

A Machine Learning-based Approach to Identify Therapists who can Perceive Client-Side Effects in Psychotherapy

Lijun Yao

Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine

Zhiwei Xu

Fudan University

Xudong Zhao

Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine

Yang Chen

Fudan University

Liang Liu

Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine

Xiaoming Fu

University of Göttingen

Fazhan Chen (✉ chenfz@shspdjw.com)

Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine

Research Article

Keywords: Side effects, psychotherapy, therapist, machine learning, artificial intelligence

Posted Date: December 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1150073/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Side effects in psychotherapy are sometimes unavoidable. Therapists play a significant role in the side effects of psychotherapy, but there have been few quantitative studies on the mechanisms by which therapists contribute to them.

Methods: We designed the Psychotherapy Side Effects Questionnaire-Therapist Version (PSEQ-T) and released it online through an official WeChat account, where 530 therapists participated in the cross-sectional analysis. The therapists were classified into groups with and without perceptions of clients' side effects. A number of features were selected to distinguish the therapists by category. Six machine learning-based algorithms were selected and trained by our dataset to build classification models. To make the prediction model interpretable, we leveraged the Shapley Additive exPlanations (SHAP) method to quantify the importance of each feature to the therapist categories.

Results: Our study demonstrated the following: 1) Of the therapists, 316 perceived the side effects of the clients in the ongoing psychotherapy sessions, with a 59.6% incidence of side effects. Among all 7 perception types of the side effects, the most common type was "make the clients or patients feel bad" (49.8%). 2) A random forest-based machine-learning classifier offered the best predictive performance to distinguish the therapists with and without perceptions of clients' side effects, with an F1 score of 0.722 and an AUC value of 0.717. 3) When "therapists' psychological activity" was considered a possible cause of the side effects in psychotherapy by the therapists, it was the most relevant feature for distinguishing the therapist category.

Conclusions: Our study revealed that the therapist's mastery of the limitations of psychotherapy technology and theory, especially the awareness and construction of their own psychological states, was the most important factor in predicting the therapist's perception of the side effects of psychotherapy.

Background

Psychotherapy is a process of actively eliminating or alleviating symptoms through the therapeutic relationship and interaction between the psychotherapist and the client and helping the client improve his or her personality, adapt to society, and promote rehabilitation [1]. Psychotherapy, as an effective method of medical treatment [2], may also result in side effects or even cause harm [3]. Many previous studies have confirmed that side effects in psychotherapy were common. The National Audit of Psychological Therapies (NAPT) conducted in England and Wales reported that 5.2% of patients had long-lasting negative effects from psychological treatment [4]. In psychotherapy, 38.5% of patients with depression ($n=135$) exhibited at least one adverse effect [5]. Our recent study reported that the incidence of side effects in psychotherapy was 31.1% (115/370), and the most common side effect was that participants "feel bad in psychotherapy" (24.6%) [6]. Among young adults with a history of psychotherapy, the incidence of psychotherapeutic side effects was approximately 41%, and too long of a treatment time

and the deterioration of the existing symptoms were the strongest predictors of poor therapeutic effects [7].

Unfortunately, the side effects of psychotherapy have not attracted much attention. Only 21% of these randomized, controlled trials of psychotherapy monitored harm to patients, and only 3% of the trials described adverse events [8]. Many therapists or clinicians fail to identify and manage these side effects, mainly due to an insufficient awareness of the side effects in psychotherapy. Boisvert et al. [9, 10] speculated that only 10% ~ 28% of therapists were aware of the worsening effect of psychotherapy. In another study on clinicians identifying the negative outcomes of psychotherapy, only 21% of the negative outcomes were identified effectively [11].

Sensitivity to the side effects in psychotherapy is a useful indicator of good therapists and can significantly improve the quality of treatments [12]. In the past decade, clinicians and researchers have gradually realized that the psychological treatment results among the same patients have deteriorated. However, clinical training rarely contains information about the side effects of psychotherapy [13]. Thus, more efforts need to be made regarding the side effects in psychotherapy from the perspective of therapists to improve both research and clinical understanding in terms of identifying and avoiding the side effects. Moreover, therapist factors are closely related to the effect of psychotherapy. The National Institute of Mental Health Treatment of Depression Collaborative Research Program [14] indicated that approximately 8% of the outcome variance in psychotherapy was attributable to the therapist. Another study showed that approximately 8% of the total variance and approximately 17% of the variance in the degree of patient improvement may be attributed to the therapist [15].

Similarly, therapist factors may affect the side effects of psychotherapy. The inappropriate narratives of therapists could destroy the outcome of psychotherapy and the therapeutic alliance, particularly therapists' controlling and challenging statements [16]. An NAPT study shows that, when patients' treatment preferences were not satisfied, patients experience more negative effects in psychological treatment [17]. The treatment preferences included the characteristics of the therapist. In our recent study, the mental state of the psychotherapist was the most relevant feature in predicting whether clients would experience side effects in psychotherapy [6]. Interestingly, when the outcome of the psychotherapy was regularly monitored and fed back to the psychotherapist, the deterioration effect of the psychotherapy was significantly reduced [18]. Early identification of treatment failure and problem-solving strategies by the psychotherapist in routine practice would also significantly improve the effect of psychotherapy [19].

In summary, therapists play a significant role in the occurrence of the side effects of psychotherapy, but the questions remain, how can therapists perceive clients' side effects, and which factors determine the degree of their perceptions. Since most of the studies on the side effects in psychotherapy have been based on the perspectives of the clients or patients, these mechanisms remain unclear.

Machine learning (ML) is a form of artificial intelligence that automatically learns from data and builds classification or predictive model. It can be used to predict outcomes for new data. In psychiatry, ML has been applied in the areas of diagnosis, progression, treatment prediction, and detection of potential

biomarkers of mental disorders [20]. Studies used electronic health records, brain imaging, cognitive testing, rating scales, genetics, electrophysiology, smartphone and social media data to predict, classify, or subgroup mental health problems including schizophrenia, depression, and suicide, et.al. [21, 22]. ML also may have a significant impact on mechanism and process, training and feedback, and technology mediated treatment modalities of psychotherapy [23]. But the applications of ML in psychotherapy related studies are still very limited [22]. Goldbery et al. used ML to predict the client-rated therapeutic alliance from session linguistic content during the psychotherapy [24]. Their results modestly predicted alliance ratings and suggested ML may be a useful tool for examining alliance in future studies. In one of our recent works, supervised ML was adopted to build a prediction model that can identify clients who may have consulting side effects in psychotherapy based on information from the psychotherapy [6]. By comparing six models based on different ML algorithms, Random Forest-based model provided the best predictive performance for psychotherapy outcomes. In the present study, we focused on the therapists in psychotherapy, and further used ML techniques to identify therapists who are aware of clients' side effects in psychotherapy.

