

# Fluvoxamine treatment response prediction in obsessive-compulsive disorder: association rule mining approach

This article was published in the following Dove Medical Press journal:  
*Neuropsychiatric Disease and Treatment*

Hesam Hasanpour<sup>1</sup>  
Ramak Ghavamizadeh  
Meibodi<sup>1</sup>  
Keivan Navi<sup>1</sup>  
Jamal Shams<sup>2</sup>  
Sareh Asadi<sup>3</sup>  
Abolhassan Ahmadiani<sup>4</sup>

<sup>1</sup>Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran; <sup>2</sup>Behavioral Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran; <sup>3</sup>Neurobiology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran; <sup>4</sup>Neuroscience Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Correspondence: Sareh Asadi  
Neurobiology Research Center, Shahid Beheshti University of Medical Sciences, PO Box 19615-1178, Tehran, Iran  
Tel +98 21 2242 9768  
Email s.asadi@sbmu.ac.ir

Abolhassan Ahmadiani  
Neuroscience Research Center, Shahid Beheshti University of Medical Sciences, PO Box 19615-1178, Tehran, Iran  
Tel +98 21 2242 9768  
Email aahmadiani@yahoo.com

**Background:** Obsessive-compulsive disorder (OCD) is a debilitating psychiatric disorder characterized by intrusive thoughts or repetitive behaviors. Clinicians use serotonin reuptake inhibitors (SRIs) for OCD treatment, but 40%–60% of the patients do not respond to them adequately. Here, we described an association rule mining approach for treatment response prediction using an Iranian OCD data set.

**Patients and methods:** Three hundred and thirty OCD patients fulfilling *DSM-5* criteria were initially included, but 151 subjects completed their pharmacotherapy which was defined as 12-week treatment with fluvoxamine (150–300 mg). Treatment response was considered as >35% reduction in the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) score. Apriori algorithm was applied to the OCD data set for extraction of the association rules predicting response to fluvoxamine pharmacotherapy in OCD patients. We considered the association of each attribute with treatment response using interestingness measures and found important attributes that associated with treatment response.

**Results:** Results showed that low obsession and compulsion severities, family history of mental illness, illness duration less than 5 years, being married, and female were the most associated variables with responsiveness to fluvoxamine pharmacotherapy. Meanwhile, if an OCD patient reported a family history of mental illness and his/her illness duration was less than 5 years, he/she responded to 12-week fluvoxamine pharmacotherapy with the probability of 91%. We also found useful and applicable rules for resistant and refractory patients.

**Conclusion:** This is the first study where association rule mining approach was used to extract predicting rules for treatment response in OCD. Application of this method in personalized medicine may help clinicians in taking the right therapeutic decision.

**Keywords:** data mining, apriori algorithm, family history, refractory patients

## Introduction

Obsessive-compulsive disorder (OCD) is a heterogeneous neuropsychiatric condition that affects 3% of the population worldwide<sup>1–3</sup> and ~1.8% in Iran.<sup>4</sup> Overall, the prevalence of OCD is slightly more in women than men with an earlier onset in the male patients.<sup>5</sup>

Evidence indicated that OCD patients compared to healthy subjects have functional impairment and poorer quality of life<sup>6</sup> which is lower for samples with higher percentages of females and was negatively associated with the age.<sup>7</sup> However, a significant number of OCD patients (40%–60%) fail to benefit sufficiently from pharmacologic treatments.<sup>8–11</sup> Different factors influence the effectiveness of OCD treatment suggesting personalized medicine as a solution. Research studies provide

vast amounts of heterogeneous data that are too complex and voluminous to be processed and analyzed by traditional methods. Although the conventional methods have their application in medical data analyses, nonetheless, they have their limitations and we need inductive approaches such as data mining methods to discover knowledge with a small number of cause–effect mechanisms.<sup>12</sup>

Data mining approach is becoming increasingly popular for providing a deeper understanding of medical data, including disease pathogenesis and treatment leading to a new discovery from the medical data set that the conventional methods are unable to process due to their limitations.<sup>13–16</sup> Data mining algorithms are classified into two main categories: supervised and unsupervised learning.<sup>17</sup> Supervised learning algorithms such as classification or regression predict the response values for a particular outcome. Unsupervised learning algorithms describe the form and hidden structure of data, using methods such as clustering and anomaly detection. One of the primary unsupervised algorithms is association rule mining that has been widely explored in different areas, including medicine. Association rule mining discovers associations and correlations among large sets of data and provides information in the form of if-then probabilistic statements.<sup>15,18</sup> Different algorithms have been proposed to generate association rules.<sup>19–21</sup> Apriori is a well-known primary association rule mining algorithm for extraction of frequent items in a set of transactions. It was first introduced by Agrawal et al.<sup>19</sup> Sharma and Om used this algorithm for early diagnosis and treatment of oral cancer.<sup>22</sup> Karabatak and Ince also proposed an expert system for detection of breast cancer based on association rules and neural network.<sup>14</sup>

Predictive Apriori,<sup>21</sup> motivated by Apriori, maximizes the accuracy of an association rule in contrast to the confidence related to focus of Apriori. Apriori ranks the rules based on confidence only, whereas predictive Apriori considers both confidence and support in the rule ranking. Nahar et al consider three rule generation algorithms – Apriori, predictive Apriori, and Tertius<sup>20</sup> – to extract cancer prevention factors.<sup>23</sup> Their results showed that Apriori is the most useful association rule mining algorithm to be used in the discovery of prevention factors. In another study by this group, they used association rule mining approach to identify factors that contributed to heart disease considering sex.<sup>24</sup> Chen et al<sup>12</sup> used association rule mining to detect possible side effects due to the exposure time to drugs during pregnancy, which resulted in the discovery of interesting information.

