

Machine Learning and Deep Learning for the Pharmacogenomics of Antidepressant Treatments

Eugene Lin^{1,2,3}, Chieh-Hsin Lin^{3,4,5}, Hsien-Yuan Lane^{3,6,7,8}

¹Department of Biostatistics, ²Department of Electrical & Computer Engineering, University of Washington, Seattle, WA, USA, ³Graduate Institute of Biomedical Sciences, China Medical University, Taichung, ⁴Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital, Chang Gung University College of Medicine, Kaohsiung, ⁵School of Medicine, Chang Gung University, Taoyuan, ⁶Department of Psychiatry, ⁷Brain Disease Research Center, China Medical University Hospital, ⁸Department of Psychology, College of Medical and Health Sciences, Asia University, Taichung, Taiwan

A growing body of evidence now proposes that machine learning and deep learning techniques can serve as a vital foundation for the pharmacogenomics of antidepressant treatments in patients with major depressive disorder (MDD). In this review, we focus on the latest developments for pharmacogenomics research using machine learning and deep learning approaches together with neuroimaging and multi-omics data. First, we review relevant pharmacogenomics studies that leverage numerous machine learning and deep learning techniques to determine treatment prediction and potential biomarkers for antidepressant treatments in MDD. In addition, we depict some neuroimaging pharmacogenomics studies that utilize various machine learning approaches to predict antidepressant treatment outcomes in MDD based on the integration of research on pharmacogenomics and neuroimaging. Moreover, we summarize the limitations in regard to the past pharmacogenomics studies of antidepressant treatments in MDD. Finally, we outline a discussion of challenges and directions for future research. In light of latest advancements in neuroimaging and multi-omics, various genomic variants and biomarkers associated with antidepressant treatments in MDD are being identified in pharmacogenomics research by employing machine learning and deep learning algorithms.

KEY WORDS: Antidepressive agents; Artificial intelligence; Deep learning; Genomics; Machine learning; Neuroimaging.

INTRODUCTION

Pharmacogenomics, an interdisciplinary field of pharmacology and genomics, is progressing into indispensable practices in medicine with the prospect of individualized clinical care for patients [1]. Namely, pharmacogenomics is one of the research fields to further the advances of precision medicine [2]. In terms of psychiatry, pharmacogenomics aims to adapt medical decisions, treatments, and

practices to individual patients with psychiatric disorders [3]. More precisely, the full population of patients with psychiatric disorders are partitioned into multiple subgroups by distinct biomarkers, where genomic data is available to serve as distinct biomarkers in light of pharmacogenomics in psychiatry. Accordingly, medications can be tailored personally to each specific patient based on comparable or applicable genomic biomarkers [3]. Nowadays, there are gradually developing genomic biomarkers that might be linked to adverse drug reactions and clinical treatment response for patients with psychiatric disorders [4]. For example, it has long been indicated that genomic biomarkers such as gene expression profiles and single nucleotide polymorphisms (SNPs) can be leveraged to assess antidepressant treatment response and remission in patients with major depressive disorder (MDD) [5,6].

In the field of pharmacogenomics, scientists incorporate

Received: February 9, 2021 / **Accepted:** April 10, 2021

Address for correspondence: Chieh-Hsin Lin

Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital, Chang Gung University College of Medicine, No.123, Dapi Road, Niasong District, Kaohsiung 83301, Taiwan
E-mail: cyndi36@gmail.com

ORCID: <https://orcid.org/0000-0001-6949-8968>

Hsien-Yuan Lane

Department of Psychiatry, China Medical University Hospital, No. 2, Yude Road, North District, Taichung 40447, Taiwan
E-mail: hylane@gmail.com

ORCID: <https://orcid.org/0000-0003-2162-8174>

© This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

various data types such as multi-omics data with state-of-the-art machine learning and deep learning models, which can subsequently pinpoint complicated patterns corresponding to experimental data [7-9]. That is, multi-omics data is utilized to represent predictive parameters (or biomarkers) to achieve the concept of pharmacogenomics by employing machine learning and deep learning models. In order to meet the pressing issues we confront currently in the field of pharmacogenomics in MDD, there is a vast demand for establishing bioinformatics frameworks in machine learning and deep learning models that can forecast antidepressant treatment outcome in clinical contexts based on next-generation multi-omics data [10].

Due to latest advancements in artificial intelligence research, machine learning and deep learning algorithms have affirmed their encouraging merits of learning and observing nonlinear and complicated hierarchical patterns corresponding to enormous large-scale multi-omics datasets [11-15]. Basically, the goal of deep learning algorithms is to construct artificial intelligence and machine learning models which utilize numerous layers of abstraction such as artificial neural networks to build a hierarchical representation for the datasets [16,17]. In other words, deep learning models for prediction tasks, such as the prediction of antidepressant treatment outcome in pharmacogenomics, are computation models for evaluating the best inference by using artificial neural networks with various layers, instead of using artificial neural networks with only one single layer [12-15,18]. Moreover, deep learning models have achieved state-of-the-art performances on a wide variety of prediction tasks in pharmacogenomics by using modern computing technologies such as the invention of general-purpose computing on graphics processing units [8,12-15,19,20].

With the recent advance in neuroimaging technologies, the integration of research on pharmacogenomics and neuroimaging in psychiatry demonstrates the massive growth possibilities of the demands for novel therapeutic strategies as well as novel medicine indications for treatment and therapeutic interventions in the future [21,22]. In addition, the usage of machine learning and deep learning approaches may play a crucial role in integrating pharmacogenomics with neuroimaging [23]. In the recent past, there were a wide range of relevant research studies for various antidepressant treat-

ments of significance for neuroimaging with consideration of machine learning approaches [24,25]. Hence, it would be remarkably intriguing to develop machine learning and deep learning models that can forecast antidepressant treatment response and remission for MDD patients by incorporating pharmacogenomics with neuroimaging. To address this challenge, machine learning and deep learning models may be able to produce suitable bioinformatics tools to accomplish the promise of pharmacogenomics by taking advantage of specific genomic biomarkers and neuroimaging data for antidepressant treatment response and remission.

