

RESEARCH ARTICLE

Model-based optimization approaches for precision medicine: A case study in presynaptic dopamine overactivity

Kai-Cheng Hsu¹, Feng-Sheng Wang^{2*}

1 Department of Neurology, National Taiwan University Hospital Yunlin Branch, Yunlin, Taiwan,

2 Department of Chemical Engineering, National Chung Cheng University, Chiayi, Taiwan

* chmfsw@ccu.edu.tw



OPEN ACCESS

Citation: Hsu K-C, Wang F-S (2017) Model-based optimization approaches for precision medicine: A case study in presynaptic dopamine overactivity. PLoS ONE 12(6): e0179575. <https://doi.org/10.1371/journal.pone.0179575>

Editor: Giovanni Maga, Istituto di Genetica Molecolare, ITALY

Received: April 5, 2017

Accepted: May 31, 2017

Published: June 14, 2017

Copyright: © 2017 Hsu, Wang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was financially supported by the Ministry of Science of Technology, Taiwan, ROC (<https://www.most.gov.tw/>) through grants MOST103-2221-E-194-045-MY3, MOST104-2627-B-194-001, and MOST105-2627-M-194-001, grant recipient Prof. Feng-Sheng Wang. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Precision medicine considers an individual's unique physiological characteristics as strongly influential in disease vulnerability and in response to specific therapies. Predicting an individual's susceptibility to developing an illness, making an accurate diagnosis, maximizing therapeutic effects, and minimizing adverse effects for treatment are essential in precision medicine. We introduced model-based precision medicine optimization approaches, including pathogenesis, biomarker detection, and drug target discovery, for treating presynaptic dopamine overactivity. Three classes of one-hit and two-hit enzyme defects were detected as the causes of disease states by the optimization approach of pathogenesis. The cluster analysis and support vector machine was used to detect optimal biomarkers in order to discriminate the accurate etiology from three classes of disease states. Finally, the fuzzy decision-making method was employed to discover common and specific drug targets for each classified disease state. We observed that more accurate diagnoses achieved higher satisfaction grades and dosed fewer enzyme targets to treat the disease. Furthermore, satisfaction grades for common drugs were lower than for specific ones, but common drugs could simultaneously treat several disease states that had different etiologies.

Introduction

The concept of precision medicine has been recognized not only by clinicians and biomedical researchers but also by patients and policy makers in recent years [1]. Precision medicine is defined as treatments tailored to individual patients on the basis of their genetic, biomarker, phenotypic, and psychosocial characteristics that distinguish a given patient from other patients with similar clinical presentations [2]. The goals of precision medicine are to identify an individual's susceptibility to disease, obtain an accurate diagnosis, and deliver an efficient treatment. Although some successful examples of precision medicine are present in oncology, relatively few examples exist in psychiatry [3, 4].

Schizophrenia is a chronic psychiatric disorder that afflicts approximately 1% of the population worldwide, but its cause remains unknown [5]. Evidence of dopamine (DA) involvement

Competing interests: The authors have declared that no competing interests exist.

in schizophrenia developed with the approval of the antipsychotic drug Reserpine in 1954, which was reported to block the accumulation of DA into secretory vesicles [6]. Other first-, second-, and third-generation antipsychotic drugs have been found to correlate with their binding affinities for D2-type DA receptors [7–10]. Furthermore, several imaging studies using positron emission tomography have shown presynaptic dopaminergic dysfunction with dopamine overactivity in schizophrenia [11–14].

The current diagnosis of psychiatric disorders, including schizophrenia, is mainly based on a clinician's assessment according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and International Statistical Classification of Diseases and Related Health Problems (ICD-10) diagnostic criteria. However, different disorders often manifest similarly, and the overlap of symptoms results in diagnostic ambiguity [15]. Diagnostic biomarkers using imaging techniques [16] and molecular biology techniques [17] may increase the accuracy of diagnostic strategies and allow disease classifications to be more effectively targeted for personalized therapy. One molecular biomarker alone may not possess a strong enough statistical power to predict disease state, especially with complex diseases. Additionally, these diseases not only require more accurate diagnostic tools but also need to be viewed more thoroughly in terms of their dynamic behaviors. Therefore, a set of biomarkers identified through systems biology approaches can be used as a panel to characterize a disease [18].

In contrast to the trial and error method, mathematical modeling and optimization are emerging technologies for drug discovery for human metabolic disorders [19–23]. In our previous studies [24, 25], we have introduced a fuzzy multiobjective optimization process for formulating the enzyme target design problem to identify drug targets. Prevention and treatment strategies for diseases should consider individual variability in precision medicine [1]. Therefore, we introduced a model-based optimization approach in this study to identify the pathogenesis of presynaptic dopamine overactivity, not only to determine a one-hit enzyme defect and two-hit enzyme defects but also their corresponding dysregulated level. Furthermore, such disorder observations were used to identify the candidates of biomarkers in order to classify different etiologies of presynaptic dopamine overactivity. Finally, a fuzzy decision-making approach was introduced to discover drug targets for various disease cases; that is, common targets for several disease states because of their different etiologies and specific targets for each disease state because of their specific etiologies were considered.

Method

This study proposes the model-based optimization approach for precision medicine for treating a patient with presynaptic dopamine overactivity as shown in Fig 1. The procedures are consisted of three parts, namely detection of pathogenesis, biomarker identification, and drug target discovery, and are similar to SOAP notes [26], which document subjective (S), objective (O), assessment (A), and plan (P) details, are a highly structured format for documenting the progress of a patient during treatment. Such notes are part of the problem-oriented medical record approach most commonly used by physicians and other healthcare professionals. The proposed computations are generally developed in advance to provide as a medical aided systems. Each computational procedure is explained as following.

Detection of pathogenesis

Sensitivity analysis is a conventional method for examining which enzyme is the most sensitive deficient in a network to trigger pathogenesis [27]. However, sensitivity analysis only changes each enzyme activity in a small amount to evaluate variation of the predicted metabolite concentrations, and cannot detect disease levels caused by enzyme defects. The most effective

enzymes of pathogenesis can be detected through the model-based optimization problem as follows:

$$\min_{\mathbf{x}, \hat{\alpha}, \mathbf{u}, \mathbf{z}} \sum_{i \in \Omega^{DS}} (x_i - x_i^{disease})^2 + \sum_{j \in \Omega^{Enz}} (1 - \hat{\alpha}_j / \hat{\alpha}_j^{basal})^2 + \sum_{j \in \Omega^{EX}} (1 - u_j / u_j^{basal})^2$$

subject to

Material balance equations :

$$\left\{ \begin{array}{l} \sum_{j=1}^r N_{ij} v_j(\mathbf{x}, \boldsymbol{\alpha}) + \sum_{j=1}^m B_{ij} u_j = 0, i \in \Omega^{SP} \\ w_j = \hat{\alpha}_j + \sum_{k=1}^{n_r} g_{jk} y_k, j \in \Omega^{Rxn} \\ v_j = \exp(w_j) \\ x_i = \exp(y_i) \\ \text{Metabolite constraints :} \\ x_i^{LB} \leq x_i \leq x_i^{UB}, i \in \Omega^{SP} \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} \text{Target constraints :} \\ \text{up-regulation : } \begin{cases} \hat{\alpha}_z^{basal} \leq \hat{\alpha}_z \leq \hat{\alpha}_z^{UB}, z \in \Omega_z^{up} \\ u_z^{basal} \leq u_z \leq u_z^{UB} \end{cases} \\ \text{down-regulation : } \begin{cases} \hat{\alpha}_z^{LB} \leq \hat{\alpha}_z \leq \hat{\alpha}_z^{basal}, z \in \Omega_z^{down} \\ u_z^{LB} \leq u_z \leq u_z^{UB} \end{cases} \\ \hat{\alpha}_z = \hat{\alpha}_z^{basal}, z \notin \Omega_z = \Omega_z^{up} + \Omega_z^{down} \end{array} \right.$$

where $x_i, i = 1, \dots, n$ are the metabolite concentrations, $\hat{\alpha}_j, j = 1, \dots, r$ are the logarithmic enzyme activities, N_{ij} is the stoichiometric coefficient of metabolite i in reaction j , v_j is the reaction rate j , and B_{ij} is the connectivity coefficient of metabolite i in external control u_j . The formulation in Eq (1) is a mixed-integer nonlinear programming (MINLP) problem whereby a set of integer variables \mathbf{z} are first detected, and their corresponding up and down regulatory levels of enzyme activities and external controls are determined to minimize the objective function. The first term of the objective function is used to account for the metabolite concentrations of the disease state as close to their disease levels, $x_i^{disease}$ as possible, the second and third terms are used to determine the detected enzymes and external inputs with as little variation to its basal level, $\hat{\alpha}_j^{basal}$ and u_j^{basal} , as possible. The MINLP problem is an NP-hard problem and is difficult to solve. Generally, the problem may contain multiple dysregulated inputs to yield the identical disease situation. Nested hybrid differential evolution (NHDE) can be applied to solve the NP-hard problem; NHDE iteratively found a set of pathogenic enzymes (Fig 2, S1 File).