There are limited studies in the field of psychiatry that use association rule mining approach for data extraction. Shen et al applied association rule mining to the Taiwan

National Health Insurance Research Database to explore associations among comorbidities of borderline personality disorder.<sup>25</sup> In another study, Schweikert et al applied Combinatorial Fusion Analysis (CFA) and association rule mining to explore the relationship between autism prevalence and lead mercury concentration, which led to a deeper understanding of its pathogenesis.<sup>26</sup> In the current study, we applied data mining algorithms to an OCD treatment response data set of Iranian patients to explore the most significant factors contributing to the treatment response. At first, we converted our data set variables which were chosen from previous studies<sup>9,27–31</sup> and based on experts' opinions to an appropriate format for application of Apriori algorithm. Next, we evaluated the association of each attribute with treatment response using support, confidence, and lift measures and found the most critical attributes that were associated with treatment response in OCD for pharmacotherapy with fluvoxamine. Finally, we used Apriori algorithm to explore the most associated rules which predict treatment response in OCD.

## Patients and methods

The preprocessing and analyses of all data were conducted using Matlab version 2014a (The Math Works Inc.).

## Subjects and treatment protocol

Three hundred and thirty outpatients with an Iranian origin and meeting *DSM-4* criteria for OCD were recruited from Imam Hossein hospital (Tehran, Iran) between 2014 and 2017. OCD diagnosis was established by a psychiatrist, and its subtypes and severity were investigated by an expert psychologist through the structured interview using Yale-Brown Obsessive Compulsive Scale (Y-BOCS) checklist and severity scale.<sup>32</sup> The sociodemographic data were collected through a questionnaire consisting of the patient's full name, address, age at assessment, the age of onset, gender, marital status, educational level, occupation, chief complaint, illness duration, history of substance usage, and the family history of psychiatric disorders, specifically OCD. Participants were in the age range of 18–65 years, having OCD symptoms for more than 1 year and with total Y-BOCS severity score more than 9 (suggested cutoff point for Iranian patients<sup>33</sup>); a total of 17 patients out of 151 had total Y-BOCS scores 9–16. The exclusion criteria were having a history of psychotic disorders or intellectual disability; reporting severe neurological pathology; history of substance usage; diagnosis with other *DSM-4* Axis I disorders except for depression, anxiety, or tic disorder; and under other pharmacotherapy or cognitive behavioral therapy (CBT).

Pharmacotherapy was defined as a 12-week treatment with fluvoxamine (150–300 mg), which has been described previously.<sup>28</sup> Briefly, the escalating fluvoxamine daily dose was initiated from 25 mg/day and increased up to 100 mg/day in the third week. Fluvoxamine daily dose for the next 3 weeks was 150 mg/day, and after the sixth week, patients were visited by the psychiatrist and received maximum tolerated dose of the fluvoxamine for the next 6 weeks. No concomitant therapy was allowed during the whole treatment period, either pharmacological or non-pharmacological. Y-BOCS severity was evaluated at the beginning of the treatment period and after 12 weeks, and based on the reduction in patients' Y-BOCS score, the patients were divided into two groups: group A (responders) included patients who showed >35% reduction in Y-BOCS scores after treatment with fluvoxamine, and group B (non-responders) was composed of patients who exhibited <35% reduction in Y-BOCS scores.<sup>34</sup> We included another group (refractory patients) consisting of patients who experienced various selective serotonin reuptake inhibitor (SSRI) trials at maximum tolerated dose during their illness period but did not respond to them adequately.<sup>34</sup>

From 330 OCD patients initially included in the study, 151 persons completed their pharmacotherapy. Others were excluded because of refusal to come for follow-up (N=108), not taking medication properly (N=38), discontinuity in pharmacotherapy due to complications, including allergy to the drug (N=8), meeting exclusion criteria during pharmacotherapy (N=6), and other reasons (N=19). Since we used the previously published data set to investigate the association rule data mining approach, the CONSORT diagram which summarizes the flow of participants through different stages of the trial was presented previously.<sup>28</sup> The excluded and included patients showed no significant differences in mean baseline scores considering their demographic (sex and marital status) and clinical characteristics (history of psychiatric disorders, age at assessment, age of onset, obsession severity, compulsion severity, and total severity) as well as symptom dimensions from the Y-BOCS symptom checklist (data not shown, all  $P>0.1$ ).

Sixty-four persons out of the 151 included patients were female, 66% were married, 58% of patients were unemployed, and 76% of them did not have academic degrees. Sixteen percentage of patients reported no history of mental illness in their family, 33% reported a history of mental illness, and 51% were from the families with a history of OCD.

Patients' age at onset (mean±SD) was 23±10.4 years and age at assessment (mean±SD) was 34.2±10.2 years. The mean of Y-BOCS scores at intake (mean±SD) was 10.84±4.7 for obsession, 9.3±5.39 for compulsion, and 20.1±8.8 for the

total severities. This research was performed in accordance with the latest version of Declaration of Helsinki and with the approval of the research ethics committee of Neuroscience Research Center, Shahid Beheshti University of Medical Sciences (IR.SBMU.PHNS.REC.1397.14). Written consent for participation was obtained from participants after being informed about the nature of the research.