At the moment, researchers have been making compelling progress in the interdisciplinary fields of pharmacogenomics, machine learning, deep learning, multi-omics, and neuroimaging [8,9,23]. The objective of machine learning and deep learning algorithms is to implement data-driven models that can generally learn from the data in the present and/or in the past by using the learned observation into analyzing predictive results (for instance, the prediction of antidepressant treatment outcome in pharmacogenomics) for any unexplored event and/or for any unknown data in the future [26-28]. In the normal contexts, the guidance of machine learning and deep learning algorithms encompasses the following three processes. First, the predictive algorithm is built from the initial input data in the beginning process. Second, the predictive algorithm is then gauged and fine-tuned in the intermediate process. Third, the predictive algorithm is utilized for generating predictive results in the final process [26-28].

In this review, we describe recent research studies in pharmacogenomics, which evaluated drug treatments by using artificial intelligence methods such as machine learning and deep learning models. Additionally, we especially focus on the integration of research on pharmacogenomics and neuroimaging for antidepressant treatment outcome in patients with MDD. Furthermore, we summarize the limitations in these research studies and show a discussion of directions as well as challenges in the future. Here, in the context of machine learning and deep learning algorithms, we present various research studies with focus on antidepressant treatment prediction in patients with MDD in terms of pharmacogenomics using multi-omics datasets (Fig. 1). Biological and/or clinical implications from this major arena may serve as decision

support aides for treatment prediction in precision psychiatry and translational psychiatry [29]. While this review does not support the full set of related research studies reported in the literature, it nonetheless demonstrates a synthesis of those that can considerably influence population health-oriented and public health-relevant applications in MDD, antidepressants, and pharmacogenomics in the near to mid-term future.

RESEARCH STUDIES IN PHARMACOGENOMICS ON THE PREDICTION OF ANTIDEPRESSANT TREATMENT OUTCOME IN PATIENTS WITH MDD

The usage of artificial intelligence and predictive algorithms in terms of predicting antidepressant treatment outcome has been a focus of attention in pharmacogenomics research. Here, we focus on antidepressant treatment outcome by using traditional machine learning algorithms as well as deep learning algorithms in this section (Table 1).

We first conducted a comprehensive search of the electronic PubMed database (2015–present) using key words such as “machine learning,” “deep learning,” “antidepressant,” “major depressive disorder,” “multi-omics,” and “pharmacogenomics”. Then, we manually screened the obtained articles with a particular focus on MDD. The multi-omics data in these studies included SNPs datasets, DNA methylation datasets, gene expression datasets, and phenotypic datasets (such as demographic and clinical datasets). In addition, the reader can refer to a recent review by Amare *et al.* [4] for genetic variants of antidepressant treatment response in MDD as well as a recent review by Perlman *et al.* [30] for predictors of antidepressant treatment response in MDD.

Traditional Machine Learning Methods

To predict antidepressant treatment response, various research studies employed traditional machine learning methods, including linear regression models, support vector machine, decision tree, random forests, and ensemble models (Fig. 1, Table 1).

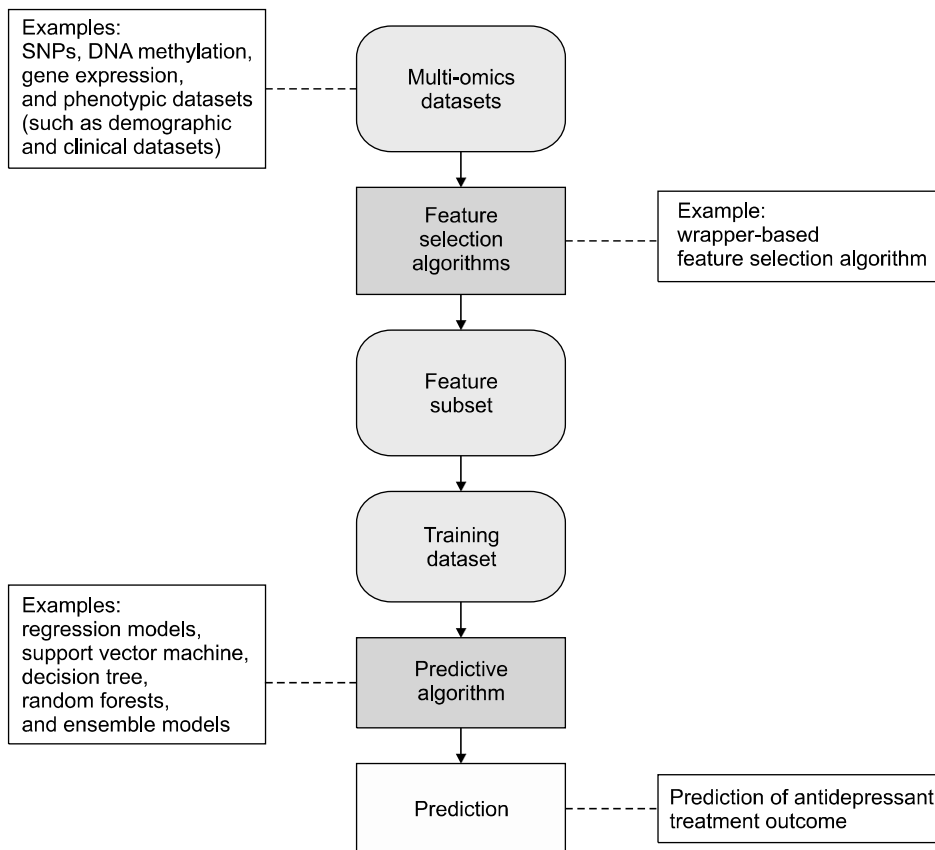


Fig. 1. An example of the machine learning framework for forecasting antidepressant treatment response. The machine learning model comprises two major components including a feature selection algorithm and a predictive algorithm. The feature selection algorithm produces a small subset of good features, which serves as the training dataset for subsequent analysis. The training dataset serves as the input for the predictive algorithm. The predictive algorithm estimates the prediction of antidepressant treatment outcome from the training dataset. SNPs, single nucleotide polymorphisms.

