;

4.2 The following values are the results for such calculations: \[ |t_U| = \text{absolute } t\text{-value for the Student’s } t\text{-test for } U\text{-vs-}C \text{ profiles}; \]

4.3 \[ |t_T| = \text{absolute } t\text{-value for the Student’s } t\text{-test for } T\text{-vs-}C \text{ profiles}. \]

In addition to the old DEI metric [51] for individual drug action, \[ 2 \cdot \left( \frac{|t_U|}{|t_U| + |t_T|} - \frac{1}{2} \right) = \frac{|t_U| - |t_T|}{|t_U| + |t_T|}, \]

we have now introduced an alternative metric, called \( \text{mirrored DEI} \), calculated as follows,

\[
\text{DEI}_M = 2 \cdot \left( \frac{2 \cdot |t_U|}{2 \cdot |t_U| + |t_T + t_U|} - \frac{1}{2} \right) = \frac{2 \cdot |t_U| - |t_T + t_U|}{2 \cdot |t_U| + |t_T + t_U|}.
\]

The \( \text{DEI}_M \) metric equals 1 when \( t_T = -t_U \); this is the maximal possible value of this metric.
Similar to the old DEI metric, \( DEI_M = 0 \), when in \( t_T = t_U \), and when \( |t_T| \gg |t_U| \), \( DEI_M = -1 \).

The third metric, balanced DEI: \( DEI_B = \frac{CDEI + CDEI_M}{2} \), is the mean value over DEI and mirrored DEI.

### 2.4. Robust Marker Gene and Pathway Analysis

We applied (see Figure 1) the leave-one-out (LOO) procedure \([90,91]\) to complete the robust core marker gene/pathway set that distinguishes:

1. Case samples from control samples (PTSD or TRD), the disease-vs-healthy, D_vs_H test;
2. Case samples after treatment or/and observation from those before treatment (untreated), the post-vs-ante, P_vs_A test;
3. Patients positively responding to the treatment from negatively or non-responding patients, the responders-vs-non-responders, R_vs_N test.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Determination of robust marker gene/pathway sets for the case of multiple markers with maximally high quality.

We called a gene or a pathway upregulated if it had an expression or activation level higher for the disease state than for the healthy state after the treatment versus before the treatment and for the responders versus non-responders. If the expression/activation levels were lower, they were downregulated.

We selected the marker genes/pathways according to the area under the receiver-operating curve (AUC) metric. The ROC (receiver-operator curve) is a widely-used graphical plot that illustrates the diagnostic ability of a binary classifier system, as its discrimination threshold is varied. The ROC is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The area under the ROC curve, called ROC AUC, or simply AUC, is routinely employed for classifier quality \([92]\).

In the resulting robust marker gene and pathway sets, we selected only those genes/pathways with AUC > 0.70 (the conventional threshold for a significant marker). The robust set of marker features (genes or pathways) should assure the reliable distinction between different classes of samples (cases vs. controls, untreated vs. treated cases, and positively responding patients vs. those with no positive response). From the theory of machine learning, we know that the number of robust marker features should not exceed the number of samples. Since the smallest cohort, GSE860, contained 33 samples, we selected 30 as the maximal number of possible marker features (Figure 1).

For the genes in robust marker set, we retrieved the description for the corresponding protein and the relevant biological processes (Gene Ontology) using Metascape, an online resource for gene annotation and analysis \([93]\).

### 2.5. Alternative Methods for Differential Gene Expression and Enrichment Analysis

To check the overall pattern for marker genes/pathways for all three tests (D_vs_N, P_vs_A, and R_vs_N), we also applied the conventional procedure for differential gene expression (DEG) analysis. According to conventional procedure \([94]\), we considered a gene differentially expressed if it simultaneously satisfies the two criteria for DESeq2 test \([79]\): the adjusted \( p \)-value for the negative binomial model, \( p_{Adjusted} < 0.05 \), and the
log2-fold change, $|\text{LFC}_2| > 0.5$. We plotted the results of the DESeq2 analysis using the EnhancedVolcano R package.

3. Results

3.1. Cohorts Used in the Analysis

We aimed to analyze all available transcriptome data for PTSD/TRD patients using our previously published DEI approach. Our goal was to test whether the DEI approach would differentiate the responders vs. the non-responders, the case vs. control samples, and treated vs. non-treated based on their gene expression levels. We curated six cohorts of 791 gene expression profiles, including 379 case samples and 413 control samples (Table 2; see more details in Table 1).

Table 2. PTSD and control cohorts curated in the current work.