## Variables and preprocessing

The analyzed OCD data set contains sociodemographic and clinical variables that were considered as independent variables (predictors) and treatment responses (responder, non-responders, refractory) which were considered as dependent variables.

Twenty-two variables were identified as predictors, including sex, marital status, family history, initial Y-BOCS obsession severity, initial Y-BOCS compulsion severity, Y-BOCS obsession subtypes, Y-BOCS compulsion subtypes, avoidance, depression score, and age of onset. These variables have to be converted to an appropriate format (true or false values) for applying associative rule mining algorithm. For example, sex translates into two columns: one is for being male and another for being female or marital status saved to one column named as married. Finally, we had a data set containing 37 columns for each sample, 34 columns for predictors, and three columns for treatment response status (Table 1). The time complexity of Apriori algorithm is exponential and takes too much time to extract associative rules,<sup>35</sup> so feature selection may help to overcome these problems. Some studies used association rules to find the most critical features and excluded redundant ones to reduce data set dimensions. Karabatak et al<sup>14</sup> developed an expert system for the detection of breast cancer based on the association rules and neural network. At first, they detected large item sets for each studied class. Next, they removed redundant features and used essential items to classify their data set. They applied this method in another study for diagnosis of erythematous-squamous diseases.<sup>36</sup> Using their strategy, we found out that being male, no history of psychiatric disorders and counting appeared in none of the frequent item sets, so we excluded them from our further analysis.

## Interestingness measures

An association rule has two parts: an antecedent (if) and a consequent (then) in the form  $X \rightarrow Y$ . This rule means that when  $X$  occurs, then  $Y$  also happens with a certain probability.<sup>19</sup> Usually, an overwhelming number of rules are created in association rule mining, and evaluating this amount of rules by a user is tedious and time-consuming.

**Table 1** Features used in the association rule mining algorithm

No	Feature	No	Feature	No	Feature	No	Feature
1	Female	2	Male	3	Married	4	Early onset (onset <18)
5	Late onset (onset ≥18)	6	No history	7	Family history of mental illness	8	Family history of OCD
9	Aggressive obsession	10	Contamination obsession	11	Sexual obsession	12	Hoarding obsession
13	Religious obsession	14	Symmetry obsession	15	Somatic obsession	16	Cleaning compulsion
17	Checking compulsion	18	Repeating compulsion	19	Counting compulsion	20	Ordering compulsion
21	Obsession severity ≤10	22	Obsession severity >10	23	Compulsion severity ≤10	24	Compulsion severity >10
25	Total severity ≤18	26	Total severity >18	27	Presence of insight	28	No insight
29	Presence of avoidance	30	Lack of avoidance	31	Gaf ≤3	32	Gaf >3
33	Illness duration ≤5 years	34	Illness duration >5 years	35	Responder	36	Non-responder
37	Refractory						

**Abbreviations:** Gaf, global assessment of functioning; OCD, obsessive-compulsive disorder.

Interestingness measures are beneficial and play an essential role in the detection of strong and interesting rules from the set of discovered rules. These indexes are intended for selecting and ranking rules according to their potential interest to the user. Confidence and support are two traditional measures of rule interestingness that reflect the certainty and usefulness of a rule, respectively.<sup>37</sup> The support is defined as a number of transactions in “data set” that include all items in the antecedent (if) and consequent (then) parts of the rule. The support is sometimes expressed as the percentage of the total number of records in the database (Equation 1). Item sets whose frequencies are greater than a predefined minimum support are called frequent item sets. Confidence is the conditional probability that a randomly selected transaction will consist of all the items in the consequent given that the transaction includes all the items in the antecedent (Equation 2). Confidence is the accuracy of the rule and is used in Apriori to rank the rules. The task of association rule mining is generation off all association rules from a set of transactions that have a support value greater than or equal to minimum support (Minsup) and confidence value greater than or equal to minimum confidence (Minconf). The values of Minsup and Minconf are set by the users and are parameters of the Apriori algorithm for association rule generation.

Rules with the high support and high confidence values are not necessarily impressive. Most of the rules, resulting from applying association rules mining to a medical data set, are irrelevant and do not have clinical importance.<sup>38</sup> So another parameter such as lift is used as a metric for

evaluation of medical significance and reliability of association rules. Lift is nothing but the ratio of confidence to expect confidence (Equation 3). A lift ratio larger than 1.0 implies that the relationship between the antecedent and the consequent is more significant than it would be expected and makes those rules potentially useful for predicting the consequent in the future. The larger the lift ratio, the more significant is the association. The support, confidence, and lift values were calculated using the following formulas:

$$Support(X \Rightarrow Y) = P(X \cap Y) \tag{1}$$

$$Confidence(X \Rightarrow Y) = P(Y|X) = (Support(X \cap Y))/(Support(X)) \tag{2}$$

$$Lift(X \Rightarrow Y) = (P(X \cap Y))/(P(X) * P(Y)) \tag{3}$$

where *X* and *Y* are two non-overlapping subsets of items. *P*(*X*) denotes the probability that a randomly selected transaction from data set will include all the items in *X*. *X* ∩ *Y* refers to all transactions that contain both *X* and *Y*.