In this study, combined with a self-designed *Questionnaire*, statistics, ML, and SHapley Additive exPlanations (SHAP) method [25], we aim to investigate the types of psychotherapeutic side effects perceived by the therapist, identify the predictive factors that determine whether the therapists can perceive clients' side effects in psychotherapy, and use different ML algorithms to establish the best model that can distinguish between therapists who can and cannot perceive the side effects of psychotherapy. The findings of this research is to develop a scientific framework to improve the therapists' ability to identify and manage the side effects of psychotherapy, and to provide a more definite basis for the professionalization of psychotherapy.

Methods

Psychotherapy Side Effects Questionnaire-Therapist Version (PSEQ-T)

Based on previous research results [26, 27], we designed the *Psychotherapy Side Effects Questionnaire-Therapist Version (PSEQ-T)*. In the PSEQ-T, side effects in psychotherapy were defined as unwanted events that the therapist perceived during the course of psychotherapy that were inconsistent with the expected goals and had a negative impact on the clients. The side effects in psychotherapy were assessed according to the respondents' answers to the question, "Has the psychotherapy you are currently conducting caused side effects or harm to your clients or patients?". The answer "yes" was considered as the therapists' ability to perceive the side effects; otherwise, there was no indication of the side effects. Seven questions in the PSEQ-T were designed to assess these side effects perceived by therapists from three dimensions: symptoms, relationships, and social function (Table 2). Three questions were designed to assess the presence of new symptoms, namely, negative emotions ("Has your psychotherapy made your clients or patients feel bad?"), bad behaviors ("Has your psychotherapy made

your clients or patients behave badly?"'), and physical discomfort ("Has your psychotherapy caused discomfort in your clients or patients' physical health?"). One question was used to assess the original problem ("Has your psychotherapy worsened the symptoms of your clients or patients?"). Two questions were used to assess negative changes in family relationships ("Has your psychotherapy strained your clients' or patients' family relationship?") and interpersonal relationships ("Has your psychotherapy strained the relationship outside your clients' or patients' family?"). The last question was used to assess negative changes in social function ("Has your psychotherapy worsened your clients' or patients' working conditions?").

To predict the perception of the side effects by psychotherapists in psychotherapy, we collected the following features of each participant in the PSEQ-T: demographics, clinical practice information, occupational information, and the possible causes of side effects in psychotherapy. Table 1 lists the detailed information for each feature.

The questionnaire was sent to ten examiners for content revision and then revised based on the feedback to form the final version of the PSEQ-T. The Cronbach's α of the PSEQ-T was 0.667, indicating acceptable internal consistency for this survey[28–31].

The procedure of data collection

The questionnaire was edited and released on the WeChat platform on February 11, 2019. Participants read and decided whether to complete the questionnaire according to the inclusion criteria and chose informed consent before submitting the questionnaire. The questionnaire was anonymous. Participants completed the questionnaire online using the mobile phone-based interface provided by WeChat. The completion time for each questionnaire was approximately three to five minutes. The information collected by the questionnaire was automatically populated into an Excel form. Data collection ceased on June 6, 2019.

Participants' entry requirements

Participants were enrolled through the online questionnaire (PSEQ-T) published on the official WeChat account from February 11 to June 6, 2019. Inclusion criteria were that the participants 1) carried out at least one session of psychotherapy in the last month, 2) had a licensed practice qualification of psychotherapy issued by the government, 3) were 18–70 years old, and 4) gave informed consent. Exclusion criteria included the participants' having 1) serious mental disorders or physical illnesses, 2) ethical faults, or 3) disagreements with the public release of the research data.

Machine Learning-based classification of therapists

In this study, we used supervised machine-learning technologies to build a classifier that could predict whether the therapist could perceive client's side effects in psychotherapy. In our dataset, we chose therapists "with perceptions of clients' side effects" category as the positive class and therapists "without perceptions of clients' side effects" category as the negative class. All the features used to build the

classifier are listed in Table 1. The workflow of the raw data preprocessing, machine-learning algorithm training and model performance evaluation are described in Fig. 1.

The final dataset included 316 therapists who reported having perceptions of clients' side effects in psychotherapy and 214 therapists who did not. To solve the unbalanced sample problem, we oversampled the minority type to 316 by the SMOTE technique [32]. Then, we randomly split the entire, balanced dataset into a training and validation dataset and a test dataset. We used 70% of the data for training and validation and the remaining 30% of the data for testing. We further used the 5-fold cross-validation method, in which the training dataset was randomly divided into 5 subsets with equal sample sizes. The cross-validation process was repeated 5 times. In each round, 4 subsets were used for training, with the remaining one was retained as validation data used to evaluate the model.

We have selected traditional ML algorithms that are widely used, such as random forest [33], logistic regression [34], support-vector machine (SVM) [35], AdaBoost [36], and algorithms with excellent predictive effects developed in recent years, such as XGBoost [37] and CatBoost [38], to train our data. To achieve the best predictive performance for each algorithm, an optimal set of parameters needs to be determined by the following three steps. First, based on the training dataset, we applied a grid search to go through the parameter space. We selected a finite set of values for each parameter to form the parameter space, and the grid search iterated through a set of parameter combinations. Second, for each combination, we evaluated the predictive performance using the F1 score. Finally, we recorded the parameters leading to the maximum F1 score based on the training and validation dataset. scikit-learn, a Python-based machine-learning library, was used to train and evaluate the classification models [39].

For the model evaluation, we used precision, recall, the F1 score and the area under the ROC curve (AUC) value to evaluate the predictive performance of the trained models [40]. Precision is the fraction of the therapists classified by the model as "with perceptions of clients' side effects" who perceived clients' side effects. Recall is the fraction of the therapists "with perceptions of clients' side effects" who were correctly identified by the model. The F1 score is the harmonic mean of precision and recall, calculated as follows:

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

An F1 score reaches its best value at 1 and worst value at 0. A higher value of the F1 score indicates a better overall predictive performance of a classifier. AUC is another important evaluation metric for examining the performance of a classifier and denotes the probability that a classifier will rank a random positive instance higher than a randomly chosen negative instance. The value of AUC is also between 0

and 1. In our study, when the AUC value was higher, the model was better able to distinguish the therapists with or without perceptions of clients' side effects in psychotherapy.

Analysis of the Feature Importance and Statistics

To make the prediction model interpretable, we used the SHAP method to quantify the importance of each feature in the prediction model. SHAP is a representative method to explain the predictions of supervised ML-based classifiers. We applied this method to quantify the importance of each feature to the categories of the therapists (with or without perceptions of clients' side effects). The aim of the SHAP model is to explain the prediction of a selected instance by calculating the contribution of each feature to the prediction. For each feature, we used the metric mean ($|SHAP\ value|$), i.e., the average value of the absolute values of the SHAP values of all the therapists, to obtain the value of the feature importance. When the $|SHAP\ value|$ is higher, the feature makes a greater contribution to the prediction model. We applied the Python programming language to implement the calculation. The P-values in Table 1 were calculated by the chi-square test. $p < 0.05$ was considered statistically significant.

