Therapists and psychotherapy side effects in China: A machine learning-based study

Lijun Yao, Zhiwei Xu, Xudong Zhao, Yang Chen, Liang Liu, Xiaoming Fu, Fazhan Chen

ARTICLE INFO

Keywords:
Side effects
Psychotherapy
Therapist
Machine learning
Artificial intelligence

ABSTRACT

Objective: Side effects in the 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. 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 clients’ side effects in psychotherapy, with a 59.6% incidence of 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. 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. We leveraged the Shapley Additive exPlanations (SHAP) method to quantify the importance of each feature to the therapist categories.

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

1. Introduction

Psychotherapy is a process of actively eliminating or alleviating symptoms through the therapeutic relationship and interaction between psychotherapists and clients, helping clients improve their personality, adapt to society, and promote rehabilitation (Kreuzer et al., 2018). As an effective method of medical treatment (Dragioti et al., 2017), psychotherapy may also result in side effects or even cause harm (Lilienfeld, 2007). For example, discussing a client’s early-life trauma during psychotherapy may worsen his/her symptoms in the short term (Cloitre et al., 2010). In addition, the therapist’s deep empathy for the client may increase the client’s dependence on psychotherapy resulting in prolonged treatment (Feng et al., 2020; Linden and Schermuly-Haupt, 2014). Many previous studies have confirmed that side effects in psychotherapy were common. The National Audit of Psychological Therapies (NAPT) undertaken in Wales and England reported that 5.2% of the patients had lasting adverse effects from psychological treatment (Crawford et al., 2016). In psychotherapy, 38.5% of patients with depression (n = 135) exhibited one adverse effect (Peth et al., 2018). Our recent study reported that the incidence of psychotherapy side effects was 31.1% (115/370), and “feel bad in psychotherapy” was the most common side effect (24.6%) (Yao et al., 2020). In group psychotherapy, 43.7% of patients experienced severe and extremely severe side effects or burdens (Linden et al., 2020). For young adult clients, the incidence of psychotherapy side

* Corresponding author.
E-mail address: chenzf@shspdjw.com (F. Chen).
effects was approximately 41%, too long of a treatment, and the deterio-
ration of the existing symptoms were the strongest predictors of poor
therapeutic effects (Lorenz, 2021).

Unfortunately, the side effects of psychotherapy have not attracted
much attention. Only about 21% of these randomized controlled psy-
chotherapy trials monitored patient-perceived harm, and about 3% described
adverse events (Jonsson et al., 2014). Many clinicians or therapists fail to
uncover and handle these side effects, mainly because of not having
enough awareness of the side effects of psychotherapy. Boisvert and Faust
(2003, 2006) 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 (Haffield et al., 2010).

Sensitivity to the client’s side effects in psychotherapy is a valuable in-
dicator of a good therapist. Enhancing the therapist’s sensitivity can greatly
improve the treatment quality (Linden, 2013). 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 on the side effects of psychotherapy
(Rozental et al., 2018). Thus, more efforts need to be made in psycho-
therapy from a therapist’s perspective to improve research and clinical
awareness of identifying and avoiding side effects (Mc Glanaghy et al.,
2022). Moreover, the therapist factors are closely related to the efficacy
of psychotherapy. The National Institute of Mental Health Treatment of
Depression Collaborative Research Program (Kim et al., 2006) noted that
about 8% of the variance in outcomes in psychotherapy could be attributed
to the therapist. Another work demonstrated that approximately 8% of the
total variance and approximately 17% of the variance in patient improve-
ment could be attributed to the therapist (Lutz et al., 2007).

Similarly, therapist factors may influence psychotherapy side effects.
The therapist’s inappropriate narratives could undermine the outcome of
psychotherapy and the therapeutic alliance, particularly therapists’ con-
trolling and challenging statements (Kadur et al., 2020). A NAPT study
shows that patients experience more negative effects in psychological
treatment when their treatment preferences are not satisfied (Williams
et al., 2016). The treatment preferences included the characteristics of
the therapist. In our recent study, the mental state of the psychotherapist
was the most crucial feature in predicting whether a client would experi-
ence side effects in psychotherapy (Yao et al., 2020). Interestingly,
when the psychotherapy outcome was regularly monitored and fed back
to the psychotherapist, the deterioration effect of the psychotherapy was
significantly reduced (Lambert et al., 2002). Early identification of
treatment failure and problem-solving strategies by the psychotherapist
in routine practice would also significantly improve the effectiveness of
psychotherapy (Whipple et al., 2003).

In summary, the therapist plays an essential role in the occurrence of
psychotherapy side effects. However, the questions remain, how can
therapists perceive clients’ side effects, and which factors determine the
degree of their perceptions. Because most studies on psychotherapy side
effects were based on the client’s or patient’s perspective, the mechanisms
by which therapists play a role in developing side effects remain unclear.