### Applying Apriori algorithm to the OCD data set for the detection of association rules

The Apriori algorithm is a well-known association rule mining algorithm that works well on complete data sets with discrete values. Since we need to find the associations between attributes and treatment response, we considered

**Table 2** Pseudo code for Apriori algorithm that is suitable for supervised learning

```

Function rules=Apriori(T, minsup, minconf, classitem)
% T is the transaction data
% class item is the item that corresponds to class label

F2=Add class item to each element of F1
K=3
Repeat
); {generate candidate k-item sets}
Ck=prune(Ck); {prune candidate item sets}

If t contains c
Count(c)=count(c)+1
end for
end for

Until
//rule generation section
For each candidate item set f
A=f - class item % all items except class item, antecedent part of the rule
B=class item % consequent part of the rule
If confidence(A→B)>minconf
Add A→B to rules
end if
end for
return rules
    
```

OCD data set attributes as the antecedent parts of  $A \rightarrow B$  rules and the treatment response class (responder, non-responder, and refractory) as the consequent part.

The applied association rule mining algorithm consisted of two sections. In the first step, frequent patterns (item sets with the supports greater than predefined Minsup) were generated. In the second step, frequent patterns' confidences were calculated and those with confidence greater than Minconf were selected as the final rules. The pseudo-code of the used algorithm is presented in Table 2.

Choosing a single cutoff point for Minsup and Minconf may cause some problems in conducting Apriori algorithm.

If we used high Minsup threshold, Apriori algorithm would not detect association rules for the low-frequency classes, while with the low Minsup, Apriori algorithm took hours to complete and resulted in a lot of rules that most of them were redundant. To deal with this challenge, multiple minimum supports for different classes were assigned. Moreover, minimum support for each class was set by considering the proportion of its distribution over the whole data set.<sup>39,40</sup>

## Results

### Association of OCD data set attributes with the treatment response classes

Table 3 summarizes the most important attributes that associated with different treatment response classes (rules with lift <1 and support <0.08 were excluded and those left were sorted based on their confidence). We find rules in the form of  $A \rightarrow B$  such that  $A$  is a single attribute and  $B$  is the treatment response class (responder, non-responder, and refractory). As the results showed, no single attribute had very high confidence that could be used alone for the detection of treatment response class.

Results showed that low obsession severity (confidence=0.72), low compulsion severity (confidence=0.67), family history of mental illness (confidence=0.72), illness duration less than 5 years (confidence=0.68), being married (confidence=0.64), and being female (confidence=0.65) have strong associations with treatment responsiveness. For example, if an OCD patient has low obsession severity, he/she will be a responder to fluvoxamine pharmacotherapy with a probability equal to 0.72.

The most critical attributes associated with non-responsiveness were the sexual obsession, illness duration more than 5 years, being male, high compulsion severity, and having contamination obsession. High obsession severity, family history of OCD, early-onset OCD, and the presence of avoidance also showed association with treatment refractory class.

**Table 3** Most important attributes associated with treatment response (support >0.08, lift > 1; sorted based on confidence)

Responder		Non-responder		Refractory	
Feature	Confidence	Feature	Confidence	Feature	Confidence
Obsession severity ≤ 10	0.72	Sexual obsession	0.38	Obsession severity > 10	0.23
Family history of mental illness	0.72	Illness duration > 5 years	0.32	Family history of OCD	0.20
Illness duration ≤ 5 years	0.68	Being male	0.30	Early onset	0.20
Compulsion severity ≤ 10	0.67	Compulsion severity > 10	0.30	Presence of avoidance	0.18
Married	0.65	Contamination obsession	0.30		
Being female	0.64				

**Abbreviation:** OCD, obsessive-compulsive disorder.

However, the confidence values for associated attributes with non-responder and refractory classes were low (0.38–0.18).

## Association rules for OCD treatment response

The most important rules associated with each class that were extracted by Apriori algorithm are summarized in Table 4. For responder class, selected rules were those with confidence levels above 70%, support levels above 15%, and lifts greater than 1. The most associated rule with response to fluvoxamine pharmacotherapy (the family history of mental illness, aggressive obsessions, and illness duration less than 5 years → responder) implies that these three attributes are good predictors of being a responder. If someone has these characteristics simultaneously, he/she will be a responder with the probability of 92%. Comparing this rule with the following rules in Table 4 and results of the single attribute association analysis with treatment response class (Table 3) showed that a family history of mental illness and illness duration less than 5 years are the most important predictors of responder class. Late-onset OCD, being female, and married as well as low obsession and compulsion severities are other attributes that appeared in the subsequent rules, showing their importance in treatment responsiveness.

As previously mentioned, we set different Minsup and Minconf for each class. So for the non-responder class, Minsup and Minconf were considered as 0.07 and 0.6, respectively, and the most important rules are mentioned in Table 4. As it is shown in Table 4, the number of attributes that appeared in the antecedent part of the rules is more than the attributes in the responder class rules. For instance, if someone has five attributes, including contamination obsession, cleaning, repeating compulsions, and illness duration more than 5 years, he/she will not respond to treatment with the probability of 73%. High obsession and compulsion severities were the other attributes that appeared in the rules of non-responder class.

For refractory class, we set Minsup =0.03 and Minconf =0.6, and extracted rules were sorted based on their confidence (Table 4). The most important rule was detected as being female, early onset, aggressive obsession, and obsession severity more than 10. This rule implies that if a woman has an early-onset OCD with great obsession severity and aggressive obsession, she will not respond to most of the treatment strategies adequately with the probability of 83%. Early-onset OCD is an attribute that appeared in almost all the rules in this class, implicating the importance of this attribute for the refractory class.