<table>
<thead>
<tr>
<th>Reference</th>
<th>[80]</th>
<th>[81]</th>
<th>[81]</th>
<th>[82]</th>
<th>[30,83]</th>
<th>[84]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE ID</td>
<td>GSE860</td>
<td>GSE64814- GPL6244</td>
<td>GSE64814- GPL11154</td>
<td>GSE81761</td>
<td>GSE97356</td>
<td>GSE185855</td>
</tr>
<tr>
<td>Cases</td>
<td>17</td>
<td>48</td>
<td>94</td>
<td>50</td>
<td>91</td>
<td>46</td>
</tr>
<tr>
<td>Controls</td>
<td>16</td>
<td>48</td>
<td>94</td>
<td>33</td>
<td>233</td>
<td>21</td>
</tr>
<tr>
<td>Number of genes</td>
<td>9068</td>
<td>12,631</td>
<td>10,184</td>
<td>22,878</td>
<td>25,830</td>
<td>55,178</td>
</tr>
</tbody>
</table>

3.2. Comparison of DEI Options for Different Approaches to the Assessment of Pathway Activation Levels

We have attempted three methods for calculation of PAL:

1. SPIA with differential expression filter [51];
2. SPIA without differential expression filter;
3. Oncobox PAL with the pathway databased OncoboxPD [88].

The DEI value we suggested [51] is a normalized indicator of treatment success. We measured the deviation of PAL values from the healthy reference using the Student’s $t$-test, which produces the metrics $t(U)$ and $t(T)$ for the untreated and treated cases, respectively. It equals 1 when the cell’s physiological conditions after treatment are indistinguishable from the healthy state. Higher DEI values may indicate that the treatment was efficient and that the post-treatment state is close to a healthy reference in terms of gene expression. The positive DEI value means that the post-treatment conditions are less dysregulated than the pre-treatment conditions, while a negative DEI value means that the post-treatment conditions are more dysregulated than the pre-treatment conditions. A strongly dysregulated post-treatment state may have a DEI value approaching $-1$.

To select the best PAL metric, we required that the corresponding DEI value is positive as frequently as possible. According to this criterion, we obtained the best results using the SPIA method with a differential expression filter; see Table 3.

However, for the TRD dataset GSE185855 [84], all the DEI scores for responders to ketamine treatment were unexpectedly negative, although they were insufficiently reliable in terms of the differential Student’s $t$-test with pathway activation values, at least for the patients without positive response to the treatment.

Additionally, for the datasets GSE81761 [82] and GSE185855 [84], the DEI values for the patients with no improvement were higher than for the patients with positive changes.

These negative values may indicate the following disadvantage of the DEI metric: it does not consider the sign flip from positive to negative, which indicates a transition through the zero value.

Perhaps, to overcome the disease, the treatment should change the pathway perturbation and the most effective treatment, which inverses the sign of the SPIA deviation from the control level. Thus, the best treatment should result in $t_T = -t_U$ rather than $t_T = 0$. 


Table 3. DEI metrics for different calculation methods of pathway activation levels.