Machine learning (ML) is a form of artificial intelligence that auto-
matically learns from data and builds classification or predictive models.
It could be adopted to predict outcomes for a new data instance. In
psychiatry, ML has been applied in diagnosing, progression, treatment
prediction, and detecting potential biomarkers of mental disorders
(Aaifees-van Doorn et al., 2021). Studies used electronic health records,
brain imaging, cognitive testing, rating scales, genetics, electrophysi-
ology, smartphone, and social media data to predict, classify, or subgroup
mental health problems including schizophrenia, depression, and sui-
cide, et al. (Chekroud et al., 2021; Graham et al., 2019). ML also may
significantly impact mechanism and process, training and feedback, and
technology-mediated psychotherapy treatment modalities (Imel et al.,
2017). Nevertheless, the applications of ML in psychotherapy-related
studies are still minimal (Chekroud et al., 2021). Goldberg et al. used
ML for predicting the client-therapist alliance from linguistic content
during psychotherapy (Goldberg et al., 2020). Their results modestly
predicted alliance ratings and suggested ML technologies would help examine alliances in future studies. In one of our recent works, super-
vised ML was adopted to build a prediction model that can uncover cli-
ents who might have consulting side effects in psychotherapy based on
information from the psychotherapy (Yao et al., 2020). By comparing six
models based on different ML algorithms, the model using the Random
Forest algorithm performed the best in predicting psychotherapy out-
comes. In the present study, we focused on the therapist factors in psy-
chotherapy and further used ML techniques to distinguish therapists with
different perceptions of client-side effects in psychotherapy.

In this study, combined with a self-designed Questionnaire, statistics,
ML, and Shapley Additive exPlanations (SHAP) methodology (Lundberg
and Lee, 2017), we aim to investigate the types of psychotherapy 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 per-
ceive the side effects of psychotherapy. The findings of this research
are to develop a scientific framework to improve the therapists’ ability to
uncover and handle psychotherapy’s side effects and provide a more solid
basis for the professionalization of psychotherapy.

2. Methods

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

According to previous research results (Chen and Zhao, 2017; Linden
and Schermuly-Haupt, 2014), we designed the Psychotherapy Side Effects
Questionnaire-Therapist Version (PSEQ-T). In the PSEQ-T, psychotherapy is
the process by which a trained professional therapist uses guided con-
versations to promote changes in a client’s thoughts, feelings, and be-
haviors; side effects in psychotherapy are defined as unwanted events
perceived by the therapist during psychotherapy that are inconsistent with
the expected goals and negatively impact the client. The side effects of
psychotherapy were assessed by referring to the answers provided by the
respondents 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 the therapists’ ability to perceive the side
effects; otherwise, we conclude that there was no indication of the side
effects. Seven questions in the PSEQ-T were designed to assess therapist-perceived client-side effects across three different dimensions of
symptoms, relationships, and social functions (Table 2). We designed three
questions 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 psycho-
therapy caused discomfort in your clients or patients’ physical health?”).
We used one question to assess the original problem (“Has your psycho-
therapy worsened the symptoms of your clients or patients?”). We used
two questions to assess negative changes in family relationships (“Has
your psychotherapy strained your clients’ or patients’ family relation-
ship?”) and interpersonal relationships (“Has your psychotherapy strained
the relationship outside your clients’ or patients’ family?”). We used the
last question to assess negative changes in social functions (“Has your
psychotherapy worsened your clients’ or patients’ working conditions?”).