**Table 4** Best discovered rules for different treatment responses

Class	Rule number	Antecedent	Supp	Conf
Responder	1	Family history of mental illness and Aggressive and Illness duration $\leq 5$	0.16	0.92
	2	Family history of mental illness and Illness duration $\leq 5$	0.19	0.91
	3	Late onset and Family history of mental illness and Illness duration $\leq 5$	0.16	0.89
	4	Female and Married and Late onset and Compulsion severity $\leq 10$	0.16	0.85
	5	Married and Obsession severity $\leq 10$ and Compulsion severity $\leq 10$	0.19	0.82
	6	Married and Hoarding and Obsession severity $\leq 10$	0.16	0.82
Non-responder	7	Contamination and Cleaning and Repeating and Illness duration $> 5$	0.08	0.73
	8	Contamination and Obsession severity $> 10$ and Total severity $> 18$ and Illness duration $> 5$	0.08	0.69
	9	Cleaning and Illness duration $> 5$	0.08	0.69
	10	Cleaning and checking and Obsession severity $> 10$ and Compulsion severity $> 10$	0.07	0.68
	11	Cleaning and checking and Obsession severity $> 10$ and Compulsion severity $> 10$ and Total severity $> 18$	0.07	0.68
Refractory	12	Female and Early onset and Aggressive and Obsession severity $> 10$	0.04	0.83
	13	Female and Married and Early onset and Checking and Gaf $\leq 3$	0.04	0.71
	14	Family history of OCD and Obsession severity $> 10$ and Presence of avoidance and Gaf $> 3$ and Illness duration $\leq 5$	0.04	0.71
	15	Female and Married and Early onset and aggressive and checking and Gaf $\leq 3$	0.03	0.71
	16	Female and Married and Early onset and checking and Gaf $\leq 3$ and Illness duration $> 5$	0.03	0.71

**Abbreviations:** Conf, confidence; Gaf, global assessment of functioning; OCD, obsessive-compulsive disorder; Supp, support.

## Discussion

This research has evaluated the efficacy of association rule mining for the first time to predict treatment response in OCD for pharmacotherapy with fluvoxamine using real data. In this study, we applied Apriori algorithm to the OCD pharmacotherapy data set and extracted the most associated features and the most important rules for each treatment response class based on the highest confidence level. Since the results of association rule data mining are in the form of if-then statements, their interpretability is high and easy to use in personalized medicine.

Personalized medicine centers on the assumption that each person has a unique physiologic characteristic which plays a significant role in both disease susceptibility and response to specific treatment. So the significant goals of personalized medicine are the prediction of an individual's vulnerability to developing an illness, achieve an accurate diagnosis, and optimize the most efficient response to treatment with the least side effects.<sup>41</sup>

Personalized medicine has vital importance in psychiatric disorders, since they are multifactorial diseases with a variable clinical course and considerable inter-individual variability in patient responses to therapies. About a third of patients respond to antidepressants, so psychiatrists use their clinical skill and experience to adapt the medical protocol for each patient.<sup>42</sup> However, in many cases, patients do not respond adequately to the medication or they report considerable side effects that result in a change in pharmacotherapy protocol. So, a series of drug alterations increase treatment duration as well as social and economic costs. Detection of predictors of treatment response may help psychiatrists to prescribe the most effective treatment with the least side effects considering individual characteristics.<sup>43</sup>

The first-line medications for OCD are antidepressants such as SSRIs; however, ~50% of OCD patients show a poor response.<sup>44</sup> Genetic factors as well as environmental (eg, smoking, foods, and beverages) and personal (eg, age, gender, ethnicity, and illnesses) variables affect the response to treatment.<sup>45</sup> Various studies detected different predictors for poor treatment response, including living without a partner,<sup>46</sup> sexual subtype,<sup>47</sup> higher OCD severity at intake,<sup>48,49</sup> poor insight into obsessions, symmetry/hoarding and contamination/washing dimension, and the presence of specific personality disorders.<sup>50</sup> Moreover, anterior cingulate cortex, orbitofrontal cortex, caudate, putamen, and thalamus volume changes and their altered activity were reported as predictors of poor response.<sup>50</sup> There are several pharmacogenetic reports related to the importance of genes involving serotonergic and glutamatergic signaling

pathways, but their results are inconsistent owing to the small sample size, lack of standardized methodology, and rigorous statistical analyses.<sup>44</sup> In the case of non-responsiveness, there are other strategies to augment the treatment response ranging from adding CBT to pharmacotherapy augmentations, such as antipsychotics, a combination of another serotonin reuptake inhibitors (SRIs), anticonvulsants, or even the use of neurosurgical procedures.<sup>51</sup> For example, the presence of anger associated with sexual obsessions and contamination may require the use of mood stabilizers as augmentation.<sup>52</sup> OCD patients may also benefit from cognitive (CT) and behavioral therapy (BT) separately or in combinations (CBT) which have been proven to be efficacious in multiple studies.<sup>53,54</sup> A recent meta-analysis comparing the efficacy of cognitive and behavioral therapies separately and in combination (CBT) in OCD treatment found no predictors for treatment response, but they indicated that improved patients had lower baseline depression and stronger beliefs about responsibility/threat and importance/control of thoughts. They showed that higher baseline scores on depression adversely affected outcomes for BT but not CT or CBT, and higher education and lower OCD severity were associated with better response to CT only.<sup>55</sup>