<table>
<thead>
<tr>
<th>GSE ID, Clinical Remarks</th>
<th>GSE860</th>
<th>GSE64814-GPL16244 (MH)</th>
<th>GSE64814-GPL11154 (NGS)</th>
<th>GSE81761 No Improvement; Non-Responders</th>
<th>GSE81761 Improvement; Responders</th>
<th>GSE97356</th>
<th>GSE185855 TRD: Non-Responders</th>
<th>GSE185855 TRD: Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td>(U)</td>
<td>−5.27</td>
<td>−1.38</td>
<td>0.00</td>
<td>2.17</td>
<td>2.17</td>
<td>−2.29</td>
<td>−2.48</td>
<td>−0.53</td>
</tr>
<tr>
<td>(T)</td>
<td>3.35</td>
<td>0.49</td>
<td>0.00</td>
<td>−1.18</td>
<td>−2.12</td>
<td>0.49</td>
<td>1.21</td>
<td>−1.89</td>
</tr>
<tr>
<td>DEI</td>
<td>0.22</td>
<td>0.47</td>
<td>0.00</td>
<td>0.30</td>
<td>0.01</td>
<td>0.65</td>
<td>0.35</td>
<td>−0.56</td>
</tr>
<tr>
<td>DEIm</td>
<td>0.69</td>
<td>0.51</td>
<td>0.00</td>
<td>0.63</td>
<td>0.98</td>
<td>0.43</td>
<td>0.15</td>
<td>−0.39</td>
</tr>
<tr>
<td>DEIb</td>
<td>0.46</td>
<td>0.49</td>
<td>0.00</td>
<td>0.46</td>
<td>0.50</td>
<td>0.54</td>
<td>0.25</td>
<td>−0.47</td>
</tr>
<tr>
<td>p-value</td>
<td>$5 \times 10^{-8}$</td>
<td>0.11</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.16</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td>(U)</td>
<td>−0.19</td>
<td>5.50</td>
<td>−3.71</td>
<td>2.96</td>
<td>2.96</td>
<td>−1.20</td>
<td>−1.08</td>
<td>−0.56</td>
</tr>
<tr>
<td>(T)</td>
<td>5.69</td>
<td>−7.83</td>
<td>3.29</td>
<td>2.50</td>
<td>−3.71</td>
<td>−1.24</td>
<td>−1.56</td>
<td>−3.52</td>
</tr>
<tr>
<td>DEI</td>
<td>−0.94</td>
<td>−0.17</td>
<td>0.06</td>
<td>0.08</td>
<td>−0.11</td>
<td>−0.02</td>
<td>−0.19</td>
<td>−0.73</td>
</tr>
<tr>
<td>DEIm</td>
<td>−0.87</td>
<td>0.65</td>
<td>0.89</td>
<td>0.04</td>
<td>0.78</td>
<td>−0.01</td>
<td>−0.10</td>
<td>−0.57</td>
</tr>
<tr>
<td>DEIb</td>
<td>−0.90</td>
<td>0.24</td>
<td>0.48</td>
<td>0.06</td>
<td>0.33</td>
<td>−0.01</td>
<td>−0.14</td>
<td>−0.65</td>
</tr>
<tr>
<td>p-value</td>
<td>$9 \times 10^{-6}$</td>
<td>$9 \times 10^{-17}$</td>
<td>$5 \times 10^{-7}$</td>
<td>$3 \times 10^{-5}$</td>
<td>$1.4 \times 10^{-6}$</td>
<td>$0.70$</td>
<td>$0.51$</td>
<td>$0.004</td>
</tr>
</tbody>
</table>

Oncobox Pathway Database [88] without differential expression filter [51]

| (U)                      | −0.75  | 11.21                   | −8.63                    | 4.52                                     | 4.52                            | −0.67    | 0.02                          | −3.06                    |
| (T)                      | 15.79  | −16.64                  | 7.41                     | 4.93                                     | −5.41                           | 2.26     | 0.95                          | −11.53                   |
| DEI                      | −0.91  | −0.20                   | 0.08                     | −0.04                                    | −0.09                           | −0.54    | −0.95                         | −0.58                    |
| DEIm                     | −0.82  | 0.61                    | 0.87                     | −0.02                                    | 0.82                            | −0.08    | −0.90                         | −0.41                    |
| DEIb                     | −0.86  | 0.21                    | 0.47                     | −0.03                                    | 0.37                            | −0.31    | −0.93                         | −0.49                    |
| p-value                  | $3 \times 10^{-38}$ | $6 \times 10^{-12}$           | $5 \times 10^{-27}$       | $6 \times 10^{-11}$                      | $4 \times 10^{-11}$             | $0.03$   | $0.30$                        | $1.0 \times 10^{-14}$     |

\(t(U)\)—Student’s t-test result for the untreated samples, \(t(T)\)—for the treated samples, DEI—drug efficiency index, DEIm—mirrored DEI, DEIb—balanced DEI, an average of the DEI and DEIm, p-values were calculated for the one-sided differential treated vs. untreated Student’s t-test with the pathway activation values.

Therefore, in addition to the old DEI [51], we introduced an alternative metric, called mirrored DEI (DEIm), as well as the average of the DEI and DEIm, called balanced DEI (DEIb) (see Materials and Methods).

Note that the DEIm and DEIb metrics helped the GSE81761 cohort but not the GSE185855 cohort, where the results for non-responders were always statistically unconvincing according to the Student’s t-test (Table 3).

3.3. Robust Marker Gene/Pathway Analysis, Differentially Expressed Gene (DEG) Analysis, and Gene Ontology (GO) Enrichment Analysis

To have a closer and more in-depth view of the pattern of genes and pathways that discriminate different types of samples from one another, we identified robust marker sets [90,91] of genes and pathways for three different comparison groups described above: D_vs_H, P_vs_A and R_vs_N.

For these three tests, we considered a gene/pathway upregulated when the expression/activation level was higher for the case samples than the control samples, for the post-treatment cases than the pre-treatment counterparts, and for the patients with clinical improvements than the persons with no improvement.

Using the online gene analytical portal Metascape [93], we specified the protein description and biological functions (Gene Ontology) for the identified robust marker genes.

