Expert Review of Pharmacoeconomics & Outcomes Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ierp20

A roadmap for applying machine learning when working with privacy-sensitive data: predicting non-response to treatment for eating disorders

Vegard G Svendsen, Ben F. M. Wijnen, Jan Alexander De Vos, Ravian Veenstra, Silvia M. A. A. Evers & Joran Lokkerbol

To cite this article: Vegard G Svendsen, Ben F. M. Wijnen, Jan Alexander De Vos, Ravian Veenstra, Silvia M. A. A. Evers & Joran Lokkerbol (2023): A roadmap for applying machine learning when working with privacy-sensitive data: predicting non-response to treatment for eating disorders, Expert Review of Pharmacoeconomics & Outcomes Research, DOI: 10.1080/14737167.2023.2230368

To link to this article: https://doi.org/10.1080/14737167.2023.2230368

EXPERT REVIEW OF PHARMACOECONOMICS & OUTCOMES RESEARCH
https://doi.org/10.1080/14737167.2023.2230368

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

A roadmap for applying machine learning when working with privacy-sensitive data: predicting non-response to treatment for eating disorders

Vegard G Svendsen [a,b,c], Ben F. M. Wijnen [a,d], Jan Alexander De Vos [e], Ravian Veenstra[e], Silvia M. A. A. Evers [a,d] and Joran Lokkerbol [a]

[a]Center of Economic Evaluation & Machine Learning, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, The Netherlands; [b]Department of Health Services Research, Care and Public Health Research, CAPHRI, Maastricht University, Maastricht, The Netherlands; [c]Norwegian Centre for Addiction Research, SERAF, University of Oslo, Oslo, Norway; [d]Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, Maastricht, The Netherlands; [e]Human Concern, Centre for Eating Disorders, Human Concern, Amsterdam, The Netherlands

ABSTRACT

Objectives: Applying machine-learning methodology to clinical data could present a promising avenue for predicting outcomes in patients receiving treatment for psychiatric disorders. However, preserving privacy when working with patient data remains a critical concern.

Methods: In showcasing how machine-learning can be used to build a clinically relevant prediction model on clinical data, we apply two commonly used machine-learning algorithms (Random Forest and least absolute shrinkage and selection operator) to routine outcome monitoring data collected from 593 patients with eating disorders to predict absence of reliable improvement 12 months after entering outpatient treatment.

Results: An RF model trained on data collected at baseline and after three months made 31.3% fewer errors in predicting lack of reliable improvement at 12 months, in comparison with chance. Adding data from a six-month follow-up resulted in only marginal improvements to accuracy.

Conclusion: We were able to build and validate a model that could aid clinicians and researchers in more accurately predicting treatment response in patients with EDs. We also demonstrated how this could be done without compromising privacy. ML presents a promising approach to developing accurate prediction models for psychiatric disorders such as ED.

ARTICLE HISTORY
Received 9 December 2022
Accepted 23 June 2023

KEYWORDS
Eating disorder; machine learning; prediction; synthetic data; treatment success

1. Introduction

There is a need for more reliable, accurate and clinically validated prediction models in psychiatry [1,2]. In short, the overreliance on theory-driven approaches with the purpose of between-group *inference* at the level of specific predictors (i.e. a coefficient for independent variables in a regression analysis) might have expanded our theoretical understanding of psychiatric disorders, but found only limited applicability in clinical practice [3,4]. An alternative to prediction models built for *inference*, are models built for *accuracy* [5,6]. In contrast to theory-driven inference models, in which a hypothesis is tested and evaluated based on stringent statistical procedures, a model built for *accuracy* is typically data-driven, taking whatever relevant input is available, and ultimately assessed how well it performs in identifying or predicting whatever it was intended to identify or predict [7]. With increasing access to both rich patient data and powerful computers, such models can take immense quantities of data as input and use it to make low-cost, accurate and clinically useful predictions at the level of the individual patient.

For such purposes, the application of machine learning (ML) techniques has become increasingly popular [8], also for psychiatric research [9–11]. In a broad sense, ML is the study of tools and methods for identifying patterns in data, which can then help us make predictions about the future [12]. ML draws on concepts from various fields including statistics, computer science, and optimization, and is best conceptualized as an approach to data analysis – rather than merely a set of (statistical or computerized) tools. A central difference between ML and more traditional statistical procedures is that in a ML (or data-driven) approach, one starts with the data which outputs a model that can then be applied to new data, whereas in a traditional (theory-driven) approach one starts with the model assumptions that constitute the output [12].