To predict the perception of the side effects of psychotherapists in
psychotherapy, we extracted each participant’s following features in the
PSEQ-T: demographics, clinical practice information, occupational infor-
mation, and the possible causes of side effects in psychotherapy. The
detailed information for each feature was listed in Table 1. We sent the
questionnaire to ten examiners for content revision. We then revised it
based on the feedback to form the final version of the PSEQ-T. In the PSEQ-
T, Cronbach’s α was 0.667, indicating acceptable internal consistency for
this survey (Churchill, 1979; Fleming, 2011; Gallais et al., 2017; Setbon
and Raude, 2010).
| Features                                      | With perception of the side effects (n=316) | Without perception of the side effects (n=214) | Overall (%) | P-value |
|-----------------------------------------------|--------------------------------------------|-----------------------------------------------|-------------|---------|
| **Gender**                                   |                                            |                                               |             |         |
| Male                                          | 87 (27.5%)                                 | 33 (15.4%)                                    | 120 (22.6%) | 0.304   |
| Female                                        | 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   |
| Single                                        | 27 (8.5%)                                  | 17 (7.9%)                                     | 44 (8.3%)   |         |
| Single with partner                          | 16 (5.1%)                                  | 11 (5.1%)                                     | 27 (5.1%)   |         |
| Married                                       | 257 (81.3%)                                | 177 (82.7%)                                   | 434 (81.9%) |         |
| Divorced, separated or widowed               | 16 (5.0%)                                  | 9 (4.2%)                                      | 25 (4.7%)   |         |
| **Education**                                 |                                            |                                               |             |         |
| College and below                            | 16 (5.1%)                                  | 15 (7.0%)                                     | 31 (5.8%)   |         |
| Undergraduate                                 | 157 (49.2%)                                | 120 (56.1%)                                   | 277 (52.3%) |         |
| Master’s degree                              | 114 (36.1%)                                | 68 (31.8%)                                    | 182 (34.3%) |         |
| PhD                                          | 29 (9.2%)                                  | 11 (5.1%)                                     | 40 (7.5%)   |         |
| **Working years of psychotherapy**           |                                            |                                               |             | 0.001***|
| <7 years                                     | 160 (50.6%)                                | 119 (55.6%)                                   | 279 (52.6%) |         |
| ≥7 years                                     | 156 (49.4%)                                | 95 (44.4%)                                    | 251 (47.4%) |         |
| **Practice qualification**                   |                                            |                                               |             |         |
| Licensed national second/third level psychological counselor | 237 (75.0%) | 184 (86.0%) | 421 (79.4%) | 0.164 |
| Licensed psychotherapist                     | 76 (24.1%)                                 | 41 (19.2%)                                    | 117 (22.1%) | 0.240 |
| Licensed psychiatrist                        | 91 (28.8%)                                 | 36 (16.8%)                                    | 127 (24.0%) | 0.006* |
| Licensed psychologist in educational system  | 45 (14.2%)                                 | 44 (20.6%)                                    | 89 (16.8%)  | 0.081  |
| **Working places for psychotherapy**         |                                            |                                               |             | 0.008**|
| Hospital                                     | 130 (41.1%)                                | 63 (29.4%)                                    | 193 (36.4%) |         |
| School                                       | 49 (15.5%)                                 | 44 (20.6%)                                    | 93 (17.5%)  |         |
| Counseling agency                            | 112 (35.4%)                                | 71 (33.2%)                                    | 183 (34.5%) |         |
| Network platform                             | 12 (3.8%)                                  | 9 (4.2%)                                      | 21 (4.0%)   |         |
| Other                                        | 13 (4.1%)                                  | 27 (12.6%)                                    | 40 (7.3%)   |         |
| Have professional supervisor                 |                                            |                                               |             | 0.746  |
| Yes                                          | 258 (81.6%)                                | 168 (78.5%)                                   | 426 (80.4%) |         |
| No                                           | 58 (18.4%)                                 | 46 (21.5%)                                    | 104 (19.6%) |         |
| Have professional personal experience        |                                            |                                               |             | 0.714  |
| Yes                                          | 236 (74.7%)                                | 152 (71.0%)                                   | 388 (73.2%) |         |
| No                                           | 80 (25.3%)                                 | 62 (29.0%)                                    | 142 (26.8%) |         |
| **Professional background**                  |                                            |                                               |             |         |
| Psychoanalysis or psychodynamic therapy      | 171 (54.1%)                                | 107 (50.0%)                                   | 278 (52.5%) | 0.521  |
| Cognitive behavioral therapy                 | 137 (43.4%)                                | 99 (46.3%)                                    | 236 (44.5%) | 0.623  |
| Humanistic therapy                           | 86 (27.2%)                                 | 62 (29.0%)                                    | 148 (27.9%) | 0.707  |
| Family therapy                               | 172 (54.4%)                                | 128 (59.8%)                                   | 300 (56.6%) | 0.419  |
| Narrative therapy                            | 36 (11.4%)                                 | 33 (15.4%)                                    | 69 (13.0%)  | 0.207  |
| Others                                       | 48 (15.2%)                                 | 45 (21.0%)                                    | 93 (17.5%)  | 0.115  |
| Assessment of possible side effects in psychotherapy | 279 (88.3%) | 181 (84.6%) | 460 (86.8%) | 0.581  |
| Yes                                          | 279 (88.3%)                                | 181 (84.6%)                                   | 460 (86.8%) |         |
| No                                           | 7 (2.2%)                                   | 9 (4.2%)                                      | 16 (3.0%)   |         |
| Not sure                                     | 30 (9.5%)                                  | 24 (11.2%)                                    | 54 (10.2%)  |         |
| Possible causes of side effects in psychotherapy | 148 (46.8%) | 62 (29.0%) | 210 (39.6%) | 0.001***|
| Characteristics of psychotherapy techniques  | 148 (46.8%)                                | 62 (29.0%)                                    | 210 (39.6%) |         |
| Improper use of psychotherapy techniques     | 226 (71.5%)                                | 115 (53.7%)                                   | 341 (64.3%) | 0.012* |
| Limited professional abilities of the therapist | 258 (81.6%) | 153 (71.5%) | 411 (77.5%) | 0.193  |
| Clients’ psychological activity              | 200 (63.3%)                                | 111 (51.9%)                                   | 311 (58.7%) | 0.092  |
| Therapists’ psychological activity           | 196 (62.0%)                                | 84 (39.3%)                                    | 280 (52.8%) | 0.0004***|
| Other unpredictable factors                  | 207 (65.5%)                                | 153 (71.5%)                                   | 360 (67.9%) | 0.412  |