Results

Participants' demographics

A total of 570 therapists completed the online questionnaire PSEQ-T. Seven participants were unwilling to share their information with the public, and 33 participants were excluded from further analysis because of irregular data input. As a result, 530 participants were included in the final dataset (Fig. 1(1)). Each instance has 12 main features, in which each feature was either numerical or categorical. The detailed number and percentage of each feature are shown in Table 1. The mean age of the therapists included in the analysis was 41.3 years ($SD=9.00$ years), and the mean age of the therapists with perceptions of clients' side effects was 40.70 years ($SD=8.71$ years), which was slightly younger than the therapists without perceptions of clients' side effects (mean=42.28 years, $SD=9.31$ years). The average number of years working as a therapist was 7.81 years ($SD=5.89$ years) for those with perceptions of clients' side effects, while the average number of years working for the therapists without perceptions of clients' side effects was 6.99 years ($SD=5.18$ years), which was also statistically significant.

Types of client-side effects perceived by the therapist

In our study, 316 therapists perceived the side effects of the clients in their current psychotherapeutic offerings, and the incidence was 59.6%. Among all 7 types of side effects, the most common was "made the clients or patients feel bad" (49.8%), which was far more common than the second most common side effect, "strained the clients' or patients' family relationship" (18.3%). The least common side effect was "worsened the clients' or patients' working conditions" (5.8%). The detailed types and incidences of each side effect perceived by the therapists are described in Table 2.

Distinguishing between the therapists with and without perceptions of clients' side effects in psychotherapy by ML

In this part, we used supervised machine-learning algorithms to distinguish between the perceptive and nonperceptive therapists. Six representative machine-learning algorithms—i.e., random forest, XGBoost, CatBoost, logistic regression, AdaBoost and SVM—were selected to build the classification models. Then, we compared each classification model's predictive performance to achieve the best classifier. Our results showed that the F1 scores of the six models—random forest, XGBoost, CatBoost, logistic regression, AdaBoost and SVM—were 0.722, 0.681, 0.681, 0.680, 0.653 and 0.647, respectively (Table 3). The models' precision and recall are also described in Table 3. The AUC values of the six models—random forest, XGBoost, CatBoost, logistic regression, AdaBoost and SVM—were 0.717, 0.689, 0.694, 0.675, 0.653 and 0.629, respectively. By comparing the predictive performance of these six classifiers, the random forest-based classifier achieved the highest F1 score of 0.722 and AUC value of 0.717, showing the best performance distinguishing between the therapists with or without perceptions of clients' side effects in psychotherapy.

Important features distinguishing the therapists with or without perceptions of clients' side effects in psychotherapy

Many factors affect a therapist's perception of clients' side effects. The PESQ-T included 12 main features, which are listed in Table 1. Some main features—including subfeatures, such as "practice qualification"—were further divided into 4 educational system subfeatures: national second/third level "psychological counselor", "psychotherapist", "psychiatrist" and "psychologist" (Table 1). To drive the classification model, 25 detailed features were included. Next, we quantified each feature's |SHAP value| in the trained random forest-based classifier (Fig. 2). Based on each feature's |SHAP value|, "therapists' psychological activity" ranked 1st among all the analyzed features, thus contributing most to distinguishing the therapists with or without perceptions of clients' side effects in our classifier.

Next, to visualize the difference between the two groups of therapists, we compared the top-six ranked features based on their |SHAP values| (Fig. 3). The perceptive therapists were more likely to believe that 1) the therapist's psychological activity would affect clients' side effects (Fig. 3(1)), 2) the characteristics of psychotherapy would cause clients' side effects (Fig. 3(2)), and 3) the improper use of psychotherapeutic techniques would cause clients' side effects (Fig. 3(3)). When the psychotherapist was younger, he or she was more likely to perceive clients' side effects (Fig. 3(4)). Workplaces also affected client's side effects, and the therapists working in hospitals were more likely to perceive them (Fig. 3(5)). Among the male therapists, the percentage of the perceptive therapists was higher than that of the nonperceptive ones, while for the female therapists, the percentage of the perceptive therapists was slightly lower than that of the nonperceptive ones (Fig. 3(6)). Overall, we found that there were significant differences between the two groups of therapists in terms of the therapist's psychological activity, the

characteristics and improper use of psychotherapy techniques, the age of the therapists, the working places where the psychotherapy occurred and the gender of the therapists.

Discussion

In the present study, we used a machine-learning approach to establish prediction models and analyzed the related influencing factors of clients' side effects perceived by therapists based on a primary online survey in China. The results demonstrated that 59.6% of the therapists reported some side effects in the psychotherapy that they were carrying out, and the most common side effect perceived by the therapists was "made the clients or patients feel bad" (49.8%). The random forest-based classifier provided the best predictive performance in distinguishing between the therapists with and without perceptions of clients' side effects, with an F1 score of 0.722 and an AUC value of 0.717. The SHAP analysis further showed that "therapists' psychological activity" was the most relevant feature distinguishing the two categories of therapists.

The identification and management of side effects by therapists is the key to performing professional psychotherapy. A limited sample study ($n = 73$) showed that, although 94.5% of clinicians agreed that negative effects of psychological treatments exist and 75% of clinicians described that they had clinical experience with negative effects, only 8 (11%) of clinicians had received information about negative effects during their basic clinical training [41]. In our study, the accuracy of identifying side effects by therapists was significantly higher than that of previous studies [9–11], which indicates that the sensitivity of therapists to side effects is increasing. In China, the National Health Commission promulgated the Code of Psychotherapy in 2013, and the Chinese Psychological Society formulated the Code of Ethics for Clinical and Counseling Psychology (2nd Edition) in 2019. The professional training of psychotherapists is increasingly becoming systematic and standardized, but there is no information about the side effects of psychotherapy in the process of training therapists in China [27] or even worldwide [13]. Licensed therapists may be able to identify and manage side effects in psychotherapy to ensure the professionalism and standardization of the psychotherapy industry. This is also the primary purpose of this study, but much work remains to be done.

In our investigation, "made the clients or patients feel bad" (49.8%) was the most common perception of the side effects reported by the therapists, which is consistent with previous studies [6, 41–43]. Qualitative research focusing on the therapists' perspectives on the negative effects of psychological treatments demonstrated that the characteristics of the negative effects included "short-term negative effects", "no treatment effect", "deterioration", "dependency" and "impact on other life domains" [41]. The negative effect "deterioration" was a common side effect of psychological treatments. Among the few quantitative studies on therapists' experiences, 100 cognitive behavior therapists were interviewed for side effects in their current outpatients [42]. These therapists reported that the most frequent side effects were "negative wellbeing/distress" (27%), "worsening of symptoms" (9%) and "strains in family relations" (6%). Similar outcomes of psychotherapy have been found in other studies based on clients' experiences. A survey conducted by our team at nearly the same time as this study revealed that the most common

side effect experienced by the clients was “feel bad in psychotherapy” (24.6%) [6]. These negative emotions caused by psychotherapy might last for a long time. After an average of 3.76 years (outpatients) and 9 months (inpatients) of psychotherapy, the negative emotions elicited by the question “I was hurt by what the therapist said to me” were still the most frequently recognized side effects in outpatients (3.6%) and inpatients (20.3%) [43]. In the PSEQ-T, “feel bad” referred to a negative emotion experienced by participants in psychotherapy, such as anxiety, tension, sadness, and anger. With the emergence of new symptoms in the process of psychotherapy, these negative feelings may be related to therapists, patients and the therapeutic alliance [44]. If these side effects are not identified and managed well by therapists, psychotherapy may induce harm.