1.1. Making clinical predictions in patients with eating disorders

ED are a considerable source of ill-health, poor quality of life and premature death [13,14]. A common and complicating feature of ED is that many patients, even after receiving best available treatment, do not display clinically meaningful improvement or remission [15–18]. Assuming that knowledge about likely treatment outcome could further augment a best-informed basis for considering potential adaptations to patients' treatment regime –

CONTACT Vegard G Svendsen v.g.svendsen@medisin.uio.no Norwegian Centre for Addiction Research, SERAF, University of Oslo, Building 45, Kirkeveien 166, Oslo 0450, Norway

it might be useful to identify likely treatment outcome as early as possible in the treatment course [19,20].

While to the best of our knowledge, longitudinal treatment non-response has not yet been modeled explicitly, several ML-based models have been developed to predict other clinical outcomes in patients with ED [21]. In one recent study, Haynos et al. [22] demonstrated how ML, in combination with patient report and interview data, could be used to longitudinally predict eating disorder symptomatology. In another, Espel-Huynh et al. [23] were able to predict short-term treatment response trajectories for a large sample of patients with ED based on clinical self-report data. Especially the latter study indicates a promising prelude to more advanced and complex models that are able to predict over a longer time period and in a population with relatively high symptomatology. Hence, it is hypothesized that ML is able to more quickly identify ED patients who are unlikely to respond to a standard course of treatment.

In both these studies, the authors were able to develop clinically relevant and potentially useful prediction models. However, the transferability of these and other ML models to other patients might be limited. For instance, models developed in highly specific patient populations or treatment regimens might not generalize well to other settings. Moreover, the amount, or type of information that constitute the input for many of these models might be so extensive or particular that their application outside the original context is not feasible.

1.2. Patient privacy and synthetic data

Assuming that most clinicians and researchers looking to develop their own prediction models might be unfamiliar with ML principles and how to apply them, this article will demonstrate how a clinician with access to clinical data, but with limited experience with programming and ML, can collaborate with colleagues from the disciplines of modeling and data science to develop a clinically useful prediction model. As this process will necessarily involve the sharing of patient data, we will also have to contend with the practical and ethical concerns of using sensitive and confidential data [24]. While privacy is a critical issue with all patient data, large datasets comprised of potentially large numbers of identifiable patient features (such as the batteries of psychometric questionnaires often adopted in ROM) might engender an even higher cause for concern and should be handled with the utmost care.

A promising approach for working with sensitive data while respecting privacy is the use of simulated or *synthetic* data [25]. Synthetic data aims to generate new, anonymous data from an original dataset while retaining the underlying associations between variables. As such, synthetic data can be used for purposes like data sharing, data exploration and secondary analysis [26,27] while reducing the risk of identity disclosure to near zero [28]. Applying ML to clinical data, while minimizing the risks of compromising privacy by using synthetic data, could represent a viable avenue for research in clinical samples such as patients with ED.

The goal of this study is to provide a showcase for how to adapt the typical prediction modeling process in the case of working with privacy-sensitive data using data from patients with eating disorders (ED). The present study demonstrates this by facilitating a pragmatic, step-wise ML scheme, in which we use anonymized, routinely collected patient-reported data from three different time points to predict the absence of reliable improvement in ED psychopathology 12 months after entering clinical treatment. The process is considered pragmatic in the sense that it will allow a primary ED researcher or clinician with access to clinical data, regardless of previous experience with ML or programming software, to – in tandem with a researcher trained in ML – train and validate a clinically useful ML prediction model without having to share any sensitive data.

2. Patients and methods

2.1. Data-analytic strategy

This study employed a ML (or data-driven) approach in which the dataset was split into a *training set*, which is used to develop the best model given the observations in this subset, and a *test set*, which is used to determine the performance of the trained models on unseen data (i.e. to validate the trained models as a large discrepancy in model performance between the training and test set could be an indication of overfitting of the model). The process of data preparation, model training and model validation is illustrated in Figure 1. Beginning in the stage of *data preparation*, the clinician shared with the researcher an overview of the variables included in the original clinical data, together with a description of which variables could be candidates for inclusion in a prediction model.