Table 1. Relevant studies on the predictive models of evaluating antidepressant treatment response and/or remission

Study	Model	Reference	Dataset	Results
Pharmacogenomics	Decision trees	Maciukiewicz <i>et al.</i> [32]	SNPs	An accuracy of 57% (response)
Pharmacogenomics	SVM	Maciukiewicz <i>et al.</i> [32]	SNPs	An accuracy of 64% (response)
Pharmacogenomics	Random forest	Kautzky <i>et al.</i> [37]	SNPs and clinical datasets	An accuracy of 25% (response)
Pharmacogenomics	Random forest	Athreya <i>et al.</i> [38]	SNPs and clinical datasets	AUC > 0.7 and accuracy > 69% (response)
Pharmacogenomics	Ensemble	Lin <i>et al.</i> [42]	SNPs and clinical datasets	AUC = 0.83 (response); AUC = 0.81 (remission); Comparable performance to MFNNs and logistic regression; Better performance than SVM, C4.5 decision tree, naive Bayes, and random forests.
Pharmacogenomics	Deep learning model	Lin <i>et al.</i> [45]	SNPs and clinical datasets	AUC = 0.82 (response); AUC = 0.806 (remission)
Neuroimaging	Decision trees	Patel <i>et al.</i> [49]	Structural MRI and clinical datasets	An accuracy of 89% (response)
Neuroimaging pharmacogenomics	Linear regression	Chang <i>et al.</i> [60]	MRI, SNPs, DNA methylation, and demographic data	An accuracy of 84% (response)
Neuroimaging pharmacogenomics	SVM	Pei <i>et al.</i> [61]	Functional MRI and SNPs	An accuracy of 86% (response)

SVM, support vector machine; SNPs, single nucleotide polymorphisms; MRI, magnetic resonance imaging; AUC, area under the receiver operating characteristic curve; MFNN, multi-layer feedforward neural network.

Decision Trees

The decision tree model, a traditional machine learning approach, builds up tree structures using top-down processing and prune the tree structures based on the concept of information entropy [31].

For example, Maciukiewicz *et al.* [32] used a classification-regression tree model, an example of decision tree models, to foresee antidepressant treatment response and to achieve an accuracy of 57% based on multi-omics data such as SNPs. In the first step, Maciukiewicz *et al.* [32] conducted a genome-wide association study (GWAS) to find genetic variants that were associated with antidepressant treatment response in a hypothesis-free manner. Next, they used the least absolute shrinkage and selection operator (LASSO) regression model [33], the regularization and variable selection scheme, to identify key predictive features such as the rs2036270 SNP in the *RARB* gene and the rs7037011 SNP near the *LOC105375971* gene [32]. In other word, the LASSO model serves as a feature selection algorithm to provide a small subset of good features for treatment prediction.

Support Vector Machines

The support vector machine (SVM) models [34] apply kernel functions to transform training datasets into higher dimensional data spaces and then determine a linear sep-

arating hyperplane based on the maximal margin [35].

For example, Maciukiewicz *et al.* [32] also used an SVM model, a traditional machine learning approach, to foresee antidepressant treatment response and to achieve an accuracy of 64% based on multi-omics data such as SNPs.

Random Forests

The random forest models construct a collection of many decision tree models using a random subset of training data and then produce the prediction by committee among the individual decision trees [36]. Namely, the random forest models, which consist of multiple decision trees to perform prediction tasks, are an example of ensemble learning approaches.

For instance, Kautzky *et al.* [37] employed a random forest model, a traditional machine learning approach, to correctly predict 25% of responders for antidepressant treatment outcome based on multi-omics data such as SNPs and clinical datasets. Remarkably, Kautzky *et al.* [37] found several key predictive features including a clinical variable known as melancholia and three SNPs such as the rs6265 SNP in the *BDNF* gene, the rs6313 SNP in the *HTR2A* gene, and the rs7430 SNP in the *PPP3CC* gene.

In addition, Athreya *et al.* [38] used a random forest

model to forecast antidepressant therapy response and to achieve an accuracy of $> 69\%$ (with the area under the receiver operating characteristic curve [AUC] value of > 0.7) based on multi-omics data such as SNPs and clinical datasets such as total baseline depression severity. In particular, Athreya *et al.* [38] identified the top six SNPs in the related GWAS analysis as key predictive features, including the rs5743467, rs2741130, and rs2702877 SNPs in the *DEFB1* gene, the rs696692 SNP in the *ERICH3* gene, the rs17137566 SNP in the *AHR* gene, and the rs10516436 SNP in the *TSPAN5* gene.

Ensemble Models

The boosting ensemble algorithm is commonly used to solve complex prediction applications in medicine and may possess numerous advantages such as greater prediction, better consistency, robust generalization, and the prevention of overfitting [39,40]. The LogitBoost algorithm is an example of the boosting ensemble algorithm, which organizes the performance of many base predictive algorithms (also referred to as weak predictive algorithms) to construct a robust prediction algorithm with higher accuracy [41]. Furthermore, the LogitBoost algorithm utilizes a binomial log-likelihood scheme that adjusts the predictive error linearly, thereby tending to be robust in handling outliers and noisy data [41]. For example, a base predictive algorithm can be a multi-layer feedforward neural network (MFNN), which may consist of one input layer, one hidden layer, and one output layer.

For example, Lin *et al.* [42] used the LogitBoost model, a boosting ensemble algorithm, to predict antidepressant treatment response and remission by using multi-omics data such as SNPs and clinical datasets in MDD patients in the Taiwanese population. In the first step, the wrapper-based feature selection algorithm [43] was performed to identify a subset of key predictive features from 10 genetic variants and 6 clinical variables, where the feature selection algorithm acts as a wrapper around the predictive algorithm. Particularly, Lin *et al.* [42] identified 15 biomarkers out of 16 biomarkers and eliminated one clinical variable, namely marital status, using the wrapper-based feature selection algorithm. Moreover, Lin *et al.* [42] compared the performance of the LogitBoost model to widely-used machine learning models, including MFNNs, logistic regression, SVM, C4.5 decision tree, naive Bayes, and random forests. It is indicated that the

LogitBoost model with the wrapper-based feature selection algorithm (using fewer biomarkers) led to equal performance to MFNNs and logistic regression [42]. In addition, the proposed model had better performance than SVM, C4.5 decision tree, naive Bayes, and random forests [42].

In brief, the best-first search is conducted by the wrapper-based method to identify a good subset of biomarkers by utilizing the predictive algorithm itself as part of the process for evaluating biomarker subsets [31,43]. The best-first search starts with an empty set of biomarkers and then searches forward to select a likely subset of biomarkers by a greedy hill-climbing technique extended with a backtracking algorithm [31,43].