Additionally, for these three comparison groups, we analyzed the differentially expressed genes using the DESeq2 tool [79].

We analyzed the intersection of robust marker genes and pathway sets between differential tests for all cohorts (Figure 2) and between cohorts for all comparison groups (Figure 3).

Table 4 shows the number of robust marker genes (marked as Exp) and pathways. We calculated the pathway activation levels using the SPIA method with the differential expression filter (marked as SPIA). These genes and pathways are also listed in Supplementary Files S1–S6.
Table 4. Number of robust marker genes (Exp) and pathways for three differential tests: D_vs_H, P_vs_A, and R_vs_N.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Exp_D_vs_H</th>
<th>Exp_P_vs_A</th>
<th>Exp_R_vs_N</th>
<th>SPIA_D_vs_H</th>
<th>SPIA_P_vs_A</th>
<th>SPIA_R_vs_N</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE860</td>
<td>12</td>
<td>7</td>
<td>-</td>
<td>16</td>
<td>21</td>
<td>-</td>
</tr>
<tr>
<td>GSE64814/GPL6244</td>
<td>3</td>
<td>21</td>
<td>-</td>
<td>0</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>GSE64814/GPL11154</td>
<td>0</td>
<td>21</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>GSE97356</td>
<td>0</td>
<td>15</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>GSE81761</td>
<td>12</td>
<td>7</td>
<td>16</td>
<td>4</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>GSE185855</td>
<td>8</td>
<td>9</td>
<td>11</td>
<td>15</td>
<td>28</td>
<td>19</td>
</tr>
</tbody>
</table>

Exp shows the number of robust marker genes, while SPIA shows the number of robust pathways.

Figure 2. Overlaps between the dysregulated genes found within different cohorts. Intersection of robust marker gene sets (A) and SPIA pathways (B) for three comparison groups (D_vs_H, P_vs_A, and R_vs_N) for different cohorts. Numbers indicate the number of unique or overlapping genes or pathways for three comparison groups within the same cohort.
Psychoactives 2023, 2, 101

Figure 3. Overlaps between the dysregulated genes found between different cohorts. Intersection of robust marker gene sets (A) and SPIA pathways (B) for three comparison groups (D_vs_H, P_vs_A, and R_vs_N) for different cohorts. Numbers indicate the number of unique or overlapping genes or pathways for different cohorts.

A detailed description of the function of genes in various comparisons is presented in Supplementary File S7.

As can be seen from the analysis above, the number of the robust marker genes for all tests and datasets was relatively small, with little to no overlap between cohorts. To confirm this data, we also performed a classical DEG analysis, where we selected the genes with an expression level higher or lower than log2 of 0.5 and an adjusted p = value lower than 0.05 (see methods section for details).

Table 5 summarizes the data seen in Figures 3–5, and Supplementary File S8 has a list of all genes significantly (p < 0.05) altered by log2 > 0.5 or log2 < −0.5. Cohort GSE81761 had a large number of genes in two categories, A_vs_P and R_vs_N; in both, most of the genes were downregulated. Cohort GSE64814-GPL6244 also contains a large number of genes in A_vs_P category.
Table 5. Number of up- and down-regulated genes per DESeq2.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>D_vs_H Up</th>
<th>D_vs_H Down</th>
<th>A_vs_P Up</th>
<th>A_vs_P Down</th>
<th>R_vs_N Up</th>
<th>R_vs_N Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE185855</td>
<td>5</td>
<td>93</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSE81761</td>
<td></td>
<td></td>
<td>759</td>
<td>16</td>
<td>567</td>
<td></td>
</tr>
<tr>
<td>GSE64814-GPL11154</td>
<td>25</td>
<td>2</td>
<td>567</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSE64814-GPL6244</td>
<td>33</td>
<td>1</td>
<td>341</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSE860</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>GSE97356</td>
<td>6</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Differential gene expression analysis for the D_vs_H test, i.e., PTSD/TRD vs. healthy datasets in GSE860 (A), GSE185855 (B), GSE64814-GPL6244 (C), GSE64814-GPL11154 (D), GSE81761 (E), GSE97356 (F) cohorts. In green, genes with expression values of log2 > 0.5 and log2 < −0.5; in blue, genes significantly (p < 0.05) differently expressed; in red, genes that are significantly differently expressed with the differences of log2 > 0.5 and log2 < −0.5.
Figure 5. Differential gene expression analysis for the P_vs_A test, i.e., PTSD/TRD patients after (post) observation and/or treatment and before (ante) in GSE860 (A), GSE185855 (B), GSE64814-GPL6244 (C), GSE64814-GPL11154 (D), GSE81761 (E), GSE97356 (F) cohorts. In green, genes with the expression values of log2 > 0.5 and log2 < −0.5; in blue, genes significantly (p < 0.05) differently expressed; in red, genes that are significantly differently expressed with the differences of log2 > 0.5 and log2 < −0.5.