Using this information, and working in the programming software *R* the researcher then prepared a coding script for (1) recoding and restructuring the original data, (2) randomly splitting the original data into a *training set* (70%) and a *test set* (30%), (3) creating an anonymized (synthetic) copy of the original *training set* which still maintained the correlation structure of the data using the R package *synthpop* [26] (4) further ensuring confidentiality by removing any chance replication in the synthetic dataset of data combinations similar to those in the actual dataset (this was done using a built-in option in the *synthpop* package that does not require anyone to actually inspect the data), and (5) creating a visual output summary detailing the variable means and distributions of the anonymized training data in comparison with the actual training data. The clinician then applied this script to the original data, producing three separate datasets: a *training set* to be used for further model training using the original data, a synthetic copy of the training set (*anonymized training data*) to be forwarded to the researcher for further model training using synthetic data, and a *test set* to be kept unseen until the validation stage at the very end of the data-analytic process.

Before forwarding the anonymized (synthetic) training data to the researcher, its relative comparability to the original data was assessed by the clinician via inspection of the visual output summary (see (5) above). The purpose of this step was to establish some level of face validity of the synthetic data, and the analysis was continued only if the clinician (who in this

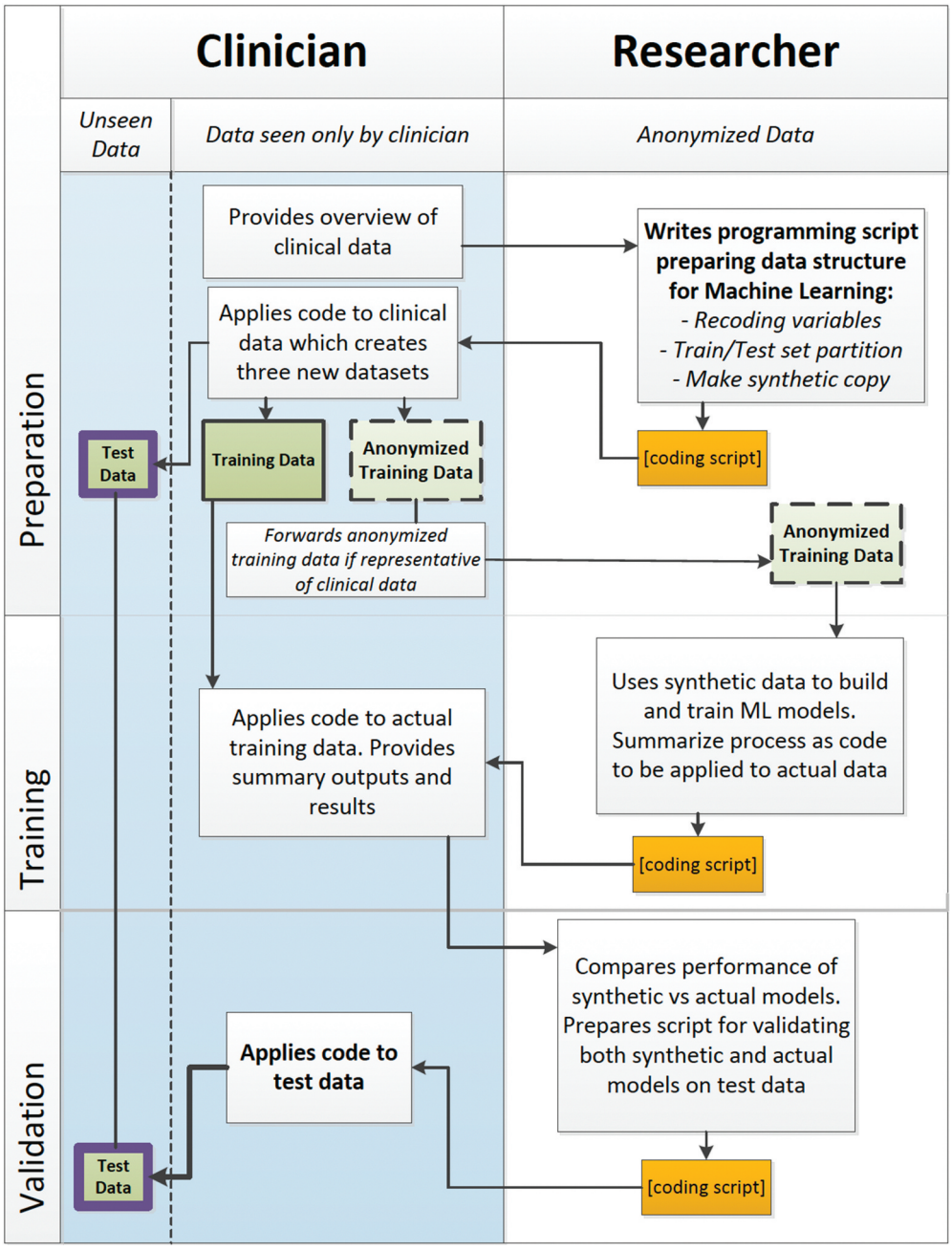


Figure 1. Illustration detailing the process of data preparation, model training and model validation when working with sensitive and synthetic data