Logistic regression models often serve as a basis for comparison when benchmarking is performed for predictive modeling applications. The C4.5 decision tree models build up tree structures using top-down processing and prune the tree structures based on the concept of information entropy [31]. The naive Bayes model estimates the probability that a particular sample appears to be in a specific domain (for instance, “non-response” or “response”) by using the Bayes’ theorem [41].

Deep Learning

In this subsection, we focus on antidepressant treatment outcome using deep learning strategies. The usage of deep learning methods is still in its infancy in terms of forecasting antidepressant treatments in MDD because scant human studies have been conducted to explore the deep learning models.

In general, the deep learning models are comprised of MFNNs with more than one hidden layer, where an MFNN model consists of one input layer, one or multiple hidden layers and one output layer [16,17]. That is, deep learning models generally leverage an MFNN model with multiple layers, instead of an MFNN model with only one single layer. The MFNN model normally utilize the back-propagation algorithm [44] in the learning stage. The merit of the deep learning framework is that the MFNN models have the strengths of nonlinear representations, integrated structures, fault tolerance, and real-time processing [45].

To predict antidepressant treatment response and remission in MDD, Lin *et al.* [45] leveraged a deep learning framework by using multi-omics data including genetic

datasets such as SNPs, demographic datasets such as marital status, age, and sex, as well as clinical datasets such as suicide attempt status, baseline Hamilton Rating Scale for Depression score, and depressive episodes for MDD patients. In the first step, Lin *et al.* [45] conducted a GWAS to determine the key predictive SNPs which were significantly associated with antidepressant treatment response and remission in a hypothesis-free approach. Secondly, MFNNs were employed to estimate the possible sophisticated interaction between predictive variables and antidepressant treatment response and remission [45]. Remarkably, Lin *et al.* [45] reported an MFNN model with 2 hidden layers for antidepressant treatment response with the AUC value of 0.82 (sensitivity = 0.75; specificity = 0.69). It is also revealed that another MFNN model with 3 hidden layers for antidepressant remission with the AUC value of 0.81 (sensitivity = 0.77; specificity = 0.66) [45].

INTEGRATION OF RESEARCH ON PHARMACOGENOMICS AND NEUROIMAGING FOR ANTIDEPRESSANT TREATMENT OUTCOME IN PATIENTS WITH MDD

The integrated approach of pharmacogenomics and neuroimaging is still in its infancy in terms of forecasting antidepressant treatments in MDD due to the fact that scant human studies have explored the integrated approach of evaluating drug treatment response. Here, we focus on antidepressant treatment outcomes using integrated strategies of pharmacogenomics and neuroimaging in this section (Fig. 2).

We first conducted a comprehensive search of the elec-

tronic PubMed database (2015–present) using key words such as “machine learning,” “deep learning,” “antidepressant,” “major depressive disorder,” “neuroimaging,” and “pharmacogenomics”. Then, we manually screened the obtained articles with a particular focus on MDD.

Neuroimaging for Antidepressant Treatment Outcome

Predicting antidepressant treatment outcome with neuroimaging data has been a focus of attention for at least two decades, from employing simple statistical approaches to using machine learning and deep learning methods [24,25]. There are a wide variety of neuroimaging modalities for examining predictive variables of antidepressant treatment response, including structural magnetic resonance imaging (MRI), diffusion tensor imaging, electroencephalography, functional MRI, positron emission tomography, single-photon emission computed tomography, near-infrared spectroscopy, and proton magnetic resonance spectroscopy [46,47]. By employing these imaging modalities, several predictive variables have been identified to be associated with antidepressant treatment outcome, such as the prefrontal cortex, anterior cingulate cortex, hippocampus, amygdala, and insula [46].

In brief, several machine learning methods have been utilized to evaluate predictive variables for forecasting antidepressant treatment outcome with neuroimaging data [48]. These machine learning methods include linear regression, SVM, random forests, and alternating decision tree models [48]. For instance, Patel *et al.* [49] suggested an alternating decision tree model, a traditional machine learning approach, to forecast antidepressant treatment response and to achieve an accuracy of 89% based on structural MRI data (such as whole brain atrophy and

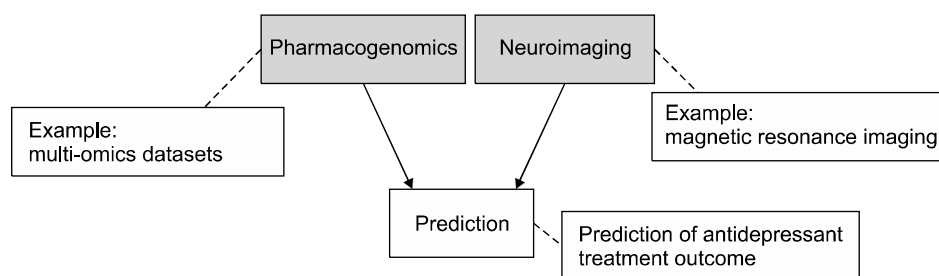


Fig. 2. An example of the integrated approach of pharmacogenomics and neuroimaging. The idea of integrating pharmacogenomics and neuroimaging is designated as neuroimaging pharmacogenomics or neuroimaging genomics. The predictive models in integrated approach can be traditional machine learning algorithms as well as deep learning algorithms.

global white matter hyper-intensity burden) and clinical datasets (such as mini-mental status examination scores and age). The alternating decision tree model, which is a generalized form of decision trees, is compatible to a boosting scheme for scaling down variance and bias principally [49].

In addition, the reader can refer to a recent review by Fonseka *et al.* [46] for predictive variables on the neuroimaging applications of antidepressant treatments in MDD and a recent review by Kang and Cho [47] for neuroimaging modalities. The reader can also refer to a recent review by Janssen *et al.* [48] for machine learning methods that were shown to be able to predict antidepressant treatment outcome with neuroimaging data.