Biomarker identification

Cluster and statistical analysis. Hierarchical cluster analysis was used to classify a given dataset into homogeneous groups that differ from each other on the basis of given parameters [28]. This analysis was performed with a between-groups linkage, and the cluster distance was

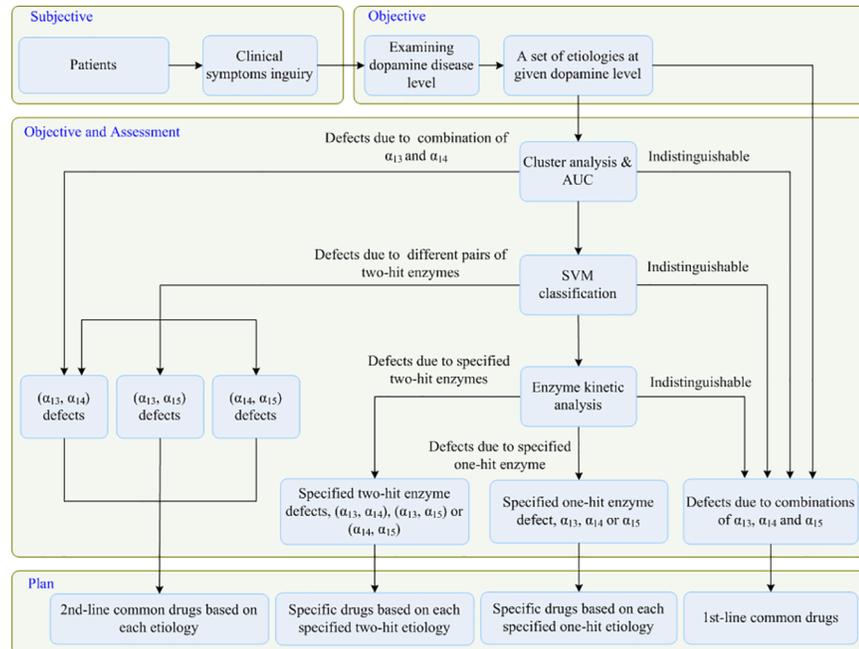


Fig 1. Roadmap of model-based optimization for precision medicine. Subjective component: The symptoms are inquired about according to DSM-IV and ICP-10 diagnostic criteria for the diagnosis of possible presynaptic dopamine overactivity. Objective component and assessment: Dopamine level is measured by image studies. Then, a single biomarker must be tested for identifying Class 1 patients with α_{13} and α_{14} defects. Furthermore, three to four biomarkers are examined for identifying Class 2 or Class 3 patients with α_{13} and α_{15} or α_{14} and α_{15} defects, respectively. Finally, if the enzyme activity can be tested, the specific disease states with precise single or multiple enzyme defects are obtained. Plan: When the dopamine level is determined, the 1st-line common drugs can be used without further information concerning enzyme defects. The 1st-line common drugs are compatible with current medications because tests for enzyme defects are not clinically available. After single or multiple biomarkers are examined to identify Class 1, 2, and 3 patients, 2nd-line common drugs are utilized. The 2nd-line common drugs are more effective than are the 1st-line common drugs. When specific enzyme defects are detected, the specific drugs can be discovered to treat patients with the highest therapeutic effect, the lowest adverse effect, and the lowest drug dose. Hence, the specific drugs are indicated for use in patients who are refractory to 1st- and 2nd-line common drugs.

<https://doi.org/10.1371/journal.pone.0179575.g001>

expressed as the square Euclidean distance [29]. A set of etiologies obtained from the pathogenesis problem (1) were clustered using a dendrogram employing the average linkage method into two groups. Between-group comparisons were made using independent-samples T tests and ANOVA tests. Receiver operating characteristic (ROC) curves were plotted and areas under curves (AUCs) were calculated. $P < 0.05$ was considered significant.

Support vector machine. Support vector machine (SVM) was applied to identify biomarkers. SVM was first introduced in 1995 [30] and has been used in many pattern recognition problems [31–33] and biological problems [33–36]. SVM determines the optimal hyperplane that has the maximum margin to classify data into two classes [34, 37–39]. The data for a two-class learning problem consists of objects labeled with one of two labels corresponding to two classes; for convenience we assume the labels y are +1 (positive patterns) or -1 (negative patterns). Let \mathbf{x} denote a vector of n -dimensional components x_i , $i = 1, \dots, n$. The object \mathbf{x}_i is called a pattern in the dataset $\{(\mathbf{x}_i, y_i), i = 1, \dots, m\}$. A linear classifier is based on a linear discriminant function of the form

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b \tag{2}$$

where the vector \mathbf{w} is the weight vector and the scalar b is the bias. Eq (2) is a hyperplane, that is,

a linear decision boundary, to linearly separate the dataset into positive or negative patterns. For a given hyperplane, we can determine the maximum margin classifier that maximizes the geometric margin, which is equivalent to minimizing the soft-margin SVM problem of the form

$$\begin{cases} \min_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{m_t} \xi_i \\ \text{subject to} \\ y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, i = 1, \dots, m_t \\ \xi_i \geq 0 \end{cases} \quad (3)$$

where m_t is the number of training sets. Once the optimal values of \mathbf{w} , b and ξ_i are found, the patterns in the test set can be classified by the discriminant function $f(\mathbf{x})$ either into the +1 or -1 class. The classifier of the problem (3) can be usually obtained by solving its dual problem.

SVM was originally designed for binary classification. To solve a multiclass problem, we decompose the multiclass problem into a series of two-class problems and construct several binary classifiers [30, 37–41]. The one-versus-all method, which was the earliest and one of the most widely used methods, was implemented. This method involves a parallel architecture comprising M-SVMs [41]. Each SVM solves a two-class problem defined by one information class. The “winner-takes-all” rule is used for the final decision, that is, the winning class is the one corresponding to the SVM with the highest output.

Triple SVM classifiers as shown in Fig 3 were applied to classify the three classes of data. First, 50% of the data from Classes 1 and 2 were randomly selected for the training of SVM1. Second,

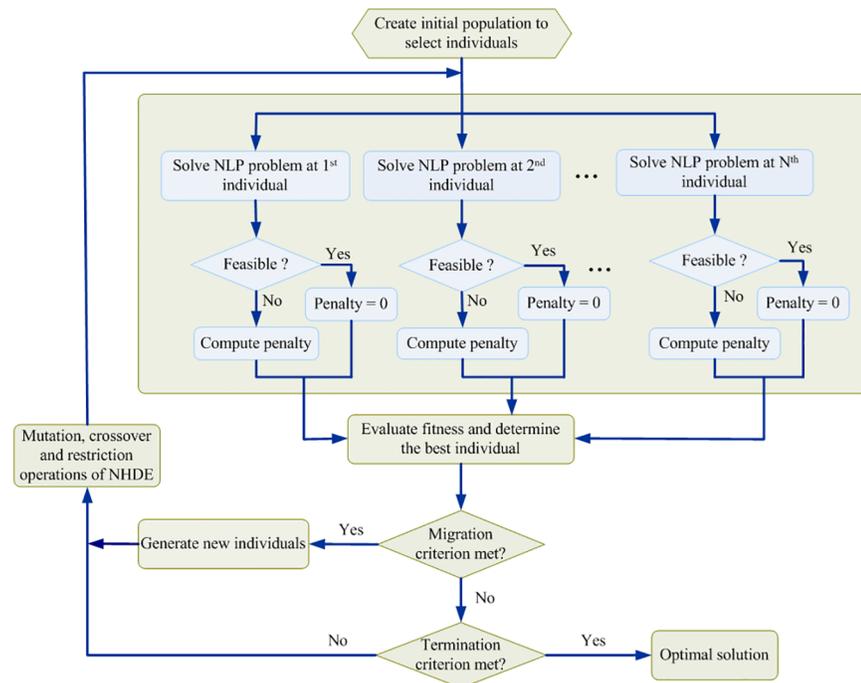


Fig 2. Flowchart for the modified algorithm for NHDE. The core procedure of the NHDE algorithm is the evaluation and selection operation as shown in the second and third block diagram of the flowchart. The evaluation step is to solve each nonlinear programming (NLP) problem produced from the maximizing decision problem for each target candidate. The fitness should be accompanied by a penalty value for infeasible solutions. The fitness of each NLP problem is computed for selecting the better individuals in the population, and then to generate the next individuals. The migration operation of NHDE is used to help all individuals escape from this local cluster. This migration operation is performed only if the measure of population diversity fails to satisfy the desired tolerance.