*P < 0.050 was considered statistically significant; **: p < 0.01; ***: p < 0.001.
2.2. The procedure of data collection

The questionnaire was published via the WeChat platform on Feb. 11, 2019. Each participant was required to decide whether to complete the questionnaire according to the pre-given inclusion criteria and chose informed consent before submitting the questionnaire. The questionnaire was anonymous. Participants completed the questionnaire using WeChat’s mobile device-based interface. For each questionnaire, the completion time was approximately three to 5 min. We used an Excel form to collect the responses from different participants. Data collection ceased on Jun. 6, 2019.

2.3. Entry requirements of the participants

Participants joined our study through the online questionnaire (PSEQ-T) published on the official WeChat account from Feb. 11 to Jun. 6, 2019. Inclusion criteria included that the participants (1) carried out at least one session of psychotherapy in the last month, (2) had a licensed practice qualification of psychological intervention issued by the government, (3) aged from 18–70 years, and (4) read informed consent. Meanwhile, exclusion criteria included the participants’ having (1) severe mental disorders or physical illnesses, (2) ethical faults, or (3) disagreements with the release of the anonymized research data to the public.

2.4. Classification of therapists using machine learning

In our work, we built a model based on supervised machine learning technologies that could predict whether the therapist could perceive the client's side effects in psychotherapy. In our collected dataset, we selected therapists “with perceptions of clients' side effects” category as positive instances and therapists “without perceptions of clients' side effects” category as negative instances. All the features used to build the classifier were listed in Table 1. The overall working procedure of the raw data preprocessing, training of the machine learning model, and model performance evaluation were described in Figure 1.

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 did not 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.

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%) |
The final dataset included 316 therapists who reported having perceptions of clients’ side effects in psychotherapy and 214 therapists who did not. Using the SMOTE technique (Chawla et al., 2002), the minority type was oversampled to 316. Afterward, the balanced dataset was randomly split into a training and validation subset and a test subset. 70% of the data was used for training and validation, and 30% was used for testing. The 5-fold cross-validation method was applied. The training and validation subset was randomly divided into 5 groups of the same size. The cross-validation procedure was repeated for 5 rounds. In each round, 4 groups were used for training, while the remaining one was chosen as the validation data used to quantify the model’s prediction performance.

We have selected traditional ML algorithms that are widely used, such as Random Forest (Breiman, 2001), Logistic Regression (Dreiseitl and Ohno-Machado, 2002), Support Vector Machine (SVM) (Hearst, 1998), AdaBoost (Freund and Schapire, 1997), and algorithms with excellent predictive effects developed in recent years, including CatBoost (Prokhorenkova et al., 2018) and XGBoost (Chen and Guestrin, 2016), 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 and validation subset, grid search was used to scan through a series of possible parameter combinations. A finite set of values for different parameters were chosen to form the parameter space, and different parameter combinations were scanned one by one. Second, for every parameter combination, we quantified the predictive performance using the F1 score. For the final step, we selected the parameter combination achieved the largest F1 score based on the training and validation subset. We applied a Python-based machine learning library, called scikit-learn, for model training and validation (Pedregosa et al., 2011).

To evaluate the predictive performance of the trained model, we adopted precision, recall, the F1 score, and the area under the ROC curve (AUC) value (Fawcett, 2006). Precision represents the fraction of the therapists classified by the model as “with perceptions of clients’ side effects” who perceived clients’ side effects. Recall means the ratio of the therapists “with perceptions of clients’ side effects” correctly uncovered by the model. The F1 score is defined as the harmonic mean of precision and recall, which could be calculated as follows:

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

The highest value of the F1 score is 1, and the smallest value of the F1 score is 0. A larger F1 score means the classifier has a better overall predictive performance. As another important evaluation metric to examine a classifier’s predictive performance, AUC represents the probability that a random positive instance will be ranked higher than a random negative instance. The largest possible value of AUC is 1, indicating a perfect prediction. In our study, when the AUC value was higher, the model was able to better distinguish between the therapists with or without perceptions of clients’ side effects in psychotherapy.