Integration of Pharmacogenomics and Neuroimaging

The original idea of integrating pharmacogenomics and neuroimaging has a long history dating back to at least 15 years ago with an aim to use the concept of intermediate endophenotypes by combining the merits of both neuroimaging and pharmacogenomics studies [50,51]. Previously, this ensemble approach is known as imaging genetics or the interchangeable term, namely imaging genomics. Nowadays, it is also known as neuroimaging genomics (or neuroimaging pharmacogenomics), which depicts the intersection of the two fields, neuroimaging and genomics (or neuroimaging and pharmacogenomics) [22]. A notable example of neuroimaging genomics is the Enhancing Neuroimaging Genetics through Meta-Analysis Consortium, which aims to understand how psychiatric disorders affect the brain by integrating neuroimaging and genomic data [52]. Another example of neuroimaging pharmacogenomics is the Canadian Biomarker Integration Network in Depression project, which aims to better forecast antidepressant treatment outcome by integrating neuroimaging data and multi-omics data such as SNPs and clinical datasets [53].

In psychiatry, the concept of intermediate endophenotypes was proposed to provide quantitative measures for behavioral phenotypes; thereby potentially creating a connection between genes and behavioral phenotypes [54,55]. In other words, the intermediate endophenotypes are the quantitative biomarkers of brain activities acquired by neuroimaging and can be used to evaluate neurobiological changes of brain functions affected by psychiatric disorders and/or treatments [56]. It is indi-

cated that the intermediate endophenotypes of neuroimaging can be leveraged to overcome the disadvantage of small effect sizes in pharmacogenomics studies [51]. Moreover, MDD-related genomic variants may manifest in the intermediate endophenotypes, which can be assessed using neuroimaging and might be more closely associated with MDD and/or antidepressant treatment outcome [57]. On the other hand, it is concerned that there are few replicated findings in the previous studies of neuroimaging genomics, which may be due to limited sample sizes, unreliable study designs, and a lack of corrections for multiple testing [58,59].

In the medical literature, such studies in the joint investigation of pharmacogenomics and neuroimaging (that is, neuroimaging pharmacogenomics or neuroimaging genomics; Fig. 2) are still rare, especially for antidepressant treatment outcome in patients with MDD. While this review does not propose to address all related studies in an exhaustive manner, it nonetheless is representative of the overall trend for current research on neuroimaging pharmacogenomics for antidepressant treatment outcome in patients with MDD.

For example, in order to identify the most effective antidepressant for an individual patient, Chang *et al.* [60] implemented an Antidepressant Response Prediction Network (ARPNNet) framework which utilized an elastic net model to pinpoint key predictive features from patient data including MRI data and multi-omics data (such as SNPs, DNA methylation, and demographic information). In other word, the elastic net model serves as a feature selection algorithm in the ARPNNet framework to generate a small subset of good features for subsequent analysis. Furthermore, the ARPNNet framework employed a linear regression model, a traditional machine learning model, and was able to achieve an accuracy of 84% by using the selected key predictive features from the patient records [60]. It is indicated that the ARPNNet model was able to assist doctors prescribe the most competent antidepressant for their patients by inferencing the similar features among patients and by recognizing the comparable antidepressant treatment records between patients [60].

Moreover, Pei *et al.* [61] reported that a neuroimaging pharmacogenomics approach, a hybrid method combining functional MRI data and multi-omics data such as SNPs, was able to predict early-stage antidepressant treatment response (namely the first 2 weeks). First, the key

predictive variables were determined by the SVM-recursive feature elimination model which serves as a feature selection algorithm [61]. The key predictive variables included 3 SNPs (including *HTR2C* rs1801412, *DRD5* rs1967550, and *TOR1A* rs3842225) and six regions of interest in functional MRI data (including left orbital part superior frontal gyrus, left hippocampus, right amygdala, right hippocampus, right posterior cingulate gyrus, and left anterior cingulate and paracingulate gyri) [61]. Next, an SVM algorithm was used to achieve the early-stage antidepressant treatment prediction with an accuracy of 86% [61].

LIMITATIONS

For the various findings as represented in the aforementioned sections, several shortcomings of these research studies should be taken into account in the interdisciplinary fields of pharmacogenomics, neuroimaging, artificial intelligence, machine learning, deep learning, and antidepressant treatments in MDD. It should be noted that the following limitations may be attributed to not only the pharmacogenomics of antidepressant treatments in MDD but also any pharmacogenomics studies.

One main disadvantage of these past studies is that there were no well-defined conclusions because common benchmark datasets may not be previously available to precisely carry out well-thought-out comparisons among various machine learning and deep learning frameworks in the previous findings [62]. Because universal benchmark frameworks have contributed to recent remarkable progresses in the field of computer vision, it is believed that the field of pharmacogenomics can also take advantages of universal benchmark frameworks [63]. Therefore, standardized benchmark frameworks are crucially needed for the pharmacogenomics of antidepressant treatments in MDD.

Moreover, it should be emphasized that we should utilize the conventional linear regression models as an essential benchmark when employing deep learning methods such as the deep MFNN frameworks [64]. In other words, deep learning methods such as the deep MFNN frameworks should be a compatible method not only to the conventional linear regression models but also to conventional non-linear models such as SVMs and random forests [63].

Furthermore, a typical pitfall is that the aforementioned machine learning and deep learning methods may not utilize the cross-validation approach to prevent the possibility of overfitting during the training procedure. For example, the repeated 10-fold cross-validation method and leave-one-out cross-validation method could be valuable approaches for exploring the generalization of machine learning and deep learning methods [41,65]. Concisely, the repeated 10-fold cross-validation method randomly divides the full data into ten subgroups, and then the machine learning and deep learning model can be trained by nine-tenths of the dataset and tested by the unused tenth of the dataset [31]. Then, the previous procedure is repeatedly performed nine more times by designating distinct nine-tenths of the dataset for training and a distinct tenth of the dataset for testing. In a similar way, the leave-one-out cross-validation method is an extreme case of the repeated 10-fold cross-validation method, where the number of folds equals the number of cases in the full data [66]. In addition, the leave-one-out cross-validation method is often leveraged when the number of cases in the full data (or in a particular subset) is limited [66]. Interestingly, it is hypothesized that the cross-validation approach might supposedly affect the long-term architecture and performance in the consequent machine learning and deep learning methods [67].