The side effects of psychotherapy have adverse effects on patients, and their occurrence should be minimized during psychotherapy. The question is how to identify patients with potential side effects by psychotherapy and therapists who can perceive the side effects of psychotherapy. In our previous work, we used ML to identify clients who may have side effects from psychotherapy. The F1 value of the model was 0.797, and the AUC was 0.804, indicating that the model had a good predictive effect [6]. Therapists can use this information to provide more suitable psychotherapy for specific patients, thereby improving the outcomes of psychotherapy. Similarly meaningful would be if we could use ML to distinguish between therapists who can and cannot perceive the side effects of psychotherapy. In this study, we used a self-compiled questionnaire to extract features from three dimensions (symptoms, relationships, and social functions). A random-forest algorithm-based model achieved an F1 score of 0.722 and an AUC value of 0.717, which demonstrated that the model can be used to distinguish among therapists with different perceptive abilities. With the information the model provides, on the one hand, our results can screen therapists and find those with better awareness of the side effects of patient psychotherapy; on the other hand, therapists with relatively poor perception can be provided the relevant training to improve the professionalism of psychotherapy.

Furthermore, we calculated the |SHAP value| of each feature in the random forest-based model. In this model, the therapists believed that their “psychological activity may be the cause of the side effects in psychotherapy”, and these therapists were the most sensitive to the side effects. Some studies based on clients’ experiences also found that the characteristics of therapists can predict the side effects in psychotherapy [6, 16, 17]. Therapist factors mediate the outcomes of psychotherapy mainly through therapeutic alliances. On average, therapists who developed stronger alliances with their patients achieved better therapeutic results [45]. According to Jennifer, Jonas and Sylke [16], the destructive therapeutic alliance was particularly evident in the therapists’ mental state performance, such as controlling and challenging statements. A good therapeutic alliance values a supportive and reinforcing context, as in a context in which fewer stressful interventions occur and the therapeutic relationship is comfortable. The therapist’s mental activity had an effect on the clients through the therapeutic relationship, which was the most critical factor related to psychotherapeutic side effects [6]. Combined with previous research results, the present study suggested that the therapist’s introspection and management of their own psychological activity will help the therapist to identify and monitor the side

effects in psychotherapy, which could significantly reduce the deterioration effect [18] and improve the effect of psychotherapy [19].

In this study, two other important predictors were the therapist's perception of the factors "characteristics of psychotherapy techniques were a possible cause of the side effects in psychotherapy" and "improper use of psychotherapy techniques was a possible cause of the side effects in psychotherapy". The essence of psychotherapy is to help people learn who they are, access their emotional basics, hold their feelings intact and be able to think even under the heaviest interpersonal pressure, which is the first and main therapeutic goal [46]. In addition to the factors of therapists and clients, the theory and technology of psychotherapy are key to the effect of treatment. Parry and her colleagues [4] believe that "using an inappropriate therapeutic method or errors in delivering a recommended therapy" might be risk factors for negative outcomes and possible mechanisms for harmful psychological therapies. Some studies have shown that the theoretical orientation of psychotherapy significantly affects the occurrence of side effects [4, 6]. For example, patients who experienced a poor therapeutic relationship, a high degree of dependency or isolation and a high burden through psychotherapy were more frequently treated in psychodynamic therapies [47]. Although such a therapeutic process is effective, it places great pressure on patients. Furthermore, an inappropriate use of psychotherapy techniques may lead to malpractice and unethical behavior in psychotherapy. At least one case of malpractice and unethical behavior was reported by 28.8% of inpatients and 7.1% of outpatients in psychotherapy [43]. Therefore, this study indicates that understanding the limitations of psychotherapy theory and technology will help therapists identify the side effects of psychotherapy.

To the best of our knowledge, this was the first study to predict the potential side effects perceived by therapists in psychotherapy using ML. The machine-learning approaches described in this study are sufficiently accurate and meaningful to be integrated into clinical psychiatry. Our research provides new methods that can be used to distinguish among therapists with different perceptions of clients' side effects and suggests important predictive factors that affect the therapist's perception. The results of this study demonstrate verify a possible technical path that can enhance the sensitivity and recognition of therapists to the side effects in psychotherapy. In this path, the stability and health of the therapist's psychological state and professional mastery, especially the mastery of the limitations of treatment theory and technology, may help to increase the recognition and management of side effects.

Limitations

This study constructed a fairly accurate model to predict the therapists who can perceive side effects in psychotherapy. However, there are still some limitations: 1) The evaluation tool PSEQ-T is a simple, self-designed questionnaire, its validity and reliability of side effects may be improved based on further use and feedback in future therapy sessions; 2) This study is a cross-sectional study, and the number and representativeness of the research samples still needs to be improved; 3) Some important factors in psychotherapy, such as the treatment dosage and the therapists characteristics, were not included in the

evaluation; 4) The present study did not cover which mental states of therapists were more likely to cause the side effects of psychotherapy. This issue would be interesting to explore in our future research.

Conclusions

In this preliminary and exploratory study, we concluded the following: 1) the therapists were sensitive to the side effects of psychotherapy, and the negative emotional experiences of the clients were commonly perceived by the therapists; 2) our random forest–based machine-learning model offered the best predictive performance of the perception of clients' side effects by the therapists, with an F1 score of 0.722 and a AUC value of 0.717; and 3) the therapist's mastery of the limitations of psychotherapy technology and theory, especially the awareness and construction of their own psychological state, were the most important factors in predicting the therapist's perception of the side effects of psychotherapy. In conclusion, our research results can provide clinicians, therapists and clients with important information to ensure that the side effects of psychotherapy are minimized or avoided in future clinical practice. They also provide an important reference for government departments or academic organizations in formulating psychotherapy service policies.

Abbreviations

PSEQ-T: Psychotherapy Side Effects Questionnaire-Therapist Version; SHAP: Shapley Additive exPlanations; ML: machine learning; NAPT: The National Audit of Psychological Therapies; SVM: support-vector machine; AUC: the area under the ROC curve.

Declarations

Acknowledgments

The authors would like to thank Yuhong Yao, Congcong Ge and Yunhan Zhao for their assistance in the data entry and questionnaire collection tasks. Moreover, the authors greatly appreciate the contributions of the participants.