case would be the best judge of the representativeness of the synthetic data) found the synthetic data to be a representative copy of the actual data. A more formal evaluation of the quality of synthetic data (relative to the actual data) would then come from comparing the results from the ML models built on either synthetic or actual data after having been applied to the *test set*.

The second stage, *model training*, involved the researcher using the anonymized training data and the R package *caret* [29], to train ML models and estimate their predictive performance with 10-fold cross-validation (for a further description of this process and its comparison to other validation strategies, see James et al., [30]). Models were trained to maximize positive predictive value (PPV: the probability that a patient identified as unlikely to improve at 12 months was correctly classified). The decision to use PPV as a performance metric was based on the clinical rationale that it was deemed more valuable to be certain about a predicted positive than about a predicted negative. Missing values were imputed with the median for continuous variables and classified as level 'unknown' for categorical variables.

In order to estimate the contribution of different strategies for model training and data inclusion, several models were built and compared for performance. Models were trained on either baseline data alone (baseline model), with the addition of data collected at three-months (three-month model) or with the addition of data collected at both three-months and six-months (six-month model). The baseline, three-month and six-

month models were also trained using two different ML techniques – resulting in six different models (3 × 2), which in turn were run on both the synthetic and the actual data.

While there are many modeling techniques applicable to ML [30], the techniques applied and compared for the purpose of this study were the *least absolute shrinkage and selection operator* (lasso) regression and *random forest* (RF). Lasso regression is a more robust alternative to ordinary least squares regression. By adding a penalty term that eventually reduces all coefficients to zero, lasso comes with a built-in feature elimination component which reduces complexity and therefore produces models that are less sensitive to overfitting and multicollinearity, and more interpretable compared to more complex modeling techniques, like RF [30,31].

RF is a tree-based ensemble technique – essentially a large collection of decision tree classifiers where each tree casts a vote for one of the two classes [32]. Unlike more rudimentary decision-tree techniques, where at each split all predictors are considered, trees in a random forest consider only a random subset – which reduces the dominance of some highly predictive variables in the model, thereby reducing correlation between the trees, and leading to a more generalizable model [30]. Our decision to include RF in our analysis is based on it being a relatively well-established and popular method in the field of predictive modeling.

Once the model code was developed and successfully run on the synthetic data, the coding script was sent to the clinician to be applied to the actual data training set. For the *validation stage* we first compared performance and the different model outputs from cross-validation. The best performing models, developed on both the synthetic and on the actual data, were eventually validated on the *test set* - which until this point had remained unseen both to the clinician and the researcher.

This strategy of ultimately validating and estimating model performance on a randomly selected hold-out *test set* ('unseen' by the trained model) would approximate the model performance on new data [3,4]. In this context, the performance on the *test set* of the models trained on the actual training data would measure the value of applying these models to make clinical predictions. The relative performance of the models trained on the synthetic data, when compared to those trained on the actual data, would provide a further measure of the representativeness, or specific utility, of the synthetic data.

2.2. Participants and procedure

The input for this model was ROM data collected from patients with ED receiving treatment at a specialized treatment center (Human Concern, center for EDs) in the Netherlands. Data was collected between March 2015 and August 2020. Inclusion criteria were as follows: 1) having a primary DSM-5 ED diagnosis at intake, 2) being 17 years or older, 3) being able to understand and fill in the questionnaires, and 4) consent to having their data used for research purposes. Exclusion criteria for treatment were as follows: 1) not being able to write and understand the Dutch language 2) severe and active automutilation, 3) active psychosis, 4) severe depression, 5) active suicidal ideation, and 6) acute somatic complications. Patients followed outpatient treatment with sessions once or twice a week with a psychologist or a clinician with lived experience of ED [33,34]. A combination of the following treatment modules was provided: insight-giving therapy, cognitive behavioral change, emotion regulation, and food/weight management. Treatment modules frequently involved support from other disciplines, such as system therapists, dieticians, or pharmacotherapy.