2.5. Analysis of the feature importance and statistics

To interpret the contributions of different features of the prediction model, we used the SHAP method to evaluate each feature’s importance in the prediction model. SHAP is a representative method to explain the model, we used the SHAP method to evaluate each feature’s importance in the prediction model. SHAP is a representative method to explain the model’s prediction performance.

Comparison the prediction performance of different machine learning algorithms—i.e., Random Forest, Logistic Regression, XGBoost, CatBoost, AdaBoost, and SVM—were selected to implement the models for classification. Then, we compared each classification model’s predictive performance to achieve the best classifier. Our findings demonstrated that the F1 scores of the selected machine learning 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 values were also shown in Table 3. The AUC values of the six machine learning models—Random Forest, XGBoost, CatBoost, Logistic Regression, AdaBoost, and SVM—were presented in Table 3. The AUC values of the six machine learning 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 classifier using the Random Forest algorithm was considered statistically significant.

3. Results

3.1. Participants’ demographics

A total of 570 therapists completed the online questionnaires. 7 participants were unwilling to share their information with the public, and 33 participants were excluded from further analysis since the input data was irregular. The final dataset has 530 participants (Figure 1(1)). Each instance has 12 main features, and each feature was either categorical or numerical. The data for each feature was presented in Table 1. The average age of therapists included in the analysis was 41.3 years (SD = 9.00 years), while the average age of therapists who perceived client-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 serving 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.

3.2. Types of client-side effects perceived by the therapist

In our study, 316 therapists perceived the clients’ side effects in their current psychotherapeutic offerings, and the incidence was 59.6%. Among all of the 7 side effect types, the most common side effect 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 types and incidences of each side effect perceived by the therapists were described in Table 2.

3.3. Distinguishing therapists with different perception of client's side effects in psychotherapy by ML

In this part, we used supervised machine learning-based classifiers to distinguish between therapists with and without the perception of the client’s psychotherapy side effects. Six classic machine learning algorithms—i.e., Random Forest, Logistic Regression, XGBoost, CatBoost, AdaBoost, and SVM—were selected to implement the models for classification. Then, we compared each classification model’s predictive performance to achieve the best classifier. Our findings demonstrated that the F1 scores of the selected machine learning 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 values were also shown in Table 3. The AUC values of the six machine learning 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 classifier using the Random Forest algorithm was considered statistically significant.

| 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 |

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.
achieved the largest F1 score of 0.722 and AUC value of 0.717, showing the best predictive performance for discriminating between the therapist's perceptions of the client's side effects in psychotherapy.

3.4. Important features distinguishing therapists with different perception of client's 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 (Figure 2). Based on each feature's |SHAP value|, “therapists’ psychological activity” ranked 1\(^{st}\) among all the analyzed features, thus contributing most to distinguishing between the therapists with or without perception of client's 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| (Figure 3). The perceptive therapists were more likely to believe that (1) the therapist's psychological activity would affect clients' side effects (Figure 3(1)), (2) the characteristics of psychotherapy would cause clients' side effects (Figure 3(2)), and (3) the improper use of psychotherapeutic techniques would cause clients' side effects (Figure 3(3)). When the psychotherapist was younger, he or she was more likely to perceive clients' side effects (Figure 3(4)). Workplaces also affected clients' side effects, and the therapists working in hospitals were more likely to perceive them (Figure 3(5)). Among the male therapists, the percentage of perceptive therapists was slightly lower than that of the nonperceptive ones, while for the female therapists, the percentage of the perceptive therapists was higher than that of the nonperceptive ones, (Figure 3(6)).

Overall, we found clear differences between the two therapist groups in terms of the following features, including therapist's psychological activity, characteristics of psychotherapy techniques and how they are used, age of the therapists, working places where the psychotherapy occurred, and gender of the therapists.

4. Discussion

In this study, we leveraged machine learning technologies 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 client-side effect perceived by therapists was “made the clients or patients feel bad” (49.8%). Among the algorithms we have explored, the classifier based on the Random Forest algorithm provided the highest predictive performance in distinguishing therapists with different perceptions of client's 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 important feature for distinguishing between 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 psychotherapy had negative effects and 75% claimed that they had clinical experience with negative effects, only 8 (11%) of clinicians have gained information about negative effects during their basic clinical training (Bystedt et al., 2014). In our study, the accuracy of identifying side effects by therapists was significantly higher than that of previous studies (Boisvert and Faust, 2003; 2006; Rieffeld et al., 2010), 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. However, there is no information about psychotherapy side effects in training therapists in China (Chen and Zhao, 2017) or worldwide (Rozental et al., 2018). Licensed therapists may be able to uncover and handle side effects in psychotherapy to ensure the professionalism and standardization of the psychological industry. This is also the primary purpose of this study, but much work remains to be done.