<https://doi.org/10.1371/journal.pone.0179575.g002>

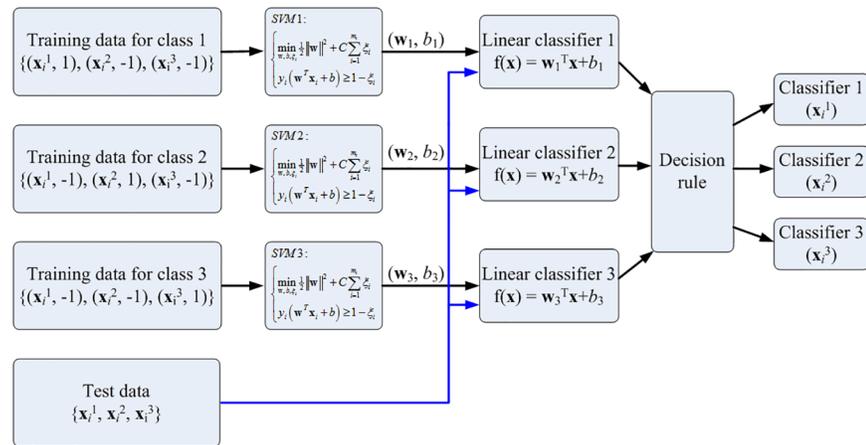


Fig 3. Flowchart for the triple SVM classifiers. 50% of data from class 1, 2, and 3 were used for training. A parallel architecture comprising triple SVMs was employed. The rest 50% of data were used for testing. Each SVM solves a two-class problem defined by one information class. The “winner-takes-all” rule is used for the final decision. The winning class is the one corresponding to the SVM with the highest output.

<https://doi.org/10.1371/journal.pone.0179575.g003>

The therapeutic and adverse objectives in the fuzzy optimization problem consider different disease states. However, the common drug targets—enzyme activities and external inputs—are identical for different diseases, so the variation effects are indiscriminate. The final objective is to minimize the overall design targets that consist of enzyme activities and external controls. The i^{th} reaction rate, v_{id} , in the material balance equation, depends on each disease state and is expressed in the power law function as follows:

$$v_{jd} = \alpha_{jd} \prod_{k=1}^n x_k^{g_{jk}} \tag{5}$$

where the rate constant or enzyme activity, α_{jd} , depends on whether the rate is normal or disease state, that is,

$$\alpha_{jd} = \begin{cases} \alpha_j, & j \in \Omega^{TE} \text{ and } j \notin \Omega^{DS} \\ \alpha_j^{basal}, & j \in \Omega^{DS} \text{ and } j \neq d \\ \alpha_j^{DS}, & j \in \Omega^{DS} \text{ and } j = d \end{cases} \tag{6}$$

The optimal drug target discovery problem in Eq (4) is a generalized optimization problem. It is used not only for detecting a common target but also for finding a specific target if the pathogenic enzyme can be completely identified. We applied a fuzzy decision making method—elicited generalized membership function to solve the problem. This approach is extended from our previous work but must handle multiple disease states. The optimization problem (4) can be transformed into a fuzzy decision-making problem as shown in supplementary (S2 File) and can then be solved using the NHDE algorithm (Fig 2 and S1 File) to obtain the optimal drug targets for different defects.

Results

We used a case of presynaptic dopamine overactivity, which may cause schizophrenia, as a case study to illustrate the performance of the proposed strategy in Fig 1. The mathematical model of the nigrostriatal dopamine network accessed from Qi et al. [42, 43] is expressed as the generalized mass action framework that consisted of 34 metabolites, 18 independent

variables, and 68 target enzymes. The detailed definition of the model is expressed in supporting information (S3 File). Here, the extracellular dopamine (DA-e) concentration was obtained from the model at steady state to be 400 (relative unit) assigned as the healthy state (HS). Other metabolite concentrations of interest include reactive oxygen species (ROS), reactive nitrogen species (RNS), and toxic species (for HS values, see S3 File).

Pathogenesis of presynaptic dopamine overactivity

After a clinical symptoms inquiry, the initial step in the objective section of Fig 1 is to examine dopamine overactive level, and then its disease value $x_i^{disease}$ is provided for pathogenesis problem (1) to determine a set of etiologies. To illustrate a wider region of pathogenesis, we considered seven disease states, that is, dopamine concentrations from 600 to 1200 (relative unit). The NHDE algorithm was applied to determine the one-hit defect to trigger increasing dopamine concentrations, and we found that two-enzyme activities decreased and one-enzyme activity increased, resulting in different overexpressed levels of dopamine (Fig 4A). In dopamine of 600 (equivalent to $\log_2(D/N) = 0.585$), we detected that three enzyme activities decreased, namely monoamine oxidase (MAO) (-26.3%), dopamine transporter (DAT) (-43.1%), and catechol-O-methyltransferase (COMT) (-93%) (not shown in Fig 4A), whereas one enzyme—vesicular monoamine transporter 2 (VMAT2)—increased (+57.3%) to reach this disease level. However, we could not find a feasible result in favor of a COMT defect when the dopamine level exceeded 600. From Fig 4A, we observed that smaller enzyme defects, resulting in higher dopamine overexpression, and MAO, were a key enzyme, compared with DAT and VMAT2, which causes presynaptic dopamine overexpression. The relative sensitivity of dopamine concentration (S1 Table) with respect to each enzyme activity was also applied to evaluate the regulation effects.

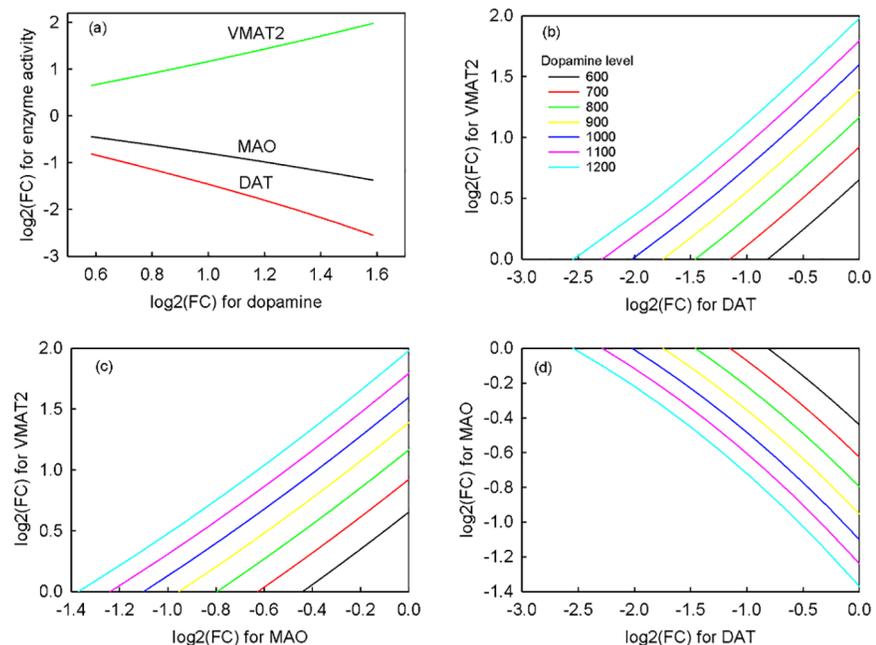


Fig 4. Dysregulated ratios of one-hit and two-hit enzymes. (a). single enzyme defect, (b). DAT and VMAT2 defects, (c). MAO and VMAT2 defects, (d). DAT and MAO defects. The enzyme activities of MAO and DAT decrease and the enzyme activity of VMAT2 increases. The combinations of two-enzyme dysregulation from the pairs of MAO, DAT, and VMAT2 cause an overexpression of various dopamine levels, namely 600, 700, 800, 900, 1000, 1100, and 1200. FC denotes the fold change between the disease state and the health state.

<https://doi.org/10.1371/journal.pone.0179575.g004>

Downregulation of MAO and DAT and upregulation of VMAT2 could raise dopamine levels. This observation was consistent with the optimization search as shown in Fig 4A.