Data mining methods are useful tools in personalized medicine for the discovery of new knowledge, appropriate to diagnosis and treatment. In the current study, Apriori algorithm that is a well-known association rule mining algorithm was used to extract the most essential rules for the prediction of treatment response in OCD patients. Results showed that low obsession and compulsion severities, family history of mental illness, illness duration less than 5 years, being married, and being female are the most associated variables with responsiveness to fluvoxamine pharmacotherapy. Our findings are consistent with previous studies.<sup>9,29,30,46,49</sup> If an OCD patient reports a family history of mental illness and his/her illness duration is less than 5 years, he/she will respond to 12-week fluvoxamine pharmacotherapy with a probability of 91%. Although other clinical variables appeared in the associated rules, including late-onset OCD, being female and married as well as low obsession and compulsion severities are good predictors of responsiveness to treatment with fluvoxamine. Extracted rules revealed that if an OCD patient has contamination obsession, cleaning and repeating compulsions, and illness duration more than 5 years, he/she will not respond to fluvoxamine pharmacotherapy with the probability of 73%. As the psychiatric disorders are multifactorial diseases, it is highly unlikely that a single variable predicts treatment response in mental disorders, so considering these rules could be very useful for making clinical decisions.

To the best of our knowledge, this is the first study that applied association rule mining to detect rules associated with treatment response. This strategy has the advantage over traditional analytical approaches because of its higher accuracy. Moreover, the formats of results are easy to use by clinicians. The other strength of this study is prescribing a single drug, fluvoxamine, which avoided having confounded factors related to the different SSRI treatments. However, similar to any other research, this study also has some limitations. First, the small sample size which makes it necessary to apply this method to similar data sets with more samples. Second, although comparison of completers with non-completers considering their mean baseline characteristics showed no significant differences between these two groups, a large number of patients who did not complete their pharmacotherapy make it difficult to generalize these rules to the Iranian OCD patients. Third, treatment response was evaluated after 12-week pharmacotherapy, while assessment of treatment response's outcome in more extended period may detect stronger predictors of fluvoxamine responsiveness. Fourth, previous pharmacological or non-pharmacological treatments may affect response to fluvoxamine treatment. Fifth, baseline systematic assessments were not performed routinely for the included subjects. Altogether, it is suggested that this approach should be applied in a similar study to achieve more accurate rules across different populations.

## Conclusion

The application of data mining methods, especially association rule data mining coupled with personalized medicine, helps to extract feasible rules considering individual clinical features in disease diagnosis and treatment.

## Acknowledgments

This research was supported by a grant from Shahid Beheshti University of Medical Sciences, Iran (grant No 14553-6). We would like to thank patients and their families for participating in this study and Imam Hossain staff for their cooperation with our research team. We also gratefully thank NRC and Tarbiat Modarres graduate students and staff for participating in this study as a control group.

## Author contributions

All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published and agree to be accountable for all aspects of the work.

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H-U. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res.* 2012;21(3):169–184. doi:10.1002/mpr.1359
- Hasler G, LaSalle-Ricci VH, Ronquillo JG, et al. Obsessive-compulsive disorder symptom dimensions show specific relationships to psychiatric comorbidity. *Psychiatry Res.* 2005;135(2):121–132. doi:10.1016/j.psychres.2005.03.003
- Subramaniam M, Abdin E, Vaingankar JA, Chong SA. Obsessive-compulsive disorder: prevalence, correlates, help-seeking and quality of life in a multiracial Asian population. *Soc Psychiatry Psychiatr Epidemiol.* 2012;47(12):2035–2043. doi:10.1007/s00127-012-0507-8
- Mohammadi MR, Ghanizadeh A, Rahgozar M, et al. Prevalence of obsessive-compulsive disorder in Iran. *BMC Psychiatry.* 2004;4(1):1. doi:10.1186/1471-244X-4-2
- Sassano-Higgins SA, Pato MT. Obsessive-compulsive disorder: diagnosis, epidemiology, etiology, and treatment. *Focus.* 2015;13(2):129–141. doi:10.1176/appi.focus.130211
- Eisen JL, Mancebo MA, Pinto A, et al. Impact of obsessive-compulsive disorder on quality of life. *Compr Psychiatry.* 2006;47(4):270–275. doi:10.1016/j.comppsy.2005.11.006
- Coluccia A, Fagiolini A, Ferretti F, Pozza A, Goracci A. Obsessive-compulsive disorder and quality of life outcomes: protocol for a systematic review and meta-analysis of cross-sectional case-control studies. *Epidemiol Biostat Public Health.* 2015;12:2.
- Alarcon RD, Libb JW, Spitzer D. A predictive study of obsessive-compulsive disorder response to clomipramine. *J Clin Psychopharmacol.* 1993;13(3):210–213.
- Ravizza L, Barzegar G, Bellino S, Bogetto F, Maina G. Predictors of drug treatment response in obsessive-compulsive disorder. *J Clin Psychiatry.* 1995;56:368–373.
- Bloch MH, McGuire J, Landeros-Weisenberger A, Leckman JF, Pittenger C. Meta-analysis of the dose-response relationship of SSRI in obsessive-compulsive disorder. *Mol Psychiatry.* 2010;15(8):850–855. doi:10.1038/mp.2009.50
- Van Ameringen M, Simpson W, Patterson B, et al. Pharmacological treatment strategies in obsessive compulsive disorder: a cross-sectional view in nine international OCD centers. *J Psychopharmacol.* 2014;28(6):596–602. doi:10.1177/0269881113517955
- Chen Y, Pedersen LH, Chu WW, Olsen J. Drug exposure side effects from mining pregnancy data. *ACM SIGKDD Explor Newsl.* 2007;9(1):22–29. doi:10.1145/1294301
- Huang Z, Li J, Su H, Watts GS, Chen H. Large-scale regulatory network analysis from microarray data: modified Bayesian network learning and association rule mining. *Decis Support Syst.* 2007;43(4):1207–1225. doi:10.1016/j.dss.2006.02.002
- Karabatak M, Ince MC. An expert system for detection of breast cancer based on association rules and neural network. *Expert Syst Appl.* 2009;36(2):3465–3469. doi:10.1016/j.eswa.2008.02.064
- Shin AM, Lee IH, Lee GH, et al. Diagnostic analysis of patients with essential hypertension using association rule mining. *Health Inform Res.* 2010;16(2):77–81. doi:10.4258/hir.2010.16.2.77
- Liao S-C, Lee I-N. Appropriate medical data categorization for data mining classification techniques. *Med Inform Internet Med.* 2002;27(1):59–67. doi:10.1080/14639230210153749
- Han J, Pei J, Kamber M. *Data Mining: Concepts and Techniques.* 3rd ed. Waltham: Elsevier; 2011.
- Ilayaraja M, Meyyappan T. Mining medical data to identify frequent diseases using Apriori algorithm. Paper presented at: Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference; Periyar University, Tamilnadu, India. 2013.



19. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. Paper presented at: Acm sigmod record; 1993. Washington, D.C., USA – May 25–28, 1993; 22 Issue 2.
20. Flach PA, Lachiche N. Confirmation-guided discovery of first-order rules with Tertius. *Mach Learn*. 2001;42(1):61–95. doi:10.1023/A:1007656703224
21. Scheffer T. Finding association rules that trade support optimally against confidence. *Principles Data Mining Knowledge Discovery*. 2001:424–435.
22. Sharma N, Om H. Extracting significant patterns for oral cancer detection using apriori algorithm. *Intell Inf Manage*. 2014;6(02):30. doi:10.4236/iim.2014.62005
23. Nahar J, Tickle KS, Ali AS, Chen YP. Significant cancer prevention factor extraction: an association rule discovery approach. *J Med Syst*. 2011;35(3):353–367. doi:10.1007/s10916-009-9372-8
24. Nahar J, Imam T, Tickle KS, Chen Y-PP. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Syst Appl*. 2013;40(4):1086–1093. doi:10.1016/j.eswa.2012.08.028
25. Shen CC, Hu LY, Hu YH. Comorbidity study of borderline personality disorder: applying association rule mining to the Taiwan national health insurance research database. *BMC Med Inform Decis Mak*. 2017; 17(1):8. doi:10.1186/s12911-016-0405-1
26. Schweikert C, Li Y, Daya D, Yens D, Torrents M, Hsu DF. Analysis of autism prevalence and neurotoxins using combinatorial fusion and association rule mining. Paper presented at: Bioinformatics and BioEngineering, 2009. BIBE'09. Ninth IEEE International Conference; 2009. Taichung, Taiwan.
27. Hasanpour H, Asadi S, Meibodi RG, et al. A critical appraisal of heterogeneity in obsessive-compulsive disorder using symptom-based clustering analysis. *Asian J Psychiatr*. 2017;28:89–96. doi:10.1016/j.ajp.2017.03.024
28. Hasanpour H, Ghavamizadeh Meibodi R, Navi K, Asadi S. Novel ensemble method for the prediction of response to fluvoxamine treatment of obsessive-compulsive disorder. *Neuropsychiatr Dis Treat*. 2018;14:2027–2038. doi:10.2147/NDT.S173388
29. Tükel R, Bozkurt O, Polat A, Genç A, Atli H. Clinical predictors of response to pharmacotherapy with selective serotonin reuptake inhibitors in obsessive-compulsive disorder. *Psychiatry Clin Neurosci*. 2006;60(4):404–409. doi:10.1111/j.1440-1819.2006.01523.x
30. Eisen JL, Sibrava NJ, Boisseau CL, et al. Five-year course of obsessive-compulsive disorder: predictors of remission and relapse. *J Clin Psychiatry*. 2013;74(3):233–239. doi:10.4088/JCP.12m07657
31. Askland KD, Garnaat S, Sibrava NJ, et al. Prediction of remission in obsessive compulsive disorder using a novel machine learning strategy. *Int J Methods Psychiatr Res*. 2015;24(2):156–169. doi:10.1002/mpr.1463
32. Goodman W, Price L, Rasmussen S, et al. Yale-brown obsessive compulsive scale (Y-BOCS). *Arch Gen Psychiatry*. 1989;46:1006–1011.
33. Rajezi Esfahani S, Motaghipour Y, Kamkari K, Zahireidin A, Janbozorgi M. Reliability and validity of the Persian version of the Yale-Brown Obsessive-Compulsive scale (Y-BOCS). *Iran J Psychiatry Clin Psychol*. 2012;17(4):297–303.
34. Pallanti S, Quercioli L. Treatment-refractory obsessive-compulsive disorder: methodological issues, operational definitions and therapeutic lines. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006;30(3): 400–412. doi:10.1016/j.pnpbp.2005.11.028
35. Pang-Ning T, Steinbach M, Kumar V. *Introduction to Data Mining*. Vol. 74. 2006.
36. Karabatak M, Ince MC. A new feature selection method based on association rules for diagnosis of erythematous-squamous diseases. *Expert Syst Appl*. 2009;36(10):12500–12505. doi:10.1016/j.eswa.2009.04.073
37. Agrawal R, Mannila H, Srikant R, Toivonen H, Verkamo AI. Fast discovery of association rules. *Adv Knowl Discov Data Min*. 1996; 12(1):307–328.
38. Chen Z, Chen G. Building an associative classifier based on fuzzy association rules. *Int J Comput Int Sys*. 2008;1(3):262–273. doi:10.1080/18756891.2008.9727623