Patients filled in questionnaires every three months as part of their treatment and to monitor recovery. Data collected at the beginning of treatment (baseline), and during treatment (at 3, 6 and 12 months) were used for the purpose of this study. The informed consent was based on another study [35], approved by the Behavioural, Management and Social Sciences Ethics Committee of The University of Twente, in which patients could opt for the possibility of having their (anonymized) ROM data made available for other scientific studies. Patients who signed for this were included in this study.

2.3. Variables to predict treatment response

Questionnares Were all self-report, and covered demographic information, anamnesis and a wide range of psychometric instruments measured at varying timepoints. These were EDE-Q [36], *The Mental Health Continuum-ShortForm* (MHC-SF [37]), *The Outcome Questionnaire* (OQ-45 [38]), *The RAND 36-ItemHealth Survey* (SF-36 [39]), *Ryff's Scales of Psychological Well-being(Positieve Geestelijke Gezondheids Schaal;* PGGS [40]), *Forms of Self-Criticism/Attacking and Self-ReassuringScale* (FSCRS [41]), *Rosenberg Self-EsteemScale* (RSE [42]) and *The Personality Inventory for DSM-V* (PID-5 [43]) (see Appendix 1 for a more detailed overview).

2.4. Outcome of interest

Outcome of interest was non-response to treatment one year after starting treatment. Based on the 12-month global scale of the Eating Disorder Examination Questionnaire (EDE-Q) [36], a dichotomized (binary) outcome measure ('reliably improved' vs. 'not improved') was created. The cutoff value for distinguishing improved patients from those who did not improve was a reliable change score of >-1.41, obtained from a representative Dutch randomized controlled trial of ED patients [44]. Only patients with a recorded EDE-Q total score at 12 months were included in the analysis.

3. Results

3.1. Study population

Between March 2015 and August 2019, 1,089 patients had received at least one year of treatment and were therefore eligible for inclusion in the model. Of these patients, 131 did not have baseline measurements, and therefore could not be included in the study. Six patients were excluded for not meeting the diagnostic criteria of qualifying for a DSM-V ED.

Table 1. Patient characteristics.

Characteristics	Total *n* (%)
Gender	
female	582 (98%)
male	11 (2%)
ED diagnosis (%)	
AN	199 (33.7%)
BN	118 (20.0%)
BED	43 (7.3%)
OSFED	230 (39.0%)
Have previously received treatment	428 (72.5%)
No reliable improvement after 12 months	393 (66.6%)
	M (SD)
Age	27 (8.9)
Years duration of ED	9.9 (9.1)

ED = Eating Disorder, AN = Anorexia Nervosa, BN = Bulimia Nervosa, BED = Binge eating disorder, OSFED = Other specified eating disorder.

A further 227 patients were excluded for not having consented to research participation. Of the 719 still eligible patients, $n = 593$ had a complete Global EDE-Q score collected after the 12-months and were included in the study. As we could not know with any certainty why any patient did not have data collected at 12 months (and therefore was excluded), we did not pursue any comparison of characteristics between included or excluded patients. As can be seen in Table 1, most of the included participants were females who had previously received treatment for ED. The majority of patients (66.6%) did not meet the requirement for reliable change after 12 months and were considered as not having responded to treatment. The prevalence of non-reliable change (~66.6%) was uniform across the different subsets (test set, training set, and anonymized (synthetic) training set) of the original data.

3.2. Synthetic data properties

Based on visual inspection, variable means and their distributions were found to be comparable in the two sets of training data, suggesting that the general utility of the anonymized (synthetic) data was good (see Appendix 4 for an overview comparing variable distribution of actual vs synthetic data). Moreover, zero replications of unique data combinations were identified.

3.3. Model performance

Results from cross-validation on the training data suggested that baseline models produced only marginal predictive gain relative to prevalence. All baseline models were therefore abandoned and not validated on the test set. Of the models that were validated on the test-set (Table 2), models trained on actual data performed better than models trained on synthetic data. Of these models, the RF models performed slightly better than the lasso models. As seen by comparing the models highlighted in boldface, the best six-month model outperformed the best three-month model by an additional reduction in error rate of 2.6% and was overall the best performing model. The EDE-Q global scale from baseline to three months seemed to be the most important variable (see Appendix 2 and 3 for model outputs on lasso and RF).