The two-hit hypothesis, also known as the Knudson hypothesis, is the hypothesis that two mutant alleles can initiate tumorigenesis [44]. Such a two-hit hypothesis can be applied to infer schizophrenia that multiple genetic alterations have been suggested to affect the development of the central nervous system in a manner that creates an abnormal signaling network [45]. In this work, the pathogenesis problem was also applied to detect the two-enzyme dysregulation in order to interpret the behaviors. All possible combinations of two-enzyme dysregulation were evaluated from the pairs of MAO, DAT, and VMAT2. Fig 4B, 4C and 4D show that different dysregulated pairs of two-hit defects cause an overexpression of various dopamine levels. These results were then used for cluster analysis and for SVMs to identify biomarkers and to discover the first- and second-line drug targets.

Biomarkers detected by cluster analysis

To identify the possible biomarkers for presynaptic dopamine overactivity, we classified the disease states into different groups depending on the similarity of metabolites. We calculated the logarithmic ratio based on two of the metabolites as the form.

$$R_{i,\log 2} = \log 2 \left(\frac{x_i^{\text{disease}}}{x_i^{\text{basal}}} \right) \quad (7)$$

Then, hierarchical clustering was employed to classify various dopamine disease levels with different etiologies separately based on $R_{i,\log 2}$. Fig 5 shows the hierarchical tree in dopamine disease level of 1200. According to the hierarchical clustering result, the presynaptic dopamine overactivity with different etiologies could be qualitatively divided into two groups: one is the one with the MAO ($\alpha 15$) defect, and the other is without the MAO defect. We compared the $R_{i,\log 2}$ of metabolites in these two groups by applying the independent-samples T tests. The ratios of the 24 metabolites were statistically different in these two groups. Furthermore, we calculated the AUC of ROC curves (Table 1, other dopamine disease levels see S2 Table), and the values were greater than 0.95 with $p < 0.001$. Consequently, these 24 metabolites could be used as biomarkers to distinguish the states of presynaptic dopamine overactivity in those with a MAO defect from those without a MAO defect.

Biomarkers detected by SVM classifier

Single biomarker obtained through cluster analysis and AUC evaluation can be used for classification (Table 1). However, they can merely be applied to classify the set of etiologies into two groups (Fig 5), and cannot fulfill the requirement of precision medicine. According to the identification of pathogenesis, the disease states of presynaptic dopamine overactivity were caused mainly by the defects of MAO, DAT, and VMAT2. Considering the one-hit and two-hit defects, we applied the SVM classifier to distinguish the disease states at the same dopamine overexpressed level into three classes. Classes 1, 2, and 3 denoted as the defective pairs of (DAT, VMAT2), (DAT, MAO), and (VMAT2, MAO), respectively. Given a dopamine disease level (e.g. 1200) in Fig 4B, 4C and 4D, we uniformly generated 100 samples for each defective pair. According to the ANOVA test, no metabolites were statistically and significantly different among these three classes. To classify these three classes, we employed a machine-learning method with SVM and feature selection proposed in our previous work [40]. In the case study, the metabolic network comprised 34 metabolites. However, all variables were applied to train a SVM classifier that is impractical in clinical trial. Feature selection was used to determine the minimum number of biomarkers. We randomly used 50 samples for training and the other

Dendrogram using Average Linkage (Between Groups)

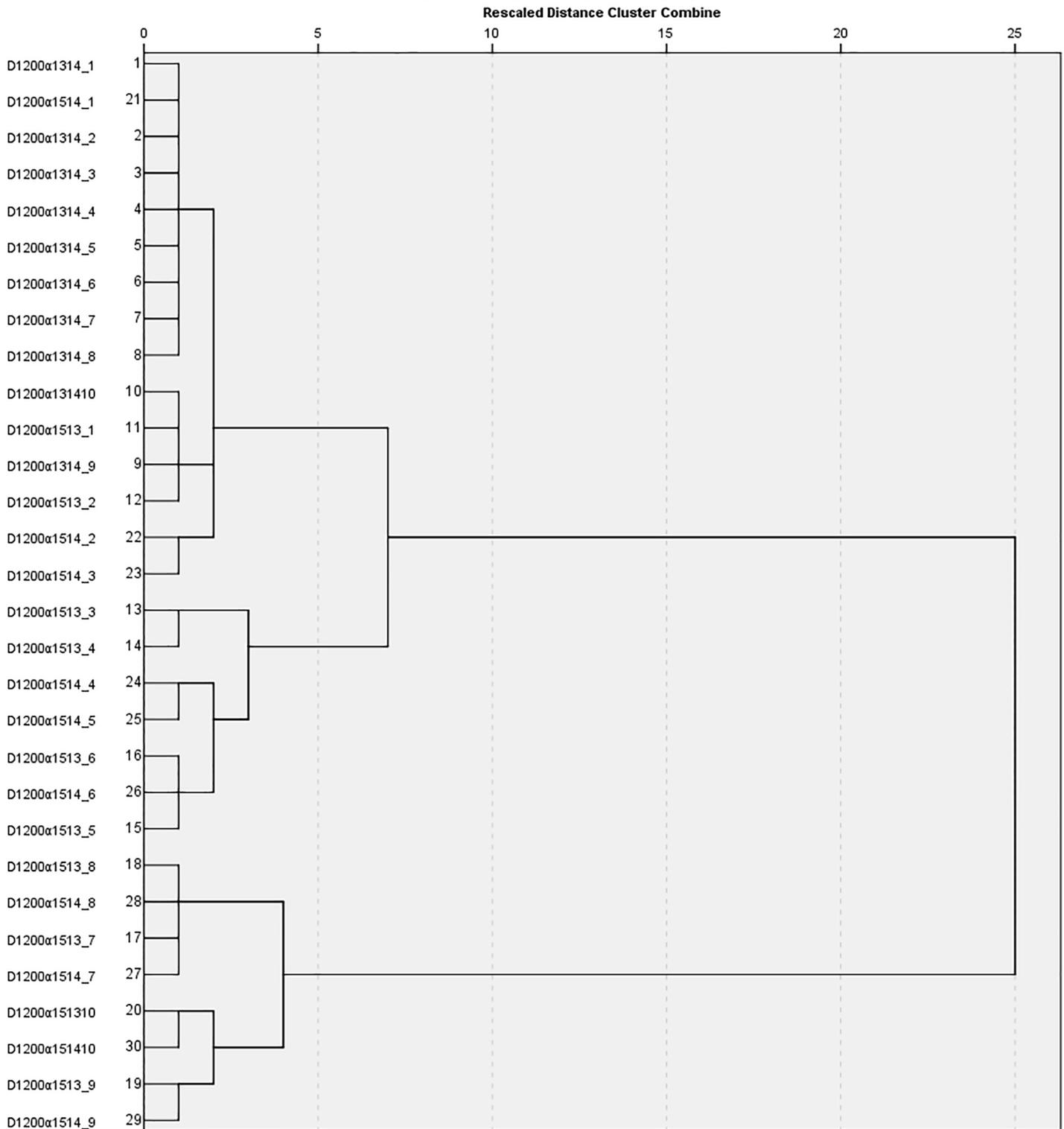


Fig 5. Hierarchical tree in dopamine disease level of 1200. The presynaptic dopamine overactivity with different etiologies could be qualitatively divided into two groups: one is the one with the MAO ($\alpha 15$) defect, and the other is without the MAO defect. D1200 $\alpha 1314_1$ denotes dopamine disease level of 1200 with DAT ($\alpha 13$) and VMAT2 ($\alpha 14$) defects, and we use 10 different combinations of two-enzyme dysregulation for cluster analysis.

<https://doi.org/10.1371/journal.pone.0179575.g005>

samples for testing. The accuracy of the 50 test samples utilizing SVM with selected biomarkers was first evaluated. The training of SVM and the selection of biomarkers were repeated until the desired accuracy was achieved, and then the optimal biomarkers were obtained. Table 2 shows the optimal biomarkers for a dopamine disease level of 1200 (other dopamine disease levels see S3 Table). The optimal accuracy of classification was 97.7% if three biomarkers were utilized for the features of SVM. By contrast, the optimal accuracy increased to 98.23% if four biomarkers were employed. We could use the optimal biomarkers to accurately assign the patients with presynaptic dopamine overactivity into three classes for furthering the second-line common drug target discovery (Fig 1).

Discovery of drug targets for different etiologies

In our previous studies [24, 25], we have introduced a fuzzy optimization method to discover the enzyme target in order to recover a disease state caused by only one enzyme defect. The method combined optimization search and decision making and considered therapeutic effects, adverse effects, and target variation effects simultaneously. This study extended this method to discover the first-line common drugs, second-line common drugs, and specific drugs according to different etiologies. The discovery of the first-line common drug targets occurred when the dopamine disease level and a set of etiologies was determined (Fig 1). We considered 33 etiological cases, which consisted of three one-hit defects and uniformly covered

Table 1. Candidates for biomarkers selected by cluster analysis and ROC curves at dopamine disease level of 1200.