In our study, “made the clients or patients feel bad” (49.8%) was the most common client-side effect reported by therapists, which is consistent with earlier findings (Bystedt et al., 2014; Gerke et al., 2020; Schermuly-Haupt et al., 2018; Yao et al., 2020). Qualitative studies focusing on the therapists’ views on the negative effects of psychotherapy showed that the characteristics of negative effects included “short-term negative effects”, “no treatment effect”, “deterioration”, “dependency”, and “impact on other life domains” (Bystedt et al., 2014). “Deterioration” was one of the common side effects of psychological treatments. In the few available quantitative studies on therapists’ experiences, cognitive behavior therapists rated the most frequent side effects as “negative wellbeing/distress” (27%), “worsening of symptoms” (9%), and “strains

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.
in family relations” (6%) (Schermuly-Haupt et al., 2018). Similar psychotherapy outcomes 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 reported by clients was “feel bad in psychotherapy” (24.6%) (Yao et al., 2020). 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%) (Gerke et al., 2020). In this study, “feel bad” refer to negative emotions experienced by clients, such as sadness, anger, anxiety, and tension. As new symptoms emerge in psychotherapy, these negative feelings may be related to therapists, patients, and the therapeutic alliance (Parry et al., 2016). 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 intervention. The question is how to identify patients with potential side effects by psychotherapy and therapists who can perceive client-side effects in intervention. In our previous work, we used ML to find clients who might have psychotherapy side effects. 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, improving psychotherapy outcomes. Similarly meaningful would be if we could use ML to distinguish between therapists who can and cannot perceive the client-side effects of psychotherapy. In our work, we used a self-compiled questionnaire to extract features from three different 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, demonstrating that the model could 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 client-side effects in 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 cause 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 psychotherapy side effects (Kadur et al., 2020; Williams et al., 2016; Yao et al., 2020). Therapist factors mediate the outcomes of psychotherapy primarily through therapeutic alliances. On average, therapists who developed stronger alliances with their clients could achieve better therapeutic outcomes. Destructive therapeutic alliances were particularly evident in the therapists’ mental state performance, such as controlling and challenging statements (Fluckiger et al., 2018). An excellent therapeutic alliance values a supportive and reinforcing context, such as when there are fewer stressful interventions, and the therapeutic relationship is comfortable. The therapist’s mental activity affected the client through the therapeutic relationship. It was essential for psychotherapeutic side effects (Yao et al., 2020). Combined with previous research results, the present study suggested that the therapist’s introspection and management of their psychological activity will help the therapist to identify and monitor the side effects in psychotherapy, which could significantly reduce the deterioration effect (Lambert et al., 2002) and improve the effectiveness of psychotherapy (Whipple et al., 2003).

In this study, two other important predictors were “characteristics of psychotherapy techniques” and “improper use of psychotherapy techniques”. The essence of psychotherapy is to help people learn who they are, access their emotional basics, hold their feelings intact, and think even under the heaviest interpersonal pressure, which is the first and main therapeutic goal (Bugliani, 2020). 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 (Crawford et al., 2016) believe that “using an inappropriate therapeutic method or errors in delivering a recommended therapy” might be risk factors for adverse outcomes and possible mechanisms for harmful psychological therapies. Studies have shown that the theoretical orientation of psychotherapy significantly affects the occurrence of client-side effects (Crawford et al., 2016; Yao et al., 2020). For example, patients with poor therapeutic relationships, high dependency or isolation, and high psychotherapy burden received treated psychodynamic therapy more often (Leitner et al., 2013). Although such a therapeutic process is effective, it places tremendous pressure on patients. Furthermore, inappropriate intervention techniques may lead to malpractice and unethical behavior in psychotherapy. 28.8% of inpatients and 7.1% of outpatients reported at least one incident of malpractice and unethical behavior in psychotherapy (Gerke et al., 2020). Therefore, this study indicates that understanding the limitations of intervention theory and technology will help therapists identify the side effects of psychotherapy.
To our best knowledge, our work is the first machine learning-based approach to predict the potential side effects perceived by therapists in psychotherapy. The supervised machine learning-based models investigated in this study are useful and practical enough to be applied in clinical psychiatry. Our research provides new methods that can be used to differentiate therapists with different client-side effects susceptibility and suggests important predictive factors that affect the therapist’s perception. The study demonstrates a possible technical path that can enhance the sensitivity and recognition of therapists to the side effects of 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 increase the recognition and management of side effects.

4.1. Limitations of the study

This study constructed a fairly accurate model to predict the therapists who can perceive client-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) the perceived side effects entirely come from the report by the therapists and cannot completely rule out the harm caused by anti-ethical problems; (3) this study is a cross-sectional study, and the number and representativeness of the research samples still need to be improved; (4) some important factors in psychotherapy, such as the treatment dosage and the therapist’s characteristics, were not involved in our study; (5) the present work did not cover which mental states of therapists were more likely to cause side effects of psychotherapy. This issue would be interesting to explore in our future research.