39. Liu B, Ma Y, Wong CK. Improving an association rule based classifier. Paper presented at: European Conference on Principles of Data Mining and Knowledge Discovery; 2000. Lecture Notes in Computer Science, vol 1910. Springer, Berlin, Heidelberg.
40. Alcalá-Fdez J, Alcalá R, Herrera F. A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. *IEEE Trans Fuzzy Syst*. 2011;19(5): 857–872. doi:10.1109/TFUZZ.2011.2147794
41. Ozomaro U, Wahlestedt C, Nemeroff CB. Personalized medicine in psychiatry: problems and promises. *BMC Med*. 2013;11(1):132.
42. Costa e Silva JA. Personalized medicine in psychiatry: new technologies and approaches. *Metabolism*. 2013;62:S40–S44. doi:10.1016/j.metabol.2012.08.017.
43. Ozomaro U, Cai G, Kajiwaraya Y, et al. Characterization of SLITRK1 variation in obsessive-compulsive disorder. *PLoS One*. 2013;8(8):e70376. doi:10.1371/journal.pone.0070376
44. Zai G, Brandl EJ, Müller DJ, Richter MA, Kennedy JL. Pharmacogenetics of antidepressant treatment in obsessive-compulsive disorder: an update and implications for clinicians. *Pharmacogenomics*. 2014; 15(8):1147–1157. doi:10.2217/pgs.14.83
45. Crettol S, de Leon J, Hiemke C, Eap CB. Pharmacogenomics in psychiatry: from therapeutic drug monitoring to genomic medicine. *Clin Pharmacol Ther*. 2014;95(3):254–257. doi:10.1038/clpt.2013.221
46. Marcks BA, Weisberg RB, Dyck I, Keller MB. Longitudinal course of obsessive-compulsive disorder in patients with anxiety disorders: a 15-year prospective follow-up study. *Compr Psychiatry*. 2011;52(6): 670–677. doi:10.1016/j.comppsy.2011.01.001
47. Mataix-Cols D, Marks IM, Greist JH, Kobak KA, Baer L. Obsessive-compulsive symptom dimensions as predictors of compliance with and response to behaviour therapy: results from a controlled trial. *Psychother Psychosom*. 2002;71(5):255–262. doi:10.1159/000064812
48. Catapano F, Perris F, Masella M, et al. Obsessive-compulsive disorder: a 3-year prospective follow-up study of patients treated with serotonin reuptake inhibitors: OCD follow-up study. *J Psychiatr Res*. 2006;40(6):502–510. doi:10.1016/j.jpsychires.2005.04.010
49. Fineberg NA, Hengartner MP, Bergbaum C, Gale T, Rössler W, Angst J. Remission of obsessive-compulsive disorders and syndromes; evidence from a prospective community cohort study over 30 years. *Int J Psychiatry Clin Pract*. 2013;17(3):179–187. doi:10.3109/13651501.2013.777744
50. Hazari N, Narayanaswamy JC, Arumugham SS. Predictors of response to serotonin reuptake inhibitors in obsessive-compulsive disorder. *Expert Rev Neurother*. 2016;16(10):1175–1191. doi:10.1080/14737175.2016.1199960
51. Soleimani R, Jalali M, Keshtkar A, Jalali S. Protocol for a systematic review and meta-analysis of lithium, anticonvulsive or atypical antipsychotic drugs for treatment of refractory obsessive-compulsive disorder. *J Med Life*. 2017;10(4):208–215.
52. Piacentino D, Pasquini M, Tarsitani L, et al. Anger is associated with aggressive, contamination, and sexual obsessions in severe OCD outpatients. *Eur Psychiatry*. 2016;33:S239–S240. doi:10.1016/j.eurpsy.2016.01.480
53. McKay D, Sookman D, Neziroglu F, et al. Efficacy of cognitive-behavioral therapy for obsessive-compulsive disorder. *Psychiatry Res*. 2015;225(3):236–246. doi:10.1016/j.psychres.2014.11.058
54. Öst L-G, Havnen A, Hansen B, Kvale G. Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993–2014. *Clin Psychol Rev*. 2015;40:156–169. doi:10.1016/j.cpr.2015.06.003
55. Steketee G, Siev J, Yovel I, Lit K, Wilhelm S. Predictors and moderators of cognitive and behavioral therapy outcomes for OCD: A patient-level mega-analysis of eight sites. *Behav Ther*. 2019;50(1):165–176. doi:10.1016/j.beth.2018.04.004

## Neuropsychiatric Disease and Treatment

Dovepress

### Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS,

and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>