Metabolites	p-value	AUC
Tyrosine (x_1)	<0.001	0.98
Dihydrobiopterin (x_2)	<0.001	0.95
Tyramine (x_3)	<0.001	0.95
L-DOPA produced from tyrosine (x_4)	<0.001	0.95
Dopaquinone (x_5)	<0.001	0.95
Pyrrolo-quinoline quinone (x_6)	<0.001	0.95
Intracellular dopamine (x_7)	<0.001	0.95
Prostaglandin H ₂ (x_{15})	<0.001	0.95
Dopamine quinone (x_{16})	<0.001	0.95
Dopamine chrome (x_{17})	<0.001	0.95
L-Dopachrome (x_{19})	<0.001	0.95
5,6-Dihydroxyindole (x_{20})	<0.001	0.95
Indole-5,6-quinone (x_{21})	<0.001	0.95
5,6-Dihydroxyindole-2-carboxylate (x_{22})	<0.001	0.95
Melanin (x_{23})	<0.001	0.95
Intracellular DOPAL (x_{24})	<0.001	0.95
Intracellular DOPAC (x_{25})	<0.001	0.95
DOPAC quinone (x_{26})	<0.001	0.95
Superoxide O ₂ ⁻ (x_{27})	<0.001	0.95
Intracellular hydrogen peroxide H ₂ O ₂ (x_{28})	<0.001	0.95
hydroxyl radical HO· (x_{30})	<0.001	0.95
Peroxynitrite HO·-NO ₂ ⁻ (x_{31})	0.001	0.95
Oxidized glutathione (x_{33})	<0.001	0.95
Dehydroascorbate (x_{34})	<0.001	0.95

ROC: Receiver operating characteristic, AUC: area under ROC curve

<https://doi.org/10.1371/journal.pone.0179575.t001>

Table 2. Selected biomarker and accuracies of the machine-learning method. Dopamine disease level is 1200. 50 samples were randomly selected for training and the rest 50 samples were used for testing. After 100 computations, the average accuracy was listed in this table.

Biomarkers				Accuracy (%)			
				Class 1	Class 2	Class 3	Overall
x_8	x_{14}	x_{30}		90.46	99.02	99.96	96.48
x_8	x_{26}	x_{34}		92.36	99.96	98.16	96.83
x_8	x_{21}	x_{32}		97.64	99.96	95.50	97.70
x_6	x_8	x_{11}	x_{21}	96.78	99.00	98.16	97.98
x_8	x_{13}	x_{15}	x_{32}	96.90	99.98	97.82	98.23

x_6 : Pyrrolo-quinoline quinine, x_8 : Dopamine packed in vesicles, x_{11} : Extracellular DOPAL, x_{13} : HVA, x_{14} : S-Adenosyl-L-homocysteine, x_{15} : Prostaglandin H₂, x_{21} : Indole-5,6-quinone, x_{26} : DOPAC quinine, x_{30} : hydroxyl radical HO·, x_{32} : Nitrogen dioxide NO₂, x_{34} : Dehydroascorbate.

<https://doi.org/10.1371/journal.pone.0179575.t002>

all possible combinations of two-hit defects. The fuzzy optimization method was applied to find enzyme targets among 33 etiological cases to achieve the defined objectives. In this case study, the identified enzyme targets are shown in Table 3 for remedying different dopamine disease levels. For the severest disease ($x_9 = 1200$), we found that a 0.79-fold change (FC) of tyrosine uptake (u_t) and 0.32 FC of tyrosine hydroxylase activity (α_1) were downregulated, and 1.71 FC of COMT activity (α_{26}) and 2.74 FC of dehydroascorbate transporter (α_{60}) were upregulated to yield 52.03% satisfaction levels for the therapeutic, adverse, and target variation effects. The satisfaction grades obtained from the optimization problem meant that the optimal enzyme targets could remedy the 33 etiological cases, and the worst case among these defects achieved an overall satisfaction grade of 0.52. The satisfaction grades were defined according to a fuzzy membership function (S2 File). This table shows that the lower the overexpression of dopamine disease levels was, the higher the overall satisfaction grade was and the fewer enzyme targets there were. We also considered only the therapeutic effect objective for identifying the first-line common drug targets of each deficiency level; the optimal results are shown in Table 3. Even though the therapeutic objective could achieve a 100% satisfaction grade for all situations, the adverse effect was completely unsatisfactory.

Utilizing the machine-learning classifier with three or four biomarkers, the disease states of each dopamine disease level could be classified into Classes 1, 2, and 3, respectively. For each class of disease state, we considered 12 etiological cases simulated by different dysregulated values to cover all possible combinations of the corresponding two-hit enzyme defects. Here, the drug target discovery problem for each disease case employed the candidate target

Table 3. First-line common drug targets for the set of etiologies.

Disease level	Triple objectives				Single objective			
	T*	A*	V*	Detected target	T*	A	V	Detected target
600	1	0.91	0.91	α_0, α_8	1	0	0.35	α_{54}
700	1	0.79	0.79	α_0, α_8	1	0	0.02	α_{23}, α_{51}
800	1	0.77	0.77	$\alpha_0, \alpha_8, \alpha_{61}$	1	0	0.4	α_0, α_{54}
900	0.97	0.65	0.65	$\alpha_0, \alpha_1, \alpha_{61}$	1	0	0.29	α_0, α_{54}
1000	0.88	0.5	0.5	$\alpha_0, \alpha_1, \alpha_{60}$	1	0	0.33	$\alpha_0, \alpha_1, \alpha_{54}$
1100	0.87	0.63	0.63	$\alpha_0, \alpha_1, \alpha_8, \alpha_{60}$	1	0	0.35	$\alpha_0, \alpha_8, \alpha_{22}, \alpha_{50}$
1200	0.52	0.52	0.52	$\alpha_0, \alpha_1, \alpha_{26}, \alpha_{60}$	1	0	0.29	$\alpha_0, \alpha_8, \alpha_{22}, \alpha_{42}$

T, A, and V denote the satisfaction grade for therapeutic, adverse, and variation effectiveness of drug target discovery problems, respectively.

The symbol * denotes the optimal grade. The value is obtained from the worst grade among 33 etiological cases. α_j is the rate constant of the reaction rate j .

<https://doi.org/10.1371/journal.pone.0179575.t003>

Table 4. Second-line common drug targets for class 1, 2 and 3, and the specific drug targets at dopamine disease level of 1200. In the second-line situations, the value is obtained from the worst grade among 12 etiological cases. The drug target discovery problem for each disease case employs the candidate target set $\Omega^{TE} \cup \Omega^{EX} = \{\alpha_0, \alpha_1, \alpha_{13}, \alpha_{14}, \alpha_{15}, \alpha_{26}, \alpha_{60}, \alpha_{61}\}$ obtained from the union of targets in Table 3.

Disease case	Triple objectives				Single objective			
	T*	A*	V*	Detected target	T*	A	V	Detected target
Class 1	0.84	0.77	0.77	$\alpha_8, \alpha_{15}, \alpha_{26}$	0.89	0.51	0.71	α_1, α_{15}
Class 2	0.65	0.65	0.65	$\alpha_0, \alpha_1, \alpha_{14}$	0.95	0	0.52	α_1, α_{14}
Class 3	0.95	0.62	0.62	$\alpha_0, \alpha_1, \alpha_{13}$	1	0	0.45	α_0, α_{13}
α_{13} (0.64), α_{14} (2.65)	1	0.74	0.74	α_8, α_{15}	1	0	0.7	α_{15}, α_{60}
α_{13} (0.35), α_{15} (0.73)	1	0.74	0.74	α_0, α_{14}	1	0	0.5	α_{14}
α_{14} (2.38), α_{15} (0.73)	1	0.77	0.77	$\alpha_0, \alpha_8, \alpha_{13}$	1	0	0.35	α_{13}
α_{13} (0.17)	1	0.81	0.81	α_{14}, α_{15}	1	0.97	0.54	α_{14}
α_{14} (3.95)	1	0.75	0.75	α_{13}, α_{15}	1	0.97	0.33	α_{13}
α_{15} (0.22)	1	0.86	0.86	$\alpha_0, \alpha_{13}, \alpha_{14}$	1	0	0.42	α_8, α_{14}

T, A, and V denote the satisfaction grade for therapeutic, adverse, and variation effectiveness of drug target discovery problems, respectively. The superscript * denotes as the optimal grade. α_j is the rate constant of the reaction rate j. α_j (FC) denotes the fold change of enzyme activity between the disease state and the health state.