Declarations

Author contribution statement

Lijun Yao: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.
Zhiwei Xu; Yang Chen; Xiaoming Fu: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.
Xudong Zhao: Conceived and designed the experiments. Liang Liu: Performed the experiments; Contributed reagents, materials, analysis tools or data.
Fazhan Chen: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

Funding statement

This study was supported by the National Key Research and Development Program of China (2021ZD0202000), the Training Plan of Health System Academic Leader of Shanghai Pudong Municipality Health Commission (Grant Number: PWYjd2019-08), the Medical Discipline Construction Project of Pudong Health Committee of Shanghai (Grant No.: PWYgy2021-02), the Special Clinical Research Project of Shanghai Municipality Health Commission (Grant Number: 202040475), and the Shanghai Key Lab of Intelligent Information Processing (IIIP201911).

Data availability statement

Data will be made available on request.

Declaration of interest’s statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The authors would also 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.

References

Aafjes-van Doorn, K., Kamsteeg, C., Bates, J., Aafjes, M., 2021. A scoping review of machine learning in psychotherapy research. Psychother. Res. 31, 92–116.
Boisvert, C.M., Faust, D., 2003. Leading researchers’ consensus on psychotherapy research findings: implications for the teaching and conduct of psychotherapy. Prof. Psychol. Res. Pract. 34, 508–513.
Boisvert, C.M., Faust, D., 2006. Practicing psychologists’ knowledge of general psychotherapy research findings: implications for science-practice relations. Prof. Psychol. Res. Pract. 37, 708–716.
Brezina, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Bogliani, A., 2020. Wrongness: social side-effects in psychotherapy. Psychoanal. Inq. 40, 253–261.
Bystrode, S., Rouzet, A., Anderson, G., Boettcher, J., Carbring, P., 2014. Clinicians’ perspectives on negative effects of psychological treatments. Cognit. Behav. Ther. 43, 319–331.
Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.
Chekroud, A.M., Bondar, J., Delgado, J., Doeberty, G., Waul, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., et al., 2021. The promise of machine learning in predicting treatment outcomes in psychiatry. World Psychiatr. 20, 154–170.
Chen, F., Zhao, X., 2017. Side effects of psychotherapy. Chin. Ment. Health J. 31, 72–76.
Che, T.Q., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794.
Churchill, G.A., 1979. A paradigm for developing better measures of marketing constructs. J. Market. Res. 16, 64–73.
Cloitre, M., Stovall-McClough, K.C., Noonan, K., Zorbas, P., Cherry, S., Jackson, C.L., Gan, W., Petkova, E., 2010. Treatment for PTSD related to childhood abuse: a randomized controlled trial. Am. J. Psychiatr. 167, 915–924.
Crawford, M.J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., Clarke, J., Parry, G.D., 2016. Patient experience of negative effects of psychological treatment: results of a national survey. Br. J. Psychiatr. 208, 260–265.
Dragioti, E., Karathanos, V., Gerdle, B., Evangelou, E., 2017. Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. Acta Psychiatr. Scand. 136, 236–246.
Dreiseitl, S., Ohno-Machado, L., 2002. Logistic regression and artificial neural network classification models: a methodology review. J. Biomed. Inf. 35, 352–359.
Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27, 861–874.
Feng, Q., Zhao, X., Liu, L., Liu, Y., Chen, F., 2020. Quantitative research of side effects in psychotherapy and counseling based on client’s experience. Chin. Ment. Health J. 34, 903–910.
Fleming, R., 2011. An environmental audit tool suitable for use in homeless facilities for people with dementia. Australas. J. Ageing 30, 108–112.
Fluckiger, C., Del Re, A.C., Wampold, B.E., Horvath, A.D., 2018. The alliance in adult psychotherapy: a meta-analytic synthesis. Psychotherapy 55, 316–340.
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139.
Gallais, B., Gagnon, C., Forgues, G., Coste, L., Laberge, L., 2017. Further evidence for the reliability and validity of the fatigue and daytime sleepiness scale. J. Neurol. Sci. 375, 23–26.
Gerke, L., Meyrowitz, A.K., Ludwig, I., Rief, W., Nestoriuc, Y., 2020. Frequencies and results of a national survey. Br. J. Psychiatr. 208, 260–265.
Goldberg, S.B., Flemotomos, N., Martinez, V.R., Tanana, M.J., Kuo, P.B., Pace, B.T., 2020. SMOTE: a scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794.
Graham, S., Depp, C., Lee, E.E., Nebeeker, C., Tu, X., Kim, H.C., Jeste, D.V., 2019. Artificial intelligence for mental health and mental illnesses: an overview. Curr. Psychiatr. Rep. 21, 116.
Hatfield, D., McCullough, L., Frantz, S.H., Krieger, K., 2010. Do we know when our clients get worse? An investigation of therapists’ ability to detect negative client change. Clin. Psychol. Psychother. 17, 25–32.
Heard, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Hearst, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Heard, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Heard, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Heard, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Hearst, M.A., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13, 18–21.
Kadur, J., Ludemann, J., Andreas, S., 2020. Effects of the therapist's statements on the patient's outcome and the therapeutic alliance: a systematic review. Clin. Psychol. Psychother. 27, 168–178.