4. Discussion

By applying ML methodology to clinical data, we were able to build and validate a model that improved substantially on chance prediction and could aid in more accurately predicting treatment response in patients with ED. In combination with clinical expert knowledge, such predictions could inform the development of personalized treatment plans and optimize treatment trajectories. Moreover, we have provided a showcase for how clinicians, with access to clinical data, could pursue similar efforts – without compromising privacy. Using routinely collected patient-reported data from baseline and three months after entering treatment, we were able to build a precise (PPV = 77%) and sensitive (sensitivity = 88%) model that performed over 31% better than chance in predicting which ED patients would be non-responsive to treatment after 12 months. While it was possible to improve predictive ability even further by including data collected at six months, this relatively small increase must be weighed against the advantage of being able to make predictions earlier. With this model, clinicians could already make, after three or six months, predictions at the patient level about which individuals are at significant risk of their current treatment regime not being effective. Such predictions could also inform the development of personalized treatment plans.

In general, it is our belief that the prediction modeling efforts reported here only scratch the proverbial surface of the opportunities offered by ML. First, it is possible that by including other sources of data (i.e. open text fields from the EDE-Q or daily reports by clinicians), our model might have performed even better. In addition, some research has indicated that biological markers, such as immunological differences, may also play a crucial role underlying mental disorders [45]. On the other hand, it should also be noted that complexity does not necessarily generate more accuracy [7,23] and that a simpler model also could have been pursued.

Second, our modeling strategy was subject to several methodological decisions – based largely on considerations of what was deemed appropriate for the current

Table 2. Performance of all models on test-set.

		Three-month models			*Six-month models*		
Model Type	Data used to train model	PPV	Sensitivity	Error rate reduction[†]	PPV	Sensitivity	Error rate reduction[†]
Lasso	Actual	.767	0.846	30.7%	.764	0.940	29.8%
RF	**Actual**	.769	**0.880**	**31.3%**	.778	**0.897**	**33.9%**
Lasso	Synthetic	.711	0.778	14.0%	.675	0.675	3,3%
RF	Synthetic	.703	0.949	11.6%	.710	0.940	13.7%

[†]Relative to a base error rate of 1 – prevalence.

purpose. For instance, our choice of PPV as the primary model metric was based on a wish to employ a measure of model performance that was both intuitive and clinically meaningful. Other decisions, such as limiting the number of modeling techniques to two, and not incorporating computationally heavy procedures for imputing missing values or performing nested cross-validation [30], were largely out of practical considerations. For every strategic and methodological aspect of the modeling effort presented in this paper, we could have chosen differently, and consequently, received different results. We would encourage readers considering similar projects to take this into consideration and to pursue context appropriate strategies for assembling and training their own ML models.

4.1. Strengths & limitations

The purpose of this article was to provide a roadmap for how to apply ML in clinical prediction modeling while also conserving privacy. To do this, we have used clinical data from a sample of patients with ED to build a model that aims to predict lack of treatment response one year into the future. The demonstration of synthetic data as a viable technology for anonymizing sensitive information is a potential strength of this study. The use of synthetic data facilitated the sharing of confidential data between two groups of researchers, where one had access to clinical data and the other provided technical competence – such that we could develop and debug coding scripts before applying them to actual data. While the utility of synthetic data has previously been established on a variety of clinical datasets [25,27], there is, to the best of our knowledge, a paucity of evidence in regarding the use of synthetic data for the purpose of training psychiatric ML models. As such, our efforts here should be considered exploratory. Although the synthetic data in general was found to be highly representative of the original data (as evident in the comparison of variable distributions and overall performance of synthetic models on the test-set), we had made no targeted effort at maximizing the utility of the synthetic dataset. Therefore, the synthetic data was used solely as a means for anonymization and had no influence on our analytic strategy or on the reading of results. If future researchers want to be more confident in the performance of models built on synthetic data and less dependent on validating their results on the primary data source, further steps might be required in the direction of increasing the utility of the synthetic dataset [27].

This model is subject to several important limitations. First, the input used for this model was whatever relevant data was available to us at the time the prediction was made – and not the result of a process guided by previous research. This greatly limits the opportunities for assigning causal associations between variables. In general, readers should be mindful that ML models – being a data-driven approach where all model features are governed by algorithmic operations and not the theoretical assumptions of the researcher – typically do not allow for causal inference [46], which was neither the purpose nor within the scope of this study.