<https://doi.org/10.1371/journal.pone.0179575.t004>

set $\Omega^{TE} \cup \Omega^{EX} = \{\alpha_0, \alpha_1, \alpha_{13}, \alpha_{14}, \alpha_{15}, \alpha_{26}, \alpha_{60}, \alpha_{61}\}$ obtained from the union of targets in Table 3 in order to find optimal enzyme targets. Consequently, the first-line common drug targets could also be utilized in second-line and specific drugs to avoid the cost of new drug developments. For each class, we discovered three optimal enzyme targets for remedying a dopamine disease level of 1200, and the overall satisfaction grades of Classes 1, 2, and 3 were 0.77, 0.65, and 0.62, respectively (Table 4). Finally, through enzyme kinetic analysis (Fig 1), we could diagnose which enzyme defect was involved, either MAO, DAT, or VMAT2. We then discovered the specific drug targets for each corresponding defect because of specified one-hit or two-hit defects, and obtained the two or three optimal enzyme targets for remedying dopamine 1200 for each specified defect (Table 4, other dopamine disease levels and candidate target set consisted of all enzymes see S4 Table). By considering the therapeutic effect only, we could achieve a 100% satisfaction grade for all situations, but the adverse effect was completely unsatisfactory.

Discussion

Schizophrenia is a complex disease with multiple risk factors and causes. Several hypotheses were proposed to explain the etiology of schizophrenia [17]. Presynaptic dopamine overactivity is one of the possible etiologies that have been inferred from the drugs for schizophrenia [46]. However, the causes of presynaptic dopamine overactivity remain under investigation. The proposed pathogenesis detection has identified four one-hit defective enzymes, namely MAO, DAT, VMAT2, and COMT that result in dopamine overexpression. In our simulation, the defect of COMT is almost inactive (-93%) to cause a dopamine level of 600. The involvement of deletions affecting the loci of the COMT gene has been revealed in several studies [47–50]. However, we could not find a feasible result for COMT defect when the dopamine level was greater than 600. Qi et al. [42, 43, 51–53] also revealed that MAO, DAT, VMAT2, and COMT were key components leading to dopamine dysfunction by using extensive Monte Carlo simulations. Nevertheless, our proposed method can identify the precise severity and combinations of these dysregulated enzymes.

The two-hit hypothesis has been applied to hypothesize that cancer is the result of accumulated genetic mutations to a cell [54]. In schizophrenia, the two-hit hypothesis suggests that a

combination of genetic susceptibility coupled with a distinct developmental insult can lead to the onset of the full clinical syndrome [45, 55]. Qi et al. [43] reported 50% dopamine excess caused by 9 scenarios of single, double and triple enzyme defects. According to our simulated results of the two-hit defects, we observed that different dysregulated pairs of enzymes caused an overexpression of various dopamine levels (Fig 4). Smaller defects of double enzymes comparing with single enzyme may still cause diseases because of the synergic effect of multiple enzyme dysfunctions. This phenomenon might explain the two-hit hypothesis of schizophrenia.

The major goals of personalized medicine are to predict an individual's susceptibility to developing an illness, achieve an accurate diagnosis, and optimize the most efficient response to treatment [4, 56]. Clinically, schizophrenia is commonly diagnosed on the basis of behavioral signs and symptoms; by contrast, biomarkers have recently been introduced into clinical psychiatry [57, 58]. Biomarkers can aid in staging and classification of the extent of a disease [59] allowing for disease stratification, which is the core concept of precision medicine [17]. In this work, the hierarchical clustering method and AUC could use a single biomarker to roughly classify the disease states into two groups, one with a MAO defect and the other without a MAO defect (Table 1). Furthermore, the SVM classifier was introduced to use three or four biomarkers to accurately categorize the patients with presynaptic dopamine overactivity into three classes to further second-line common drug target discovery (Table 2). The accuracy of classification exceeded 97%. Finally, such information can be provided for an enzyme kinetic analysis to detect which of the specified enzymes cause illness in order to discover a specific drug target.

There are limitations for the application of precision medicine in clinical practice. The traditional medicinal approach relies on a "one-size-fits-all" strategy, or stratified medicine that warrants tailored medicinal care only to a group or subgroup of patients with a known disease. One of the goals of precision medicine is to maximize the probability of curing a disease or ailment while minimizing the potential adverse effects of medicinal interventions in individual patients [60, 61]. In this study, the multiobjective fuzzy optimization formulation was introduced to discover the first-line common drugs, second-line common drugs, and specific drugs that have different etiologies. The approach combined an optimization search and decision making concurrently and simultaneously considered therapeutic effects, adverse effects, and target variation effects to achieve the goals of precision medicine. In the simulated results, we compared the target discovery problem considering with or without adverse effects, and observed that the therapeutic effect could achieve a 100% satisfaction grade if the problem was considered the therapeutic objective only, but the adverse effect was completely unsatisfactory. The multiobjective optimization approach was performed with a trade-off procedure for obtaining a compromise discovery to yield reduced therapeutic effect to improve adverse and target variation effects (Tables 3 and 4).

To identify the second-line and specific drug targets, the candidate target set from the first-line common drug targets was employed. Therefore, we could select different combinations and dosages of remedies among the same drugs targets, rather than develop new medications for common and specific drug targets. The satisfaction grades of first-line common drug targets are lower than that of second-line and specific drug targets. However, medications of first-line common drug targets can treat disease states with different etiologies when biomarkers are not available. As more biomarkers and enzyme activities can be examined, higher satisfaction grades can be obtained by second-line and specific drug targets (Table 4).

The SOAP note is a problem oriented medical record (Fig 1). Optimization processes in SOAP, i.e. pathogenesis identification, biomarker detection, and drug discovery, should be conducted in advance by biomedical scientists. When treating a patient, a physician just need to measure dopamine level, check biomarkers, and then choose the optimal drugs according to

the protocol. Further tests, such as DNA sequencing, gene expression levels and metabolic levels might be helpful to define pathological states for specific drugs. However, physicians can still use common drugs if advanced tests are not available. The disease states may be altered by remedy, and therefore it is indicated to follow up these tests periodically.

In our previous studies [24, 25], we introduced a fuzzy multiobjective optimization method for drug discovery. In this study, we extended this method to pathogenesis identification, biomarker detection, and designs of common and specific drugs. In addition, we implemented these results to draw a SOAP flowchart that can be employed by physicians in clinical practice easily. However, we would like to emphasize that the accuracy of this method depends on the accuracy of the model. For example, the model we used in this study focus on the presynaptic dopamine metabolism, but schizophrenia is also related to glutamatergic, GABAergic, glycinergic and serotonergic systems [62]. If these complex systems could be enrolled in the model, the proposed optimization approaches can be applied to the whole process of precision medicine.

Conclusions

This study proposed a strategy of model-based optimization for precision medicine to remedy presynaptic dopamine overexpression. The strategy consisted of pathogenesis detection, biomarker identification, enzyme kinetic analysis, and drug target discovery for treating a patient. The procedures were similar to SOAP notes used by physicians and other healthcare professionals. One-hit and two-hit enzyme defects were detected by the pathogenesis problem to determine a set of etiologies, which was similar to the results obtained by Monte Carlo computations [51]. The identified biomarkers were then applied to classify such etiologies into three classes. Moreover, the specific enzyme defects could be identified if experiments for enzyme kinetic analysis were available. Finally, the multiobjective fuzzy optimization formulation was introduced to discover the first-line common drugs, second-line common drugs, and specific drugs for different etiologies. The approach combined optimization search and decision making concurrently and considered therapeutic effects, adverse effects, and target variation effects simultaneously to achieve the goals of precision medicine.

Supporting information

S1 Table. Relative sensitivity. The relative sensitivity of dopamine concentration with respect to each enzyme activity.

(XLSX)

S2 Table. Single biomarker. Candidates of biomarkers selected by cluster analysis and ROC curves for different dopamine disease levels.

(XLSX)

S3 Table. Multiple biomarkers. Selected biomarkers and accuracies obtained by the machine-learning method for different dopamine disease levels.

(XLSX)

S4 Table. Optimal enzyme targets. Drug target discovery problem for second-line common drug targets and the specific drug targets using candidate target set consisted of all enzymes.

(XLSX)

S1 File. NHDE algorithm. The detailed computational procedure of NHDE.