On another note, the typical concerns in deep learning, namely deep neural networks or deep MFNNs, may be included in the aforementioned studies [63]. For example, the accuracy of the deep learning frameworks would be relatively flat when the cohort size is low [68]. Moreover, it is pivotal that a number of independent experiments would have generalized their findings by using universal benchmark datasets [69]. However, it is an open challenge that universal benchmark datasets may be unreachable to expedite successive analysis in deep learning experiments. Therefore, future deep learning experiments such as pharmacogenomics studies of antidepressant treatments in MDD should be accurately reproducible based on universally well-established benchmark datasets, which should be achieved by the research community in the multidisciplinary fields of pharmacogenomics, neuroimaging, artificial intelligence, machine learning, deep learning, and antidepressant treatments in MDD.

Yet another pitfall is that it is especially troublesome to evaluate the interpretation of deep learning methods such

as the deep MFNN frameworks. In principal, deep learning methods are recognized as a “black box”, which is typically problematic to interpret [70]. It is also worth mentioning that not only the deep MFNN frameworks but also other deep learning methods (such as convolutional neural network models and recurrent neural network models) have this problem [71]. It is indisputable that future interpretable deep learning methods are warranted; thereby, we would identify understandable features elicited from the deep learning architectures.

One major open challenge and emerging problem in the pharmacogenomics of antidepressant treatments in MDD using the machine learning and deep learning methods is that open-source software tools are essentially warranted due to their relative importance in replicability and reusability [72]. Second, standardized benchmarking metrics in the pharmacogenomics of antidepressant treatments in MDD using the machine learning and deep learning methods could be very challenging. It should be emphasized that the aforementioned challenges and emerging problems in the pharmacogenomics of antidepressant treatments in MDD using the machine learning and deep learning methods may also stem from any pharmacogenomics studies.

There are major open challenges and emerging problems in the machine learning and deep learning strategy itself. The first open challenge and emerging problem in the machine learning and deep learning strategy is to resolve the high-dimensionality problem, especially on the task of dimensionality reduction for the extremely unmanageable large volume of the real-world data [73]. For example, the application of new deep learning algorithms, such as the generative adversarial network (GAN) architecture [74], has been contributing to fulfill the task of dimensionality reduction for the large-scale single-cell sequencing data in the preclinical stage of the drug development pipeline [75]. In addition, the reader can refer to a recent review by Lin *et al.* [75] for the relevant applications of GAN models in drug design and discovery. The second open challenge and emerging problem in machine learning and deep learning strategy is to overcome the problem of unstructured data or un-categorized data, where the real-world data is prone to have a small quantity of structured data or categorized data [73].

CONCLUSION AND PERSPECTIVES

As indicated by the aforementioned findings, pharmacogenomics affirms to provide novel therapeutic and diagnostic approaches for antidepressant treatment prediction in MDD using the multidisciplinary approach of artificial intelligence, machine learning, deep learning, and neuroimaging. First, in terms of artificial intelligence techniques, it is of great interest that future prospective research projects should explore machine learning and deep learning strategy such as deep MFNNs and GAN models to forecast antidepressant treatment response and remission with the desired predictive variables of multi-omics data, which may further advance feasible medical solutions in global health as well as public health. Second, in terms of research on the integration of pharmacogenomics and neuroimaging, it is vital to note that machine learning and deep learning techniques might play a key role in predicting antidepressant treatment response and remission in MDD. Furthermore, machine learning and deep learning strategy such as deep MFNNs and GAN models will be undoubtedly established towards the field of pharmacogenomics in light of the compelling needs of innovative approaches in the fields of public health, global health, and population health [9]. In addition, we would envision that the recent advancements in single-cell sequencing technologies and data-intensive health sciences may assuredly trigger novel machine learning and deep learning software tools for public health, global health, and population health over the next decade [1,76]. Thereby, the general public and governments should achieve these challenges and issues with high priorities in the up-coming years [7,77]. In the next generation to come, pharmacogenomics involving machine learning, deep learning, and neuroimaging approaches would eventually become a reality in individualized clinical care when future extensive studies are able to fully investigate the relevant predictive variables as well as clinical biomarkers [78,79].

■ Acknowledgments

This work was supported by National Health Research Institutes, Taiwan (NHRI-EX109-10731NI), Ministry of Science and Technology in Taiwan (MOST 109-2622-B-039-001-CC2; 109-2314-B-039-001; 109-2314-B-039-039-MY3), Taiwan Ministry of Health and Welfare Clinical

Trial and Research Center of Excellence (MOHW109-TDU-B-212-114004), China Medical University and Hospital (DMR-109-MOST-03). The funding sources had no involvement in the article.

■ Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

■ Author Contributions

Conceptualization: Eugene Lin, Chieh-Hsin Lin, Hsien-Yuan Lane. Writing—original draft: Eugene Lin. Writing—review and editing: Eugene Lin, Chieh-Hsin Lin, Hsien-Yuan Lane.

■ ORCID

Eugene Lin <https://orcid.org/0000-0001-6984-2159>
Chieh-Hsin Lin <https://orcid.org/0000-0001-6949-8968>
Hsien-Yuan Lane <https://orcid.org/0000-0003-2162-8174>