Second, as with any clinical tool, this model is only as valuable as the extent to which it useful in clinical practice. Its utility must also be assessed in light of whatever action is associated with being correctly identified by the model – or the consequence of getting this identification wrong. Even with the most precise model presented here, about 22% of those identified as not likely to respond to treatment will actually respond. However, it should also be noted that even perfect knowledge about the future outcome does not tell us anything about how to change it – and that even an infallible model will therefore not necessarily lead to better clinical care [47].

Third, reliable change on the EDE-Q might not be the only relevant measure of treatment outcome or meaningful improvement for patients with ED. Other constructs, such as psychological well-being [48] might be equally important, especially from the patient's perspective. Thus, to equate lack of change in ED symptomology with lack of response to treatment might not be entirely accurate, as response to treatment is arguably more nuanced than what might be captured by the EDE-Q. The lack of uniform definitions of clinical outcomes and of response to treatment in ED research remains a significant issue [49,50], and future prediction model efforts should seek to harmonize operationalizations with whatever consensus exists in the literature. By focusing solely on change on the EDE-Q, it is possible that other (relevant) effects are missed. Moreover, it is also likely that our choice of a binary rather than a continuous outcome measure failed to capture potentially important nuance in patients' treatment response trajectories.

Fourth, in our analysis we did not explore the opportunity to increase prediction certainty by changing the probability threshold of the model [51]. As with most statistical procedures involving a dichotomous outcome, our model uses a default classification probability cutoff of 0.5. Raising this threshold would likely lead to increased precision, albeit with a cost to sensitivity. In clinical terms, this means that a clinician who is using our model to identify patients at risk of not improving – but who would prefer higher certainty, even at the risk of failing to detect some cases (false negatives), could attain such predictions by making only minor adjustments to the code.

Fifth, results presented in this article are mainly for the purpose of demonstration and are not intended to generalize beyond the confines of this research project. Importantly, it was also not an objective of this project to evaluate the effect or quality of the treatments described above. As already noted, the outcome measure adopted for this study is subject to limitations such that it might not represent an appropriate or generalizable

estimate of lack of treatment response. Patient and treatment characteristic are described solely for purpose of transparency.

Sixth, it is possible that an inter-diagnostic modeling approach, rather than grouping all eating disorders together, might have been more appropriate and allowed for more precise predictions. While we did not have enough data to stratify based on diagnosis, this might be a point of consideration for future research. For example, it could be that a prognostic model solely build using patients with anorexia results in better predictions for those patients given that the underlying observations may be more homogenous.

5. Conclusion

We have presented a pragmatic approach for using ML to build a clinically relevant prediction model from clinical data, without compromising privacy. Using only routinely collected patient-reported data, we were able to build and validate an accurate prediction to predict absence of response to treatment after one year. This model allows clinicians to predict which ED patients have an elevated risk of the current treatment regime not being effective in an early stage (i.e. three months after start of treatment). Predictions made with this model could inform treatment decisions and aid clinicians in the development of personalized treatment plans. It is our hope that the approach presented above can stimulate the development of prediction models in clinical settings while ensuring patients' privacy.

Funding

One author is supported by The Netherlands Organisation for Health Research and Development (ZonMw) Mental Healthcare Fellowship.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Data availability statement

The informed consent was based on another study (de Vos et al., 2021), approved by the Behavioural, Management and Social Sciences Ethics Committee of the University of Twente, in which patients could opt for the possibility to include their ROM data (anonymized) for other scientific studies. Patients who signed for this were included in this study. Due to privacy regulation, the data cannot be made publicly available.

Author contributions statement

V Svendsen, J Lokkerbol and B Wijnen were involved in all stages of developing the code, prognostic model, interpretation and writing of the article.

J A DeVos and R Veenstra were involved in generating synthetic data and running analyses on the patient data. Moreover, they were responsible for clinical interpretation of the results and writing of the manuscript.

ORCID

Vegard G Svendsen http://orcid.org/0000-0001-5802-7482
Ben F. M. Wijnen http://orcid.org/0000-0001-7993-1905
Jan Alexander De Vos http://orcid.org/0000-0001-8611-0777
Silvia M. A. A. Evers http://orcid.org/0000-0003-1026-570X
Joran Lokkerbol http://orcid.org/0000-0001-9949-5442

References

Papers of special note have been highlighted as: (•) of interest (••) of considerable interest.