We then compared the number of robust genes (as per AUC criteria) and DESeq2 genes. It appears that only for cohort GSE860 was the number of genes higher per AUC criteria. For the GSE97356 and GSE64814-GPL11154 cohorts, the numbers were comparable, while for GSE64814-GPL6244, GSE81761 and GSE185855, the numbers were higher in the DESeq2 category. The most dramatic differences were observed for the GSE81761 cohort, especially in P_vs_A and R_vs_N categories.

We then analyzed whether there are any overlapping genes in both categories (Table 6 and Supplementary File S9). For the GSE860 cohort, there was one gene overlap in the D_vs_H category: PAX3, encoding the protein involved in sensory perception of sound and mechanical stimuli, as well as muscle development. In the GSE64814-GPL6244 cohort, D_vs_H also had one gene overlap: EPSTI1—epithelial stromal interaction 1, while P_vs_A had an overlap of 17 genes. Among these, there was a large number of genes encoding snoRNAs, including SCARNA7, SCARNA17, SNORD6, SNORD24, SNORDD43, SNORD68,
and SNORA65 as well as regulators of transcription, DRAP1, PHTF1, and NELFE. For the GSE81761 cohort, there were seven gene overlaps in the R_vs_N category, all encoding Zn finger proteins involved in spindle formation, regulation of non-coding RNA biogenesis, and regulation of transcription and protein ubiquitination. For the GSE185855 cohort, there was a complete overlap of 11 genes between AUC and DESeq2 analyses for the R_vs_N category; they included proteins involved in differentiation of natural killer cells, basophils, and erythrocytes; suppression by virus of apoptotic processes; regulation of autophagy in neurons; and the adiponectin-activated signaling pathway.

Table 6. Number of robust maker genes (according to the AUC criteria), differentially expressed genes (according to the DESeq2 criteria), and overlapping genes.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>AUC D_vs_H</th>
<th>AUC P_vs_A</th>
<th>AUC R_vs_N</th>
<th>DESeq2 D_vs_H</th>
<th>DESeq2 P_vs_A</th>
<th>DESeq2 R_vs_N</th>
<th>Overlap D_vs_H</th>
<th>Overlap P_vs_A</th>
<th>Overlap R_vs_N</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE860</td>
<td>12</td>
<td>7</td>
<td>-</td>
<td>4</td>
<td>0</td>
<td>-</td>
<td>1</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>GSE64814-GPL6244</td>
<td>3</td>
<td>21</td>
<td>-</td>
<td>34</td>
<td>416</td>
<td>-</td>
<td>1</td>
<td>17</td>
<td>-</td>
</tr>
<tr>
<td>GSE64814-GPL11154</td>
<td>0</td>
<td>21</td>
<td>-</td>
<td>25</td>
<td>6</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>GSE97356</td>
<td>0</td>
<td>15</td>
<td>-</td>
<td>19</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>GSE81761</td>
<td>12</td>
<td>7</td>
<td>16</td>
<td>0</td>
<td>759</td>
<td>583</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>GSE185855</td>
<td>8</td>
<td>9</td>
<td>11</td>
<td>5</td>
<td>105</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
</tbody>
</table>

The failure of DEI, DEIm, and DEIb metrics to reflect differences between responders and non-responders for the TDR cohort (GSE185855) may be due to the small number, or lack of, differentially expressed genes in the D_vs_H and P_vs_A tests for this cohort (Figures 4 and 5, respectively), which makes the PAL/SPIA and DEI calculations highly unreliable. High p-values for the differential Student’s t-test with pathway values for untreated and treated non-responders confirm this unreliability. On the other hand, the differential genes for the R_vs_N test with the GSE185855 dataset were rather numerous (Figure 6A)—most of these genes showed a more drastic dysregulation in the responder group, thus reflecting the negative value for DEI/DEIm/DEIb (Table 3).

Figure 6. Differential gene expression analysis for the R_vs_N test, i.e., PTSD/TRD patients who were responsive or non-responsive to treatments in GSE185855 (A) and GSE81761 (B) cohorts. In green, genes with the expression values of log2 > 0.5 and log2 < −0.5; in blue, genes significantly (p < 0.05) differently expressed; in red, genes that are significantly differently expressed with the differences of log2 > 0.5 and log2 < −0.5.