Author Contributions

LY and FC made substantial contributions to the conception, design, analysis, and drafting of the manuscript, ensuring that the work was appropriately investigated and resolved. XZ, LL and XF contributed to the study design and critical review of the manuscript for important intellectual content. ZX, YC and XF were involved in implementing the machine learning algorithms and statistical analysis. All authors read and approved the final version of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (Grant Number: 81771464), Training Plan of Health System Academic Leader of Shanghai Pudong Municipality Health Commission (Grant Number: PWRd2019-08), Outstanding Clinical Discipline Project of Shanghai Pudong (Grant Number: PWYgy2018-10), Special Clinical Research Project of Shanghai Municipality Health Commission (Grant Number: 202040475), and Open Research Program of Shanghai Key Lab of Intelligent Information Processing (Grant Number: IIPL201911).

Data availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The project involving human participants was reviewed and approved by the Ethics Committee of Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine (Approval Number: 2019tjdx9). The participants completed the questionnaire anonymously and agreed to informed consent when submitting the questionnaire data. All methods performed in this study was in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Shanghai Pudong New Area Mental Health Center, Clinical Research Center for Mental Disorders, Tongji University School of Medicine, Shanghai 200124, P.R. China. ² School of Computer Science, Fudan University, Shanghai 200434, P.R. China. ³ Institute of Computer Science, University of Göttingen, Göttingen 37077, Germany.

References

1. Kreutzer J, DeLuca J, Caplan B. Encyclopedia of Clinical Neuropsychology. Cham: Springer. 2018. ISBN 978-0-387-79949-0.
2. Dragioti E, Karathanos V, Gerdle B, Evangelou E. Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. *Acta Psychiatr Scand*. 2017; 136(3):236–246. <https://doi.org/10.1111/acps.12713>.
3. Lilienfeld SO. Psychological Treatments That Cause Harm. *Perspect Psychol Sci*. 2007; 2(1):53–70. <https://doi.org/10.1111/j.1745-6916.2007.00029.x>.
4. Crawford MJ, Thana L, Farquharson L, Palmer L, Hancock E, Bassett P, Clarke J, Parry GD. Patient experience of negative effects of psychological treatment: results of a national survey. *Brit J Psychiat*. 2016; 208(3):260–265. <https://doi.org/10.1192/bjp.bp.114.162628>.
5. Peth J, Jelinek L, Nestoriuc Y, Moritz S. Adverse Effects of Psychotherapy in Depressed Patients - First Application of the Positive and Negative Effects of Psychotherapy Scale (PANEPS). *Psychother Psychosom Med Psychol*. 2018; 68(9-10):391–398. <https://doi.org/10.1055/s-0044-101952>.
6. Yao L, Zhao X, Xu Z, Chen Y, Liu L, Feng Q, Chen F. Influencing Factors and Machine Learning-Based Prediction of Side Effects in Psychotherapy. *Front Psychiatry*. 2020; 11:537442. <https://doi.org/10.3389/fpsyg.2020.537442>.
7. Lorenz TK. Predictors and impact of psychotherapy side effects in young adults. *Couns Psychother Res*. 2021; 21(1):237–243. <https://doi.org/10.1002/capr.12356>.
8. Jonsson U, Alaie I, Parling T, Arnberg FK. Reporting of harms in randomized controlled trials of psychological interventions for mental and behavioral disorders: a review of current practice. *Contemp Clin Trials*. 2014; 38(1):1–8. <https://doi.org/10.1016/j.cct.2014.02.005>.
9. Boisvert CM, Faust D. Leading researchers' consensus on psychotherapy research findings: Implications for the teaching and conduct of psychotherapy. *Prof Psychol Res Pr*. 2003; 34(5):508–513. <https://doi.org/10.1037/0735-7028.34.5.508>.
10. Boisvert CM, Faust D. Practicing psychologists' knowledge of general psychotherapy research findings: Implications for science-practice relations. *Prof Psychol Res Pr*. 2006; 37(6):708–716. <https://doi.org/10.1037/0735-7028.37.6.708>.
11. Hatfield D, McCullough L, Frantz SH, Krieger K. Do we know when our clients get worse? an investigation of therapists' ability to detect negative client change. *Clin Psychol Psychother*. 2010; 17(1):25–32. <https://doi.org/10.1002/cpp.656>.
12. Linden M. How to Define, Find and Classify Side Effects in Psychotherapy: From Unwanted Events to Adverse Treatment Reactions. *Clin Psychol Psychot*. 2013; 20(4):286–296. <https://doi.org/10.1002/cpp.1765>.
13. Rozental A, Castonguay L, Dimidjian S, Lambert M, Shafran R, Andersson G, Carlbring P. Negative effects in psychotherapy: commentary and recommendations for future research and clinical practice. *BJPsych Open*. 2018; 4(4):307–312. <https://doi.org/10.1192/bjo.2018.42>.
14. Kim DM, Wampold BE, Bolt DM. Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program

- data. *Psychotherapy Research*. 2006; 16(2):161–172. <https://doi.org/10.1080/10503300500264911>.
15. Lutz W, Leon SC, Martinovich Z, Lyons JS, Stiles WB. Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*. 2007; 54(1):32–39. <https://doi.org/10.1037/0022-0167.54.1.32>.
16. Kadur J, Ludemann J, Andreas S. Effects of the therapist's statements on the patient's outcome and the therapeutic alliance: A systematic review. *Clin Psychol Psychother*. 2020; 27(2):168–178. <https://doi.org/10.1002/cpp.2416>.
17. Williams R, Farquharson L, Palmer L, Bassett P, Clarke J, Clark DM, Crawford MJ. Patient preference in psychological treatment and associations with self-reported outcome: national cross-sectional survey in England and Wales. *BMC Psychiatry*. 2016; 16:4. <https://doi.org/10.1186/s12888-015-0702-8>.
18. Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, Goates M. Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clin Psychol Psychot*. 2002; 9(2):91–103. <https://doi.org/10.1002/cpp.324>.
19. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: The use of early identification of treatment failure and problem-solving strategies in routine practice. *Journal of Counseling Psychology*. 2003; 50(1):59–68. <https://doi.org/10.1037/0022-0167.50.1.59>.
20. Aafjes-van Doorn K, Kamsteeg C, Bate J, Aafjes M. A scoping review of machine learning in psychotherapy research. *Psychother Res*. 2021; 31(1):92–116. <https://doi.org/10.1080/10503307.2020.1808729>.
21. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, Jeste DV. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr Psychiatry Rep*. 2019; 21(11):116. <https://doi.org/10.1007/s11920-019-1094-0>.
22. Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, Fokkema M, Cohen Z, Belgrave D, DeRubeis R, Iniesta R *et al*. The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*. 2021; 20(2):154–170. <https://doi.org/10.1002/wps.20882>.
23. Imel ZE, Caperton DD, Tanana M, Atkins DC. Technology-enhanced human interaction in psychotherapy. *J Couns Psychol*. 2017; 64(4):385–393. <https://doi.org/10.1037/cou0000213>.
24. Goldberg SB, Flemotomos N, Martinez VR, Tanana MJ, Kuo PB, Pace BT, Villatte JL, Georgiou PG, Van Epps J, Imel ZE *et al*. Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *J Couns Psychol*. 2020; 67(4):438–448. <https://doi.org/10.1037/cou0000382>.
25. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems 30 (Nips 2017)*. 2017; 30. <https://dl.acm.org/doi/10.5555/3295222.3295230>.
26. Linden M, Schermuly-Haupt ML. Definition, assessment and rate of psychotherapy side effects. *World Psychiatry*. 2014; 13(3):306–309. <https://doi.org/10.1002/wps.20153>.