REFERENCES

- Lin E, Lin CH, Lane HY. *Precision psychiatry applications with pharmacogenomics: artificial intelligence and machine learning approaches*. *Int J Mol Sci* 2020;21:969.
- Klein ME, Parvez MM, Shin JG. *Clinical implementation of pharmacogenomics for personalized precision medicine: barriers and solutions*. *J Pharm Sci* 2017;106:2368-2379.
- Torres EB, Isenhour RW, Nguyen J, Whyatt C, Nurnberger JL, Jose JV, et al. *Toward precision psychiatry: statistical platform for the personalized characterization of natural behaviors*. *Front Neurol* 2016;7:8.
- Amare AT, Schubert KO, Baune BT. *Pharmacogenomics in the treatment of mood disorders: strategies and opportunities for personalized psychiatry*. *EPMA J* 2017;8:211-227.
- Lin E, Tsai SJ. *Genome-wide microarray analysis of gene expression profiling in major depression and antidepressant therapy*. *Prog Neuropsychopharmacol Biol Psychiatry* 2016;64:334-340.
- Lin E, Lane HY. *Genome-wide association studies in pharmacogenomics of antidepressants*. *Pharmacogenomics* 2015;16:555-566.
- Lin E, Tsai SJ. *Multi-omics and machine learning applications in precision medicine*. *Curr Pharmacogen Personal Med* 2017;15:97-104.
- Bzdok D, Meyer-Lindenberg A. *Machine learning for precision psychiatry: opportunities and challenges*. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2018;3:223-230.
- Lin E, Lane HY. *Machine learning and systems genomics approaches for multi-omics data*. *Biomark Res* 2017;5:2.
- Athreya AP, Iyer R, Wang L, Weinsilboum RM, Bobo WV. *Integration of machine learning and pharmacogenomic biomarkers for predicting response to antidepressant treatment: can computational intelligence be used to augment clinical assessments?* *Pharmacogenomics* 2019;20:983-988.
- Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. *A survey on deep learning in medical image analysis*. *Med Image Anal* 2017;42:60-88.
- Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. *Opportunities and obstacles for deep learning in biology and medicine*. *J R Soc Interface* 2018;15:20170387.
- Dana D, Gadhiya SV, St Surin LG, Li D, Naaz F, Ali Q, et al. *Deep learning in drug discovery and medicine; scratching the surface*. *Molecules* 2018;23:2384.
- Grapov D, Fahrman J, Wanichthanarak K, Khoomrung S. *Rise of deep learning for genomic, proteomic, and metabolomic data integration in precision medicine*. *OMICS* 2018;22:630-636.
- Zhang S, Bamakan SMH, Qu Q, Li S. *Learning for personalized medicine: a comprehensive review from a deep learning perspective*. *IEEE Rev Biomed Eng* 2019;12:194-208.
- LeCun Y, Bengio Y, Hinton G. *Deep learning*. *Nature* 2015;521:436-444.
- Hinton G. *Deep learning—a technology with the potential to transform health care*. *JAMA* 2018;320:1101-1102.
- Hulshoff Pol H, Bullmore E. *Neural networks in psychiatry*. *Eur Neuropsychopharmacol* 2013;23:1-6.
- Ekins S. *The next era: deep learning in pharmaceutical research*. *Pharm Res* 2016;33:2594-2603.
- Mamoshina P, Vieira A, Putin E, Zhavoronkov A. *Applications of deep learning in biomedicine*. *Mol Pharm* 2016;13:1445-1454.
- Falcone M, Smith RM, Chenoweth MJ, Bhattacharjee AK, Kelsoe JR, Tyndale RF, et al. *Neuroimaging in psychiatric pharmacogenetics research: the promise and pitfalls*. *Neuropsychopharmacology* 2013;38:2327-2337.
- Mufford MS, Stein DJ, Dalvie S, Groenewold NA, Thompson PM, Jahanshad N. *Neuroimaging genomics in psychiatry—a translational approach*. *Genome Med* 2017;9:102.
- Davatzikos C. *Machine learning in neuroimaging: progress and challenges*. *Neuroimage* 2019;197:652-656.
- Godlewska BR. *Neuroimaging as a tool for individualized treatment choice in depression: the past, the present and the future*. *Curr Behav Neurosci Rep* 2020;7:32-39.
- Patel MJ, Khalaf A, Aizenstein HJ. *Studying depression using imaging and machine learning methods*. *Neuroimage Clin* 2015;10:115-123.
- Iniesta R, Stahl D, McGuffin P. *Machine learning, statistical learning and the future of biological research in psychiatry*. *Psychol Med* 2016;46:2455-2465.
- Lane HY, Tsai GE, Lin E. *Assessing gene-gene interactions in pharmacogenomics*. *Mol Diagn Ther* 2012;16:15-27.
- Lin E, Hwang Y, Liang KH, Chen EY. *Pattern-recognition tech-*

- niques with haplotype analysis in pharmacogenomics. *Pharmacogenomics* 2007;8:75-83.
29. Dwyer DB, Falkai P, Koutsouleris N. *Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol* 2018;14:91-118.
 30. Perlman K, Benrimoh D, Israel S, Rollins C, Brown E, Tunteng JF, et al. *A systematic meta-review of predictors of antidepressant treatment outcome in major depressive disorder. J Affect Disord* 2019;243:503-515.
 31. Huang LC, Hsu SY, Lin E. *A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data. J Transl Med* 2009;7:81.
 32. Maciukiewicz M, Marshe VS, Hauschild AC, Foster JA, Rotzinger S, Kennedy JL, et al. *GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. J Psychiatr Res* 2018;99:62-68.
 33. Tibshirani R. *The lasso method for variable selection in the Cox model. Stat Med* 1997;16:385-395.
 34. Vapnik V. *The nature of statistical learning theory. Berlin, Heidelberg:Springer Science & Business Media;2013.*
 35. Lin E, Hwang Y. *A support vector machine approach to assess drug efficacy of interferon-alpha and ribavirin combination therapy. Mol Diagn Ther* 2008;12:219-223.
 36. Breiman L. *Random forests. Mach Learn* 2001;45:5-32.
 37. Kautzky A, Baldinger P, Souery D, Montgomery S, Mendlewicz J, Zohar J, et al. *The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. Eur Neuropsychopharmacol* 2015;25:441-453.
 38. Athreya AP, Neavin D, Carrillo-Roa T, Skime M, Biernacka J, Frye MA, et al. *Pharmacogenomics-driven prediction of antidepressant treatment outcomes: a machine-learning approach with multi-trial replication. Clin Pharmacol Ther* 2019;106:855-865.
 39. Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. *A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern Part C* 2011;42:463-484.
 40. Yang P, Hwa Yang Y, Zhou BB, Zomaya AY. *A review of ensemble methods in bioinformatics. Curr Bioinform* 2010;5:296-308.
 41. Lin E, Lin CH, Hung CC, Lane HY. *An ensemble approach to predict schizophrenia using protein data in the N-methyl-D-aspartate receptor (NMDAR) and tryptophan catabolic pathways. Front Bioeng Biotechnol* 2020;8:569.
 42. Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. *Prediction of antidepressant treatment response and remission using an ensemble machine learning framework. Pharmaceuticals (Basel)* 2020;13:305.
 43. Kohavi R, John GH. *Wrappers for feature subset selection. Artif Intell* 1997;97:273-324.
 44. Rumelhart DE, Hinton GE, Williams RJ. *Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, editors. Parallel distributed processing: explorations in the microstructure of cognition. Vol. 1: foundations. Cambridge: MIT Press;1996.*
 45. Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. *A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. Front Psychiatry* 2018;9:290.
 46. Fonseka TM, MacQueen GM, Kennedy SH. *Neuroimaging biomarkers as predictors of treatment outcome in Major Depressive Disorder. J Affect Disord* 2018;233:21-35.
 47. Kang SG, Cho SE. *Neuroimaging biomarkers for predicting treatment response and recurrence of major depressive disorder. Int J Mol Sci* 2020;21:2148.
 48. Janssen RJ, Mourão-Miranda J, Schnack HG. *Making individual prognoses in psychiatry using neuroimaging and machine learning. Biol Psychiatry Cogn Neurosci Neuroimaging* 2018;3:798-808.
 49. Patel MJ, Andreescu C, Price JC, Edelman KL, Reynolds CF 3rd, Aizenstein HJ. *Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. Int J Geriatr Psychiatry* 2015;30:1056-1067.
 50. Hariri AR, Drabant EM, Weinberger DR. *Imaging genetics: perspectives from studies of genetically driven variation in serotonin function and corticolimbic affective processing. Biol Psychiatry* 2006;59:888-897.
 51. Rabl U, Scharinger C, Müller M, Pezawas L. *Imaging genetics: implications for research on variable antidepressant drug response. Expert Rev Clin Pharmacol* 2010;3:471-489.
 52. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, et al. *The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav* 2014;8:153-182.
 53. Lam RW, Milev R, Rotzinger S, Andreazza AC, Blier P, Brenner C, et al. *Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. BMC Psychiatry* 2016;16:105.
 54. Gottesman II, Gould TD. *The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry* 2003;160:636-645.
 55. Meyer-Lindenberg A, Weinberger DR. *Intermediate phenotypes and genetic mechanisms of psychiatric disorders. Nat Rev Neurosci* 2006;7:818-827.
 56. Viviani R, Lehmann ML, Stingl JC. *Use of magnetic resonance imaging in pharmacogenomics. Br J Clin Pharmacol* 2014;77:684-694.
 57. Buch AM, Liston C. *Dissecting diagnostic heterogeneity in depression by integrating neuroimaging and genetics. Neuropsychopharmacology* 2021;46:156-175.
 58. Carter CS, Bearden CE, Bullmore ET, Geschwind DH, Glahn DC, Gur RE, et al. *Enhancing the informativeness and replicability of imaging genomics studies. Biol Psychiatry* 2017;