(DOCX)

S2 File. Solving strategy. Definition of the symbols in the pathogenesis problem and fuzzy multiobjective target discovery problem and its solving strategy.

(DOCX)

S3 File. Presynaptic dopamine model. List of dependent and independent variables and kinetic equations for presynaptic dopamine metabolic model.

(DOCX)

S4 File. Glossary of clinical terms. List of definitions of clinical terms.

(DOCX)

Acknowledgments

The financial support from Ministry of Science and Technology of Taiwan (Grant MOST103-2221-E-194-045-MY3, MOST104-2627-B-194-001 and MOST105-2627-M-194-001), is highly appreciated.

Author Contributions

Conceptualization: KCH FSW.

Data curation: KCH FSW.

Formal analysis: KCH FSW.

Funding acquisition: FSW.

Investigation: KCH FSW.

Methodology: KCH FSW.

Project administration: FSW.

Resources: KCH FSW.

Software: KCH FSW.

Supervision: FSW.

Validation: KCH FSW.

Visualization: KCH FSW.

Writing – original draft: KCH.

Writing – review & editing: FSW.

References

1. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med.* 2015; 372(9):793–5. doi: [10.1056/NEJMp1500523](https://doi.org/10.1056/NEJMp1500523). PMID: [25635347](https://pubmed.ncbi.nlm.nih.gov/25635347/).
2. Jameson JL, Longo DL. Precision medicine—personalized, problematic, and promising. *N Engl J Med.* 2015; 372(23):2229–34. doi: [10.1056/NEJMs1503104](https://doi.org/10.1056/NEJMs1503104). PMID: [26014593](https://pubmed.ncbi.nlm.nih.gov/26014593/).
3. Mehta R, Jain RK, Badve S. Personalized medicine: the road ahead. *Clin Breast Cancer.* 2011; 11(1):20–6. doi: [10.3816/CBC.2011.n.004](https://doi.org/10.3816/CBC.2011.n.004). PMID: [21421518](https://pubmed.ncbi.nlm.nih.gov/21421518/).
4. Ozomaro U, Wahlestedt C, Nemeroff CB. Personalized medicine in psychiatry: problems and promises. *BMC Med.* 2013; 11:132. doi: [10.1186/1741-7015-11-132](https://doi.org/10.1186/1741-7015-11-132). PMID: [23680237](https://pubmed.ncbi.nlm.nih.gov/23680237/); PubMed Central PMCID: [PMC3668172](https://pubmed.ncbi.nlm.nih.gov/PMC3668172/).

5. Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology. *Schizophrenia research*. 2008; 102(1):1–18.
6. Carlsson A, Lindqvist M, Magnusson T. 3, 4-Dihydroxyphenylalanine and 5-Hydroxytryptophan as Reserpine Antagonists. *Nature*. 1957; 180:1200.
7. Creese I, Burt DR, Snyder SH. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science*. 1976; 192(4238):481–3. PMID: [3854](#)
8. Seeman P. Schizophrenia as a brain disease. The dopamine receptor story. *Arch Neurol*. 1993; 50(10):1093–5. PMID: [8215969](#).
9. Seeman P, Lee T. Antipsychotic drugs: direct correlation between clinical potency and presynaptic action on dopamine neurons. *Science*. 1975; 188(4194):1217–9. PMID: [1145194](#)
10. Seeman P, Schwarz J, Chen JF, Szechtman H, Perreault M, McKnight GS, et al. Psychosis pathways converge via D2high dopamine receptors. *Synapse*. 2006; 60(4):319–46. <https://doi.org/10.1002/syn.20303> PMID: [16786561](#)
11. Hietala J, Syvalahti E, Vuorio K, Rakkolainen V, Bergman J, Haaparanta M, et al. Presynaptic dopamine function in striatum of neuroleptic-naive schizophrenic patients. *Lancet*. 1995; 346(8983):1130–1. PMID: [7475604](#).
12. Lyon GJ, Abi-Dargham A, Moore H, Lieberman JA, Javitch JA, Sulzer D. Presynaptic regulation of dopamine transmission in schizophrenia. *Schizophr Bull*. 2011; 37(1):108–17. doi: [10.1093/schbul/sbp010](#). PMID: [19525353](#); PubMed Central PMCID: PMCPMC3004182.
13. McGowan S, Lawrence AD, Sales T, Quedest D, Grasby P. Presynaptic dopaminergic dysfunction in schizophrenia: a positron emission tomographic [¹⁸F]fluorodopa study. *Arch Gen Psychiatry*. 2004; 61(2):134–42. doi: [10.1001/archpsyc.61.2.134](#). PMID: [14757589](#).
14. Reith J, Benkelfat C, Sherwin A, Yasuhara Y, Kuwabara H, Andermann F, et al. Elevated dopa decarboxylase activity in living brain of patients with psychosis. *Proc Natl Acad Sci U S A*. 1994; 91(24):11651–4. PMID: [7972118](#); PubMed Central PMCID: PMCPMC45289.
15. Tretter F, Gebicke-Haerter PJ. Systems biology in psychiatric research: from complex data sets over wiring diagrams to computer simulations. *Methods Mol Biol*. 2012; 829:567–92. doi: [10.1007/978-1-61779-458-2_36](#). PMID: [22231839](#).
16. Linden DE. The challenges and promise of neuroimaging in psychiatry. *Neuron*. 2012; 73(1):8–22. doi: [10.1016/j.neuron.2011.12.014](#). PMID: [22243743](#).
17. Alawieh A, Zaraket FA, Li JL, Mondello S, Nokkari A, Razafsha M, et al. Systems biology, bioinformatics, and biomarkers in neuropsychiatry. *Front Neurosci*. 2012; 6:187. doi: [10.3389/fnins.2012.00187](#). PMID: [23269912](#); PubMed Central PMCID: PMCPMC3529307.
18. Woods AG, Sokolowska I, Taurines R, Gerlach M, Dudley E, Thome J, et al. Potential biomarkers in psychiatry: focus on the cholesterol system. *J Cell Mol Med*. 2012; 16(6):1184–95. doi: [10.1111/j.1582-4934.2012.01543.x](#). PMID: [22304330](#); PubMed Central PMCID: PMCPMC3823072.
19. Guillen-Gosalbez G, Sorribas A. Identifying quantitative operation principles in metabolic pathways: a systematic method for searching feasible enzyme activity patterns leading to cellular adaptive responses. *BMC Bioinformatics*. 2009; 10:386. doi: [10.1186/1471-2105-10-386](#). PMID: [19930714](#); PubMed Central PMCID: PMCPMC2799421.
20. Materi W, Wishart DS. Computational systems biology in drug discovery and development: methods and applications. *Drug Discov Today*. 2007; 12(7–8):295–303. doi: [10.1016/j.drudis.2007.02.013](#). PMID: [17395089](#).
21. Stone JA, Banfield C, Pfister M, Tannenbaum S, Allerheiligen S, Wetherington JD, et al. Model-based drug development survey finds pharmacometrics impacting decision making in the pharmaceutical industry. *J Clin Pharmacol*. 2010; 50(9 Suppl):20S–30S. doi: [10.1177/0091270010377628](#). PMID: [20881214](#).
22. Vera J, Curto R, Cascante M, Torres NV. Detection of potential enzyme targets by metabolic modelling and optimization: application to a simple enzymopathy. *Bioinformatics*. 2007; 23(17):2281–9. doi: [10.1093/bioinformatics/btm326](#). PMID: [17586544](#).
23. Vera J, González-Alcón C, Marín-Sanguino A, Torres N. Optimization of biochemical systems through mathematical programming: Methods and applications. *Computers & Operations Research*. 2010; 37(8):1427–38.
24. Hsu KC, Wang FS. Fuzzy optimization for detecting enzyme targets of human uric acid metabolism. *Bioinformatics*. 2013; 29(24):3191–8. doi: [10.1093/bioinformatics/btt564](#). PMID: [24078686](#).
25. Hsu KC, Wang FS. Fuzzy Decision Making Approach to Identify Optimum Enzyme Targets and Drug Dosage for Remedying Presynaptic Dopamine Deficiency. *PLoS One*. 2016; 11(10):e0164589. doi: [10.1371/journal.pone.0164589](#). PMID: [27736960](#).
26. Quinn L, Gordon J. Functional outcomes documentation for rehabilitation: WB Saunders Co; 2003.