4. Future Perspectives

The omics-based personalized approach has been popular for cancer [52,95], hereditary diseases [96,97], endocrine disorders [98–100], immunological/infectious diseases [101–103],
neurological conditions [104], regenerative medicine/geriatrics [98,105,106], etc., for approximately a decade.

The systems bioinformatics approach to personalized drug prescription uses different paradigms. First, we can focus on pathway activation data, which have higher reliability and robustness than single/handfuls of gene expression and mutation profiles [89]. We can use aggregated (cumulative) PAL and mutation burden values as personalized markers for the analysis of drug efficiency [67]. We refer to this paradigm for personalized drug prescription as an a posteriori, or non-hypothesis-driven, approach. It relies on direct mathematical association of molecular and clinical features in the training dataset, using machine learning (ML)/artificial intelligence (AI) applications [90,91].

Another paradigm, an a priori, or hypothesis-driven, approach links molecular pathway activation with specific targets of drugs and ranks these drugs according to their ability to restore the healthy state within the cell [52]. Despite its greater complexity, we can say that this approach may be more appropriate for clinical implementation.

On the other hand, we have already proposed and verified (for the action of cannabis drugs on the skin, oral, and intestinal epithelial cells) the method and software for omics-based personalized scoring of treatment efficiency. Like many, if not most, other omics-based personalized medicine projects, our DEI (drug efficiency index) metric [51] uses the following methodology:

- The Big Data full expression profiles, obtained via ether microarray hybridization (MH) or next-generation sequencing (NGS) of mRNA are used as input data for PAL calculations;
- For each gene, we calculate the case-to-control log-fold changes (LFCs) using full expression profiles;
- We then unitize these LFCs for the assessment of PAL, including Oncobox PAL [67], or SPIA (signaling pathway impact analysis) metrics [71], according to the protein–protein interactions graphs for each pathway.

However, unlike the Oncobox [52] approach to the assessment of personalized drug efficiency, our DEI [51] method does not exploit the purely a priori paradigm. Instead, it compares the SPIA values before and after treatment to check which state (pre- or post-treatment) differs more from healthy conditions. Therefore, our DEI method is more a posteriori by type than the Oncobox approach.

To check whether our DEI method is helpful in psychiatry, we have attempted it for five PTSD cohorts and one TRD cohort. These six cohorts enlisted 379 cases and 413 control samples (overall, 791 profiles). Except for the TRD cohort (GSE185855), which contained patients treated with ketamine, all patients from the five PTSD cohorts did not receive any pharmaceutical treatment; instead, they were under observation during psychotherapy. Thus, we could not apply the Oncobox drug efficiency scoring (DES) approach [52,107] since psychotherapy has no molecular targets.

Our analysis of the DEI metric for the psychotherapy of PTSD, as well as for the ketamine therapy of TRD, implied the following differential analysis of PAL values (both SPIA and Oncobox PAL), as well as of gene expression values:

- Case samples (PTSD or TRD) from control samples;
- Case samples after treatment or/and observation from before treatment (untreated);
- Samples from patients positively responding to the treatment from responding negatively or non-responding.

To assess the adequacy of the DEI metric for treatment success, we checked whether DEI values are positive as frequently as possible. According to this criterion, we achieved the best results for the DEI with SPIA values with a differential expression filter. The SPIA values without differential expression filter and Oncobox PAL values produced worse DEI results.

Only one PTSD (GSE81761) and the only TRD cohort (GSE185855) contained information on the clinical success of the therapy. For these cohorts, we tested the quality of the
DEI metric in identifying responders versus non-responders only in these two cohorts; our analysis could differentiate, that is, provide the higher values for the responders.

We introduced two DEI modifications, mirrored DEI (DEIm) and balanced DEI (DEIb, the mean value over our original DEI, and mirrored DEI). The DEIm and DEIb values provided the proper difference between groups for the PTSD cohort. However, they did not produce meaningful results for the TRD cohort, perhaps because there were no reliable differentially expressed genes (DEG) in the disease-vs-healthy and after-vs-befetreatment tests. This lack of DEG made every PAL, SPIA, and DEI calculation questionable, despite the abundance of DEG in the responders-vs-non-responders comparison.

The analysis of dysregulated genes showed the genes falling into the following categories: gene expression regulators on the level of chromatin, transcription, splicing, mRNA stability, and histone ubiquitination; immune response to bacteria, viruses, and fungi; signal transduction; ion transport; neuron morphogenesis, autophagy, and oxidative stress; cell division and differentiation. On a molecular level, dysregulation of these genes may cause multiple symptoms of PTSD. Indeed, a number of reports documented the dysfunction of the immune system in PTSD patients, including an increase in pro-inflammatory cytokines, immunosenescence, and decreased circulation of T cells [108]. PTSD also shows changes in neuronal morphology, disrupted neural plasticity, and substantial remodeling of dendrites [109].