27. Chen F, Zhao X. Side Effects of Psychotherapy. *Chin Ment Health J.* 2017; 31(1):72–76.
28. Fleming R. An environmental audit tool suitable for use in homelike facilities for people with dementia. *Australas J Ageing.* 2011; 30(3):108–112. <https://doi.org/10.1111/j.1741-6612.2010.00444.x>.
29. Setbon M, Raude J. Factors in vaccination intention against the pandemic influenza A/H1N1. *Eur J Public Health.* 2010; 20(5):490–494. <https://doi.org/10.1093/eurpub/ckq054>.
30. Gallais B, Gagnon C, Forgues G, Cote I, Laberge L. Further evidence for the reliability and validity of the Fatigue and Daytime Sleepiness Scale. *J Neurol Sci.* 2017; 375:23–26. <https://doi.org/10.1016/j.jns.2017.01.032>.
31. Churchill GA. A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research.* 1979; 16(1):64–73. <https://doi.org/10.2307/3150876>.
32. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res.* 2002; 16:321–357. <https://doi.org/10.1613/jair.953>.
33. Breiman L. Random forests. *Mach Learn.* 2001; 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>.
34. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform.* 2002; 35(5-6):352–359. [https://doi.org/10.1016/S1532-0464\(03\)00034-0](https://doi.org/10.1016/S1532-0464(03)00034-0).
35. Hearst MA. Support vector machines. *Ieee Intell Syst App.* 1998; 13(4):18–21. <https://doi.org/10.1109/5254.708428>.
36. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci.* 1997; 55(1):119–139. <https://doi.org/10.1006/jcss.1997.1504>.
37. Chen TQ, Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 2016:785–794. <https://doi.org/10.1145/2939672.2939785>.
38. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. *Adv Neur In.* 2018; 31. <https://dl.acm.org/doi/10.5555/3327757.3327770>.
39. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V *et al.* Scikit-learn: Machine Learning in Python. *J Mach Learn Res.* 2011; 12:2825–2830. <http://doi.org/10.5555/1953048.2078195>.
40. Fawcett T. An introduction to ROC analysis. *Pattern Recogn Lett.* 2006; 27(8):861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>.
41. Bystedt S, Rozental A, Andersson G, Boettcher J, Carlbring P. Clinicians' Perspectives on Negative Effects of Psychological Treatments. *Cogn Behav Therapy.* 2014; 43(4):319–331. <https://doi.org/10.1080/16506073.2014.939593>.
42. Schermuly-Haupt ML, Linden M, Rush AJ. Unwanted Events and Side Effects in Cognitive Behavior Therapy. *Cognitive Ther Res.* 2018; 42(3):219–229. <https://doi.org/10.1007/s10608-018-9904-y>.

43. Gerke L, Meyrose AK, Ladwig I, Rief W, Nestoriuc Y. Frequencies and Predictors of Negative Effects in Routine Inpatient and Outpatient Psychotherapy: Two Observational Studies. *Front Psychol.* 2020; 11. <https://doi.org/10.3389/fpsyg.2020.02144>.
44. Parry GD, Crawford MJ, Duggan C. Iatrogenic harm from psychological therapies - time to move on. *Brit J Psychiat.* 2016; 208(3):210–212. <https://doi.org/10.1192/bjp.bp.115.163618>.
45. Fluckiger C, Del Re AC, Wampold BE, Horvath AO. The Alliance in Adult Psychotherapy: A Meta-Analytic Synthesis. *Psychotherapy.* 2018; 55(4):316–340. <https://doi.org/10.1037/pst0000172>.
46. Bugliani A. Wrongness: Social Side-Effects in Psychotherapy. *Psychoanal Inq.* 2020; 40(4):253–261. <https://doi.org/10.1080/07351690.2020.1742539>.
47. Leitner A, Märkens M, Koschier A, Gerlich K, Liegl G, Hinterwallner H, Schnyder U. Patients' Perceptions of Risky Developments During Psychotherapy. *Journal of Contemporary Psychotherapy.* 2013; 43(2):95–105. <https://doi.org/10.1007/s10879-012-9215-7>.

Tables

Table 1

Features of therapists involved in the study

Features	With perception of the side effects (n=316)	Without perception of the side effects (n=214)	Overall (%)	P-value
Gender				
<i>Male</i>	87 (27.5%)	33 (15.4%)	120 (22.6%)	0.304
<i>Female</i>	229 (72.5%)	181 (84.6%)	410 (77.4%)	
Age				0.006**
<=29	28 (8.9%)	17 (7.9%)	45 (8.5%)	
30-49	232 (73.4%)	151 (70.6%)	383 (72.3%)	
>=50	56 (17.7%)	46 (21.5%)	102 (19.2%)	
Marriage status				0.980
<i>Single</i>	27 (8.5%)	17 (7.9%)	44 (8.3%)	
<i>Single with partner</i>	16 (5.1%)	11 (5.1%)		
<i>Married</i>	257 (81.3%)	177 (82.7%)	27 (5.1%)	
<i>Divorced, separated or widowed</i>	16 (5.0%)	9 (4.2%)	434 (81.9%)	
Education				0.300
<i>College and below</i>	16 (5.1%)	15 (7.0%)	31 (5.8%)	
<i>Undergraduate</i>	157 (49.7%)	120 (56.1%)	277 (52.3%)	
<i>Master's degree</i>	114 (36.1%)	68 (31.8%)	182 (34.3%)	
<i>PhD</i>	29 (9.2%)	11 (5.1%)	40 (7.5%)	
Working years of psychotherapy				0.001***
<i><7 years</i>	160 (50.6%)	119 (55.6%)	279 (52.6%)	
<i>≥ 7years</i>	156 (49.4%)	95 (44.4%)	251 (47.4%)	
Practice qualification				
<i>Licensed national second/third level psychological counselor</i>	237 (75.0%)	184 (86.0%)	421 (79.4%)	0.164
<i>Licensed psychotherapist</i>	76 (24.1%)	41 (19.2%)		0.240
<i>Licensed psychiatrist</i>	91 (28.8%)	36 (16.8%)	117 (22.1%)	0.006*
<i>Licensed psychologist in educational system</i>	45 (14.2%)	44 (20.6%)	127 (24.0%)	0.081
Working places for psychotherapy				0.008**
<i>Hospital</i>	130 (41.1%)	63 (29.4%)	193 (36.4%)	
<i>School</i>	49 (15.5%)	44 (20.6%)		
	112 (35.4%)	71 (33.2%)	93	