27. Voit EO. A first course in systems biology: Garland Science; 2012.
28. McLachlan GJ. Cluster analysis and related techniques in medical research. *Stat Methods Med Res.* 1992; 1(1):27–48. PMID: [1341650](#). <https://doi.org/10.1177/096228029200100103>
29. Art D, Gnanadesikan R, Kettenring J. Data-based metrics for cluster analysis. *Utilitas Mathematica A.* 1982; 21:75–99.
30. Vapnik V. The nature of statistical learning theory: Springer Science & Business Media; 2013.
31. Huang C-L, Wang C-J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with applications.* 2006; 31(2):231–40.
32. Pontil M, Verri A. Support vector machines for 3D object recognition. *IEEE transactions on pattern analysis and machine intelligence.* 1998; 20(6):637–46.
33. Gillani Z, Akash MS, Rahaman MD, Chen M. CompareSVM: supervised, Support Vector Machine (SVM) inference of gene regularity networks. *BMC Bioinformatics.* 2014; 15:395. doi: [10.1186/s12859-014-0395-x](#). PMID: [25433465](#); PubMed Central PMCID: [PMC4260380](#).
34. Ben-Hur A, Ong CS, Sonnenburg S, Scholkopf B, Ratsch G. Support vector machines and kernels for computational biology. *PLoS Comput Biol.* 2008; 4(10):e1000173. doi: [10.1371/journal.pcbi.1000173](#). PMID: [18974822](#); PubMed Central PMCID: [PMC42547983](#).
35. Noble WS. What is a support vector machine? *Nature biotechnology.* 2006; 24(12):1565–7. <https://doi.org/10.1038/nbt1206-1565> PMID: [17160063](#)
36. Yang ZR. Biological applications of support vector machines. *Brief Bioinform.* 2004; 5(4):328–38. PMID: [15606969](#).
37. Bishop CM. *Pattern recognition and machine learning.* New York: Springer; 2006.
38. Duda RO, Hart PE, Stork DG. *Pattern classification:* John Wiley & Sons; 2012.
39. Friedman J, Hastie T, Tibshirani R. *The elements of statistical learning:* Springer series in statistics Springer, Berlin; 2001.
40. Hsu K-C, Yu S-N. Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm. *Computers in Biology and Medicine.* 2010; 40(10):823–30. <https://doi.org/10.1016/j.compbiomed.2010.08.005> PMID: [20832782](#)
41. Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing.* 2004; 42(8):1778–90.
42. Qi Z, Miller GW, Voit EO. Computational systems analysis of dopamine metabolism. *PLoS One.* 2008; 3(6):e2444. doi: [10.1371/journal.pone.0002444](#). PMID: [18568086](#); PubMed Central PMCID: [PMC435046](#).
43. Qi Z, Miller GW, Voit EO. A mathematical model of presynaptic dopamine homeostasis: implications for schizophrenia. *Pharmacopsychiatry.* 2008; 41 Suppl 1:S89–98. doi: [10.1055/s-2008-1080936](#). PMID: [18756426](#).
44. Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer.* 2001; 1(2):157–62. doi: [10.1038/35101031](#). PMID: [11905807](#).
45. Feigenson KA, Kusnecov AW, Silverstein SM. Inflammation and the two-hit hypothesis of schizophrenia. *Neurosci Biobehav Rev.* 2014; 38:72–93. doi: [10.1016/j.neubiorev.2013.11.006](#). PMID: [24247023](#); PubMed Central PMCID: [PMC43896922](#).
46. Baumeister AA, Francis JL. Historical development of the dopamine hypothesis of schizophrenia. *J Hist Neurosci.* 2002; 11(3):265–77. doi: [10.1076/j.hin.11.3.265.10391](#). PMID: [12481477](#).
47. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet.* 2008; 40(7):827–34. doi: [10.1038/ng.171](#). PMID: [18583979](#).
48. Crespi B, Stead P, Elliot M. Evolution in health and medicine Sackler colloquium: Comparative genomics of autism and schizophrenia. *Proc Natl Acad Sci U S A.* 2010; 107 Suppl 1:1736–41. doi: [10.1073/pnas.0906080106](#). PMID: [19955444](#); PubMed Central PMCID: [PMC42868282](#).
49. Lewandowski KE. Relationship of catechol-O-methyltransferase to schizophrenia and its correlates: evidence for associations and complex interactions. *Harv Rev Psychiatry.* 2007; 15(5):233–44. doi: [10.1080/10673220701650409](#). PMID: [17924258](#).
50. Molero P, Ortuno F, Zalacain M, Patino-Garcia A. Clinical involvement of catechol-O-methyltransferase polymorphisms in schizophrenia spectrum disorders: influence on the severity of psychotic symptoms and on the response to neuroleptic treatment. *Pharmacogenomics J.* 2007; 7(6):418–26. doi: [10.1038/sj.tpj.6500441](#). PMID: [17363961](#).
51. Qi Z, Miller GW, Voit EO. Computational analysis of determinants of dopamine (DA) dysfunction in DA nerve terminals. *Synapse.* 2009; 63(12):1133–42. doi: [10.1002/syn.20686](#). PMID: [19670315](#); PubMed Central PMCID: [PMC42759407](#).

52. Qi Z, Miller GW, Voit EO. Computational modeling of synaptic neurotransmission as a tool for assessing dopamine hypotheses of schizophrenia. *Pharmacopsychiatry*. 2010; 43 Suppl 1:S50–60. doi: [10.1055/s-0030-1248317](https://doi.org/10.1055/s-0030-1248317). PMID: [20486051](https://pubmed.ncbi.nlm.nih.gov/20486051/)
53. Qi Z, Miller GW, Voit EO. Mathematical models in schizophrenia. *Handbook of Schizophrenia Spectrum Disorders, Volume I*: Springer; 2011. p. 305–25.
54. Nordling CO. A new theory on cancer-inducing mechanism. *Br J Cancer*. 1953; 7(1):68–72. PMID: [13051507](https://pubmed.ncbi.nlm.nih.gov/13051507/); PubMed Central PMCID: [PMCPMC2007872](https://pubmed.ncbi.nlm.nih.gov/PMC/PMC2007872/).
55. Maynard TM, Sikich L, Lieberman JA, LaMantia AS. Neural development, cell-cell signaling, and the "two-hit" hypothesis of schizophrenia. *Schizophr Bull*. 2001; 27(3):457–76. PMID: [11596847](https://pubmed.ncbi.nlm.nih.gov/11596847/).
56. Montine TJ, Montine KS. Precision medicine: Clarity for the clinical and biological complexity of Alzheimer's and Parkinson's diseases. *The Journal of experimental medicine*. 2015; 212(5):601–5. <https://doi.org/10.1084/jem.20150656> PMID: [25941321](https://pubmed.ncbi.nlm.nih.gov/25941321/)
57. Domenici E, Wille DR, Tozzi F, Prokopenko I, Miller S, McKeown A, et al. Plasma protein biomarkers for depression and schizophrenia by multi analyte profiling of case-control collections. *PLoS One*. 2010; 5(2):e9166. doi: [10.1371/journal.pone.0009166](https://doi.org/10.1371/journal.pone.0009166). PMID: [20161799](https://pubmed.ncbi.nlm.nih.gov/20161799/); PubMed Central PMCID: [PMCPMC2820097](https://pubmed.ncbi.nlm.nih.gov/PMC/PMC2820097/).
58. Kraemer HC, Schultz SK, Arndt S. Biomarkers in psychiatry: methodological issues. *Am J Geriatr Psychiatry*. 2002; 10(6):653–9. PMID: [12427574](https://pubmed.ncbi.nlm.nih.gov/12427574/).
59. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001; 69(3):89–95. doi: [10.1067/mcp.2001.113989](https://doi.org/10.1067/mcp.2001.113989). PMID: [11240971](https://pubmed.ncbi.nlm.nih.gov/11240971/).
60. Barbolosi D, Ciccolini J, Lacarelle B, Barlesi F, Andre N. Computational oncology—mathematical modelling of drug regimens for precision medicine. *Nat Rev Clin Oncol*. 2016; 13(4):242–54. doi: [10.1038/nrclinonc.2015.204](https://doi.org/10.1038/nrclinonc.2015.204). PMID: [26598946](https://pubmed.ncbi.nlm.nih.gov/26598946/).
61. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov*. 2016; 15(7):473–84. doi: [10.1038/nrd.2016.32](https://doi.org/10.1038/nrd.2016.32). PMID: [26965202](https://pubmed.ncbi.nlm.nih.gov/26965202/).
62. Avila-Rojas H, Sandoval-Zamora H, Perez-Neri I. Neurotransmitter systems, neuroanatomical pathology and cell death in schizophrenia: update and perspectives. *Invest Clin*. 2016; 57(2):217–30. PMID: [28429899](https://pubmed.ncbi.nlm.nih.gov/28429899/).