- 82:157-164.
59. Pereira LP, Köhler CA, Stubbs B, Miskowiak KW, Morris G, de Freitas BP, et al. *Imaging genetics paradigms in depression research: systematic review and meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry* 2018;86:102-113.
 60. Chang B, Choi Y, Jeon M, Lee J, Han KM, Kim A, et al. *ARNNet: antidepressant response prediction network for major depressive disorder. Genes (Basel)* 2019;10:907.
 61. Pei C, Sun Y, Zhu J, Wang X, Zhang Y, Zhang S, et al. *Ensemble learning for early-response prediction of antidepressant treatment in major depressive disorder. J Magn Reson Imaging* 2020;52:161-171.
 62. Rutledge RB, Chekroud AM, Huys QJ. *Machine learning and big data in psychiatry: toward clinical applications. Curr Opin Neurobiol* 2019;55:152-159.
 63. Lin E, Lin CH, Lane HY. *Relevant applications of generative adversarial networks in drug design and discovery: molecular de novo design, dimensionality reduction, and de novo peptide and protein design. Molecules* 2020;25:3250.
 64. Kalinin AA, Higgins GA, Reamaroon N, Soroushmehr S, Allyn-Feuer A, Dinov ID, et al. *Deep learning in pharmacogenomics: from gene regulation to patient stratification. Pharmacogenomics* 2018;19:629-650.
 65. Jiao Y, Du P. *Performance measures in evaluating machine learning based bioinformatics predictors for classifications. Quant Biol* 2016;4:320-330.
 66. Wong TT. *Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit* 2015;48:2839-2846.
 67. Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A. *A primer on deep learning in genomics. Nat Genet* 2019;51:12-18.
 68. Russell LE, Schwarz UI. *Variant discovery using next-generation sequencing and its future role in pharmacogenetics. Pharmacogenomics* 2020;21:471-486.
 69. Azodi CB, Bolger E, McCarren A, Roantree M, de Los Campos G, Shiu SH. *Benchmarking parametric and machine learning models for genomic prediction of complex traits. G3 (Bethesda)* 2019;9:3691-3702.
 70. Azodi CB, Tang J, Shiu SH. *Opening the black box: interpretable machine learning for geneticists. Trends Genet* 2020;36:442-455.
 71. Eraslan G, Avsec Ž, Gagneur J, Theis FJ. *Deep learning: new computational modelling techniques for genomics. Nat Rev Genet* 2019;20:389-403.
 72. Kalamara A, Tobalina L, Saez-Rodriguez J. *How to find the right drug for each patient? Advances and challenges in pharmacogenomics. Curr Opin Syst Biol* 2018;10:53-62.
 73. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. *Deep learning applications and challenges in big data analytics. J Big Data* 2015;2:1.
 74. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. *Generative adversarial nets. Montreal:University of Montreal;2014.*
 75. Lin E, Mukherjee S, Kannan S. *A deep adversarial variational autoencoder model for dimensionality reduction in single-cell RNA sequencing analysis. BMC Bioinformatics* 2020;21:64.
 76. Lin E, Lin CH, Lai YL, Huang CH, Huang YJ, Lane HY. *Combination of G72 genetic variation and G72 protein level to detect schizophrenia: machine learning approaches. Front Psychiatry* 2018;9:566.
 77. Angermueller C, Pärnamaa T, Parts L, Stegle O. *Deep learning for computational biology. Mol Syst Biol* 2016;12:878.
 78. Lin E, Tsai SJ. *Diagnostic prediction based on gene expression profiles and artificial neural networks. In: Purohit HJ, Kalia VC, More RP, editors. Soft computing for biological systems. Singapore:Springer;2018. p.13-22.*
 79. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep* 2016;6:26286.