<i>Counseling agency</i>	12 (3.8%)	9 (4.2%)	(17.5%)	
<i>Network platform</i>	13 (4.1%)	27 (12.6%)	(34.5%)	
<i>Other</i>			21 (4.0%)	
			40 (7.5%)	
Have professional supervisor				0.746
<i>Yes</i>	258 (81.6%)	168 (78.5%)	426 (80.4%)	
<i>No</i>	58 (18.4%)	46 (21.5%)	104 (19.6%)	
Have professional personal experience				0.714
<i>Yes</i>	236 (74.7%)	152 (71.0%)	388 (73.2%)	
<i>No</i>	80 (25.3%)	62 (29.0%)	142 (26.8%)	
Professional background				
<i>Psychoanalysis or psychodynamic therapy</i>	171 (54.1%)	107 (50.0%)	278 (52.5%)	0.521
<i>Cognitive behavioral therapy</i>	137 (43.4%)	99 (46.3%)		0.623
<i>Humanistic therapy</i>	86 (27.2%)	62 (29.0%)	236 (44.5%)	0.707
<i>Family therapy</i>	172 (54.4%)	128 (59.8%)		0.419
<i>Narrative therapy</i>	36 (11.4%)	33 (15.4%)	148 (27.9%)	0.207
<i>Others</i>	48 (15.2%)	45 (21.0%)	300 (56.6%)	0.115
			69 (13.0%)	
			93 (17.5%)	
Assessment of possible side effects in psychotherapy				0.581
<i>Yes</i>	279 (88.3%)	181 (84.6%)	460 (86.8%)	
<i>No</i>	7 (2.2%)	9 (4.2%)	16 (3.0%)	
<i>Not sure</i>	30 (9.5%)	24 (11.2%)	54 (10.2%)	
Possible causes of side effects in psychotherapy				
<i>Characteristics of psychotherapy techniques</i>	148 (46.8%)	62 (29.0%)	210 (39.6%)	0.001***
<i>Improper use of psychotherapy techniques</i>	226 (71.5%)	115 (53.7%)		0.012*
<i>Limited professional abilities of the therapist</i>	258 (81.6%)	153 (71.5%)	341 (64.3%)	0.193
<i>Clients' psychological activity</i>	200 (63.3%)	111 (51.9%)	411 (77.5%)	0.092
<i>Therapists' psychological activity</i>	196 (62.0%)	84 (39.3%)		0.0004***

<i>Other unpredictable factors</i>	360 (67.9%)
* P < 0.050 was considered statistically significant; **: p < 0.01; ***: p < 0.001.	

Table 2
Types of side effects perceived by therapists

Content of the side effect	n (%)
Has your psychotherapy made clients or patients feel bad?	264 (49.8%)
Has your psychotherapy strained the clients' or patients' family relationship?	97 (18.3%)
Has your psychotherapy strained the relationship outside the clients' or patients' family?	68 (12.8%)
Has your psychotherapy worsened the symptoms of clients or patients?	67 (12.6%)
Has your psychotherapy made clients or patients behave badly?	58 (10.9%)
Has your psychotherapy made clients or patients' physical health uncomfortable?	33 (6.2%)
Has your psychotherapy worsened the clients' or patients' working conditions?	31 (5.8%)

Table 3
Comparison the prediction performance of different machine learning algorithms to distinguish psychotherapists with or without the perception of clients' side effects in psychotherapy

Classifier	Precision	Recall	F1-Score	AUC
Random Forest	0.686	0.761	0.722	0.717
XGBoost	0.677	0.685	0.681	0.689
CatBoost	0.689	0.674	0.681	0.694
Logistic Regression	0.647	0.717	0.680	0.675
AdaBoost	0.633	0.674	0.653	0.653
SVM	0.596	0.707	0.647	0.629

Figures

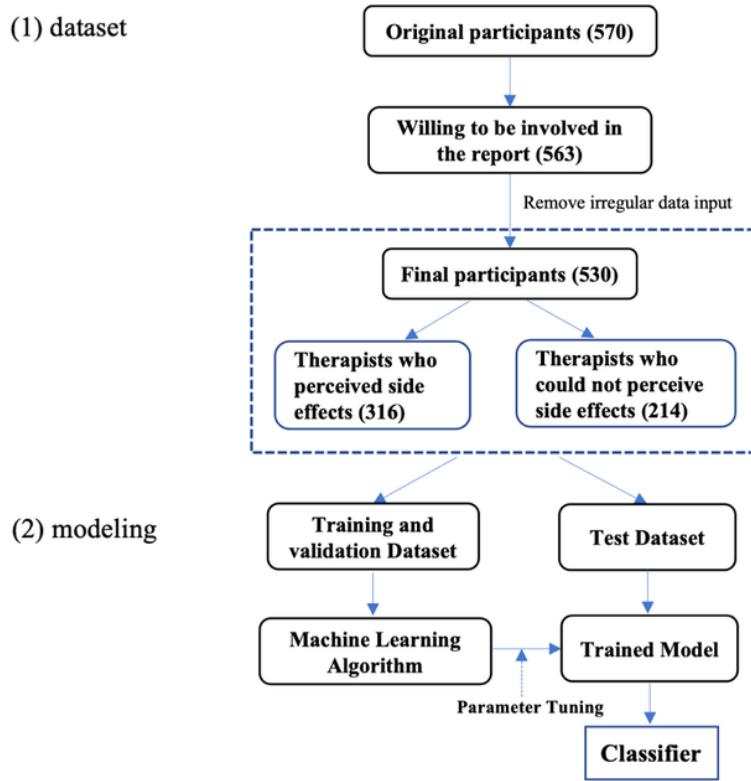


Figure 1

The workflow of data processing and machine-learning based modeling.

(1) 570 therapists were involved in the original PSEQ. By removing therapists unwilling to make their data public and with irregular data input, 530 therapists were finally involved in the dataset. 316 therapists reported that they could perceive clients' side effects in psychotherapy, and 214 therapists didn't report perceiving side effects. (2) The whole dataset was split into a training and validation dataset and a test dataset. Six different machine learning algorithms were selected for training based on the training and validation dataset. Trained models were obtained after parameter tuning. The final classifier was determined according to the comparison of each trained model's prediction performance.