Kim, D.M., Wampold, B.E., Bolt, D.M., 2006. Therapist effects in psychotherapy: a random-effects modeling of the national Institute of mental health treatment of depression collaborative research Program data. Psychother. Res. 16, 161–172.

Kreuzer, J., DeLuca, J., Caplan, B., 2018. Encyclopedia of Clinical Neuropsychology. Springer, Cham.

Lambert, M.J., Whipple, J.L., Vermeersch, D.A., Smart, D.W., Hawkins, E.J., Nielsen, S.L., Goates, M., 2002. Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. Clin. Psychol. Psychother. 9, 91–103.

Leitner, A., Martens, M., Koschier, A., Gerlich, K., Liegl, G., Hinterwallner, H., Schnyder, U., 2013. Patients' perceptions of risky developments during psychotherapy. J. Contemp. Psychother. 43, 95–105.

Lilienfeld, S.O., 2007. Psychological treatments that cause harm. Perspect. Psychol. Sci. 2, 53–70.

Linden, M., 2013. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clin. Psychol. Psychother. 20, 286–296.

Linden, M., Muschalla, B., Walter, M., 2020. Gender and side effects of group cognitive behavior psychotherapy. Arch. Psychiatr. Ment. Health 4, 14–18.

Lunden, M., Schermuly-Haupt, M.L., 2014. Definition, assessment and rate of psychotherapy side effects. World Psychiatr. 13, 306–309.

Lorenc, T.K., 2021. Predictors and impact of psychotherapy side effects in young adults. Counsell. Psychother. Res. J. 21, 237–243.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30, 30 (Nips 2017).

Lutz, W., Leon, S.C., Martinovich, Z., Lyons, J.S., Stiles, W.B., 2007. Therapist effects in outpatient psychotherapy: a three-level growth curve approach. J. Counsel. Psychol. 54, 32–39.

Mc Glanaghy, E., Jackson, J.L., Morris, P., Prentice, W., Dougall, N., Hutton, P., 2022. Discerning the adverse effects of psychological therapy: consensus between experts by experience and therapists. Clin. Psychol. Psychother. 29, 579–589.

Parry, G.D., Crawford, M.J., Duggan, C., 2016. Iatrogenic harm from psychological therapies - time to move on. Br. J. Psychiatr. 208, 210–212.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Peth, J., Jelinek, I., Nestoriuc, Y., Moritz, S., 2018. Adverse effects of psychotherapy in depressed patients - first application of the positive and negative effects of psychotherapy scale (PANEPS). Psychother. Psychosom. Med. Psychol. 68, 391–398.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A., 2018. CatBoost: unbiased boosting with categorical features. Adv. Neur. In. 31.

Rozental, A., Castonguay, L., Dumidiian, S., Lambert, M., Shafran, R., Anderson, G., Carlbring, P., 2018. Negative effects in psychotherapy: commentary and recommendations for future research and clinical practice. BJPsych Open 4, 307–312.

Schermuly-Haupt, M.L., Linden, M., Rush, A.J., 2018. Unwanted events and side effects in cognitive behavior therapy. Cognit. Ther. Res. 42, 219–229.

Setbon, M., Raude, J., 2010. Factors in vaccination intention against the pandemic influenza A/H1N1. Eur. J. Publ. Health 20, 490–494.

Whipple, J.L., Lambert, M.J., Vermeersch, D.A., Smart, D.W., Nielsen, S.L., Hawkins, E.J., 2003. Improving the effects of psychotherapy: the use of early identification of treatment failure and problem-solving strategies in routine practice. J. Counsel. Psychol. 50, 59–68.

Williams, R., Farquharson, L., Palmer, L., Bassett, P., Clarke, J., Clark, D.M., Crawford, M.J., 2016. Patient preference in psychological treatment and associations with self-reported outcome: national cross-sectional survey in England and Wales. BMC Psychiatr. 16, 4.

Yao, L., Zhao, X., Xu, Z., Chen, Y., Liu, L., Feng, Q., Chen, F., 2020. Influencing factors and machine learning-based prediction of side effects in psychotherapy. Front. Psychiatr. 11, 537442.