Another question for the further application of the DEI/DEIm/DEIb metrics concerns cohort merging. For example, the cohort GSE64814 contained two sub-cohorts, one with data obtained with microarray hybridization (MH) Affymetrix Human Gene 1.0 ST Array (GPL6244), and the other with data obtained by sequencing with the NGS device Illumina HiSeq 2000 (GPL11154). Some methods can merge the unlimited number of MH and NGS cohorts. One of the most general examples is the uniformly shaped harmonization algorithm; however, this algorithm may affect the LFC values [110]. Therefore, this may invalidate the SPIA calculations with the differential expression filter. Thus, applying DEI/DEIm/DEIb metrics after transcriptomic harmonization may require more detailed research.

5. Conclusions

There is no consensus on the best treatment strategy for PTSD; some practitioners argue that the proverb “time cures” suggests the most proper treatment. To assess the clinical aspects of PTSD at the molecular and cell physiology level, we have applied the multi-omics personalized medicine approach to PTSD treatment. This approach utilizes Big Data on multi-omics (genome, transcriptome, proteome, etc.) profiling to study the physiological properties of healthy and pathological samples at the level of different cellular pathways.

Previously, we introduced the DEI, a metric for a drug’s ability to restore a healthy state in the gene expression pathways. Here, we have applied the DEI concept to the transcriptomics data from six cohorts of PTSD and TRD patients. We found that the best treatment should not only make the pathway activation levels similar to the healthy state (as we suggested previously) but should also encourage the pathological changes to disappear by reverting the sign of pathway activation levels.

In addition, we concluded that any type of DEI is adequate only when there are significant differences between the healthy vs. pathological and treated vs. untreated gene expression profiles. The differential gene expression analysis (based on the area under the receiver-operating curve (robust AUC) and negative binomial distribution (DESeq2)) confirms this finding.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/psychoactives2020007/s1, File S1: Robust marker gene (with Gene Ontology) and pathway set for the dataset GSE860; File S2: Robust marker gene (with Gene Ontology) and pathway set for the dataset GSE64814-GPL6244; File S3: Robust marker gene (with Gene Ontology) and pathway set for the dataset GSE64814-GPL11154; File S4: Robust marker gene (with Gene Ontology) and pathway set for the dataset GSE97356; File S5: Robust marker gene (with
Gene Ontology) and pathway set for the dataset GSE81761; File S6: Robust marker gene (with Gene Ontology) path pathway set or the dataset GSE185855; File S7: Description of function of genes in various cohort comparisons; File S8: Differentially expressed genes from different cohorts and categories; File S9: Genes overlapping between the robust maker genes (AUC criteria) and the differentially expressed genes (DESeq2 criteria).