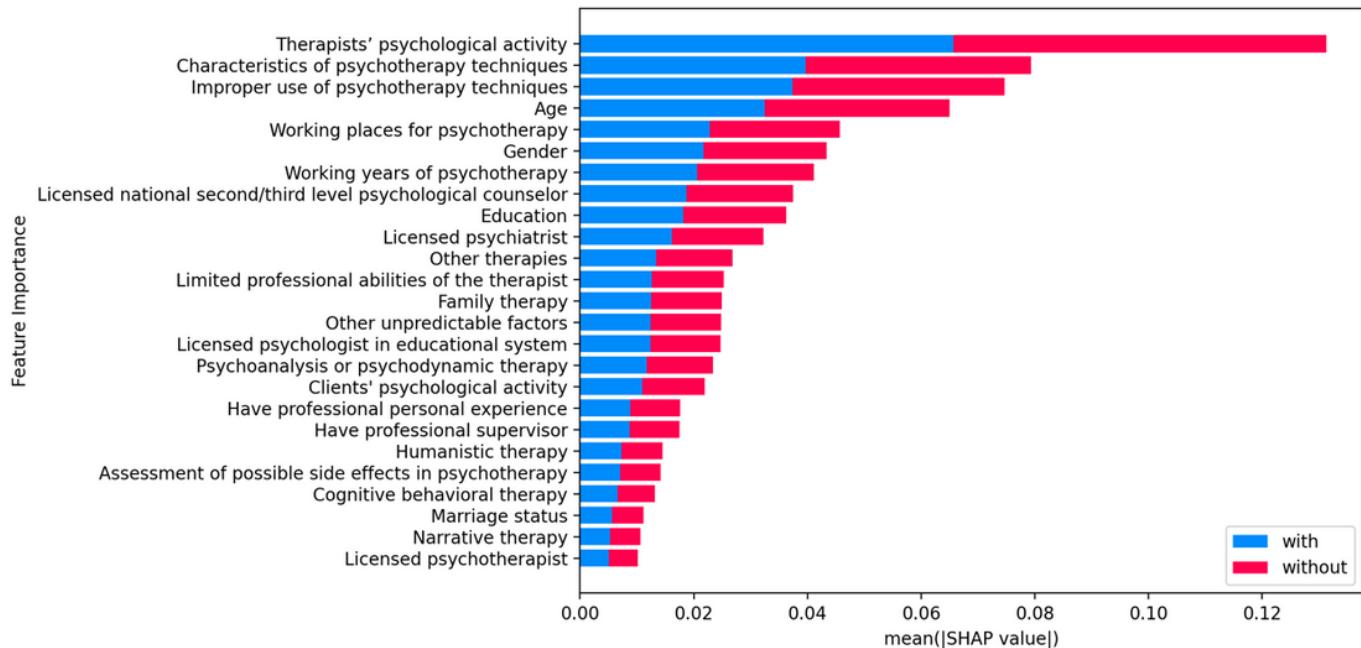


Figure 2

SHAP summary plot of the Random Forest-based classifier

The relative importance for each feature in the classifier, obtained by taking the average absolute value of each feature' SHAP value.

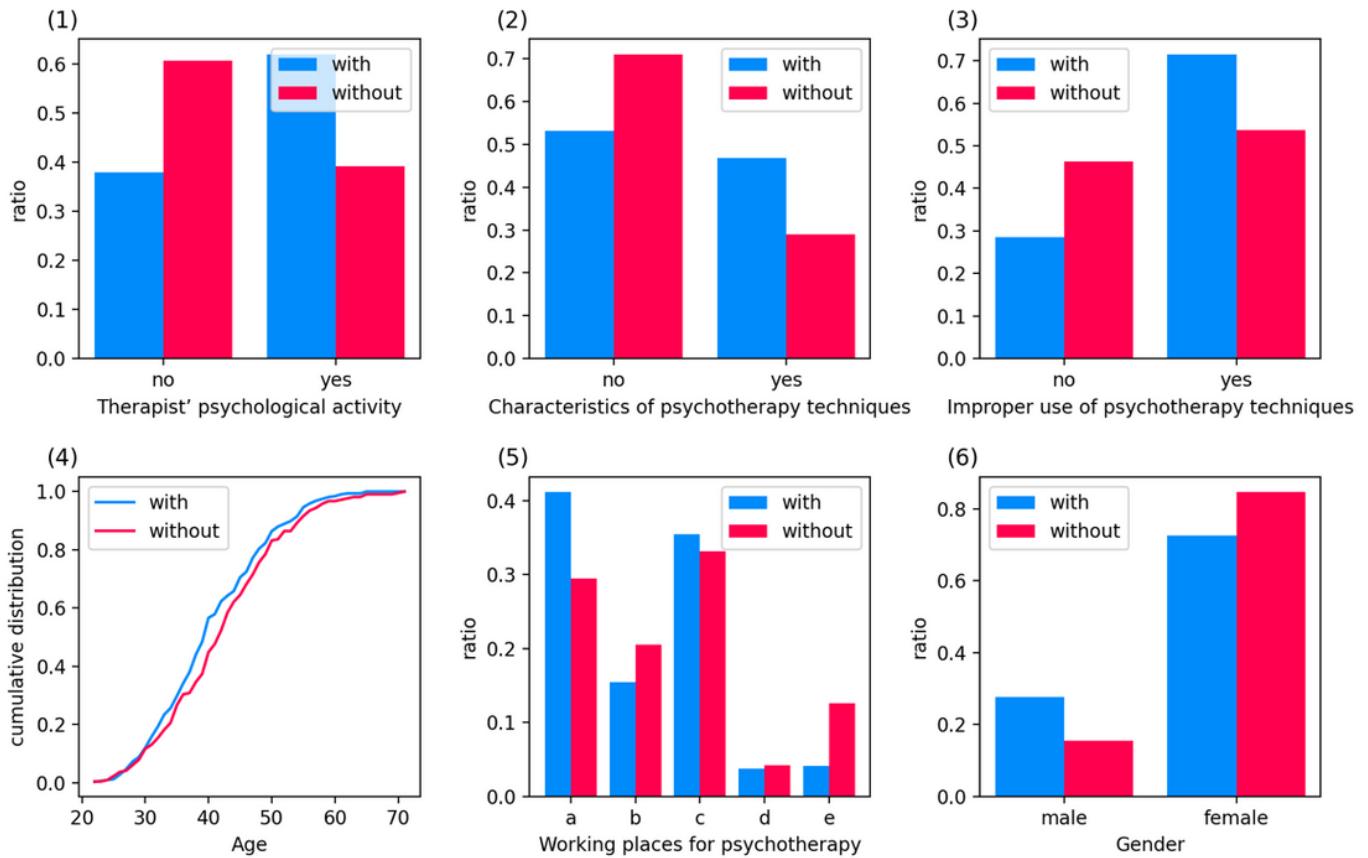


Figure 3

Comparison between therapists with or without perception of the side effects based on graph metrics.

(1) therapist's psychological activity; (2) characteristics of psychotherapy techniques; (3) improper use of psychotherapy techniques; (4) age of psychotherapists; (5) working places for psychotherapy, 'a' to 'e' denotes hospital, school, counseling agency, network platform and others, respectively; (6) gender. Blue column or line: psychotherapists with perception of clients' side effects; Red column or line: psychotherapists without perception of clients' side effects.