**Author Contributions:** Conceptualization, I.K., O.K.; methodology, N.B., B.B., Y.I.; software, B.B., N.B.; validation, N.B., I.K.; formal analysis, N.B.; resources, B.B., I.K., O.K.; data curation, N.B., I.K.; writing—original draft, N.B., I.K.; writing—review and editing, N.B., I.K., O.K.; visualization, N.B., I.K.; supervision, I.K., O.K.; funding acquisition, I.K., O.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by MITACS, grant numbers IT11447, IT15089.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data from all cohorts analyzed here have been published and listed in the Methods Section.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial intelligence</td>
</tr>
<tr>
<td>APAP</td>
<td>Automatic positive airway pressure</td>
</tr>
<tr>
<td>ARR</td>
<td>Activator/repressor role</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under curve</td>
</tr>
<tr>
<td>BEP</td>
<td>Brief eclectic psychotherapy</td>
</tr>
<tr>
<td>CBT-i</td>
<td>Cognitive behavioral therapy for insomnia</td>
</tr>
<tr>
<td>CDEI</td>
<td>Cannabis drug efficiency index</td>
</tr>
<tr>
<td>DEG</td>
<td>Differentially expressed genes</td>
</tr>
<tr>
<td>DEI</td>
<td>Drug efficiency index</td>
</tr>
<tr>
<td>DEIb</td>
<td>Balanced drug efficiency index</td>
</tr>
<tr>
<td>DEIm</td>
<td>Mirrored drug efficiency index</td>
</tr>
<tr>
<td>DES</td>
<td>Drug efficiency score</td>
</tr>
<tr>
<td>DESeq2</td>
<td>Differential gene expression analysis in sequencing 2</td>
</tr>
<tr>
<td>EMDR</td>
<td>Eye movement desensitization and reprocessing</td>
</tr>
<tr>
<td>FloWPS</td>
<td>Floating-window projective separator</td>
</tr>
<tr>
<td>GAD</td>
<td>Generalized anxiety disorder</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>iPANDA</td>
<td>In silico pathway activation network decomposition analysis</td>
</tr>
<tr>
<td>LFC</td>
<td>Log-fold change</td>
</tr>
<tr>
<td>MDD</td>
<td>Major depressive disorder</td>
</tr>
<tr>
<td>MH</td>
<td>Microarray hybridization</td>
</tr>
<tr>
<td>ML</td>
<td>Machine learning</td>
</tr>
<tr>
<td>NET</td>
<td>Narrative exposure therapy</td>
</tr>
<tr>
<td>NGS</td>
<td>Next-generation sequencing</td>
</tr>
<tr>
<td>NII</td>
<td>Node involvement index</td>
</tr>
<tr>
<td>PAL</td>
<td>Pathway activation level</td>
</tr>
<tr>
<td>PE</td>
<td>Pathway-express</td>
</tr>
<tr>
<td>PTSD</td>
<td>Post-traumatic stress disorder</td>
</tr>
<tr>
<td>QN</td>
<td>Quantile normalization</td>
</tr>
<tr>
<td>SNRI</td>
<td>Serotonin–norepinephrine reuptake inhibitor</td>
</tr>
<tr>
<td>SPIA</td>
<td>Signaling pathway impact analysis</td>
</tr>
<tr>
<td>SSRI</td>
<td>Selective serotonin reuptake inhibitors</td>
</tr>
<tr>
<td>TAPPA</td>
<td>Topology analysis of pathway phenotype association</td>
</tr>
<tr>
<td>TBscore</td>
<td>Topology-based score</td>
</tr>
<tr>
<td>TRD</td>
<td>Treatment-resistant depression</td>
</tr>
</tbody>
</table>
References


5. Edmondson, D. An Enduring Somatic Threat Model of Posttraumatic Stress Disorder Due to Acute Life-Threatening Medical Events: Enduring Somatic Threat Model. Soc. Person. Psychol. Compass 2014, 8, 118–134. [CrossRef]

6. McHugh, P.R.; Treisman, G. PTSD: A Problematic Diagnostic Category. J. Anxiety Disord. 2007, 21, 211–222. [CrossRef]


9. Breslau, N. The Epidemiology of Trauma, PTSD, and Other Posttrauma Disorders. Trauma Violence Abus. 2009, 10, 198–210. [CrossRef]


36. Sbarski, B.; Akirav, I. Cannabinoids as Therapeutics for PTSD. Pharmacol. Ther. 2020, 211, 107551. [CrossRef]

37. Stanciu, C.N.; Brunette, M.F.; Teja, N.; Budney, A.J. Evidence for Use of Cannabinoids in Mood Disorders, Anxiety Disorders, and PTSD: A Systematic Review. Psychiatr. Serv. 2021, 72, 429–436. [CrossRef]


40. Fraguas-Sánchez, A.I.; Torres-Suárez, A.I. Medical Use of Cannabinoids. Drugs 2018, 78, 1665–1703. [CrossRef]

41. Loflin, M.J.; Babson, K.A.; Bonn-Miller, M.O. Cannabinoids as Therapeutic for PTSD. Curr. Opin. Psychol. 2017, 14, 78–83. [CrossRef]


60. Li, X.; Shan, C.; Wu, Z.; Yu, H.; Yang, A.; Tan, B. Emodin Alleviated Pulmonary Inflammation in Rats with LPS-Induced Acute Lung Injury through Inhibiting the MTOR/HIF-1α/VEGF Signaling Pathway. *Inflamm. Res.* **2020**, *69*, 365–373. [CrossRef]


90. Tkachev, V.; Sorokin, M.; Mescheryakov, A.; Simonov, A.; Garazha, A.; Buzdin, A.; Muchnik, I.; Borisov, N. FLOating-Wind Projective Separator (FloWPS): A Data Trimming Tool for Support Vector Machines (SVM) to Improve Robustness of the Classifier. *Front. Genet.* 2019, 10, 717. [CrossRef]


100. Zhang, M.; Luo, H.; Xi, Z.; Rogova, E. Drug Repositioning for Diabetes Based on “omics” Data Mining. *PLoS ONE* 2015, 10, e0126082. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.