1. Koppe G, Meyer-Lindenberg A, Durstewitz D. Deep learning for small and big data in psychiatry. Neuro psycho pharmacol. 2021;46(1):176–190. doi: 10.1038/s41386-020-0767-z
2. McMahon FJ. Prediction of treatment outcomes in psychiatry—where do we stand? Dialogues Clin Neurosci. 2014;16(4):455. doi: 10.31887/DCNS.2014.16.4/fmcmahon
3. Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol. 2018;14(1):91–118. doi: 10.1146/annurev-clinpsy-032816-045037

•• **Valuable introductory paper to the topic of machine learning for clinical practise**

4. Cearns M, Hahn T, Baune BT. Recommendations and future directions for supervised machine learning in psychiatry. Transl Psychiatry. 2019;9(1):1–12. doi: 10.1038/s41398-019-0607-2

•• **Valuable introductory paper to the topic of machine learning for clinical practise**

5. Kuhn M, Johnson K. Introduction. applied predictive modelling. (NY): Springer New York; 2013. p. 1–16. doi: 10.1007/978-1-4614-6849-3_1
6. Poldrack RA, Huckins G, Varoquaux G. Establishment of best practices for evidence for prediction: a review. JAMA Psychiatry. 2020 May 1;77(5):534–540. doi: 10.1001/jamapsychiatry.2019.3671
7. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, et al. Implementing precision psychiatry: a systematic review of individualized prediction models for clinical practice. Schizophrenia Bulletin. 2021 Mar 16;47(2):284–297. doi: 10.1093/schbul/sbaa120
8. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349(6245):255–260. doi: 10.1126/science.aaa8415
9. Rutledge RB, Chekroud AM, Huys QJ. Machine learning and big data in psychiatry: toward clinical applications. Curr Opin Neurobiol. 2019;55:152–159. doi: 10.1016/j.conb.2019.02.006
10. Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. Psychol Med change. 2016;46(12):2455–2465. doi: 10.1017/S0033291716001367
11. Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3(3):243–250. doi: 10.1016/S2215-0366(15)00471-X
12. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Clin Infect Dis. 2018 Jan 6;66(1):149–153. doi: 10.1093/cid/cix731
13. Andersen AEDiagnosis and treatment of the eating disorders spectrum in primary care medicine. In: Mehler PS, and Anderson AE, editor. Eating disorders: a guide to medical care and complications 3 ed. Baltimore, Maryland, USA: Johns Hopkins University Press; 2017, p. 1–55.
14. Klump KL, Bulik CM, Kaye WH, et al. Academy for eating disorders position paper: eating disorders are serious mental illnesses. Int J Eat Disord. 2009;42(2):97–103. doi: 10.1002/eat.20589
15. Halmi KA. Perplexities of treatment resistance in eating disorders. BMC Psychiatry. 2013;13(1):1–6. doi: 10.1186/1471-244X-13-292
16. Fassino S, Abbate-Daga G. Resistance to treatment in eating disorders: a critical challenge. BMC Psychiatry. 2013 Nov 7;13(1):282. doi: 10.1186/1471-244X-13-282
17. Halmi KA, Agras WS, Crow S, et al. Predictors of treatment acceptance and completion in anorexia nervosa: implications for future study designs. Arch Gen Psychiatry. 2005 Jul;62(7):776–781.

18. Linardon J, Wade TD. How many individuals achieve symptom abstinence following psychological treatments for bulimia nervosa? A meta-analytic review. Int J Eat Disord. 2018 Apr;51(4):287–294. doi: 10.1002/eat.22838
19. Le Grange D, Accurso EC, Lock J, et al. Early weight gain predicts outcome in two treatments for adolescent anorexia nervosa. Int J Eat Disord. 2014 Mar;47(2):124–129.
20. Linardon J, de la Piedad Garcia X, Brennan L. Predictors, moderators, and mediators of treatment outcome following manualised cognitive-behavioural therapy for eating disorders: a systematic review. Eur Eat Disorders Rev. 2017 Jan;25(1):3–12. doi: 10.1002/erv.2492
21. Wang SB. Machine learning to advance the prediction, prevention and treatment of eating disorders. Euro Eating Disorders Rev. 2021 Sep;29(5):683–691. doi: 10.1002/erv.2850
22. Haynos AF, Wang SB, Lipson S, et al.Machine learning enhances prediction of illness course: a longitudinal study in eating disorders. Psychol Med. 2021;58(8):1392–1402. doi:10.1017/S0033291720000227
• **Landmark paper which inspired our own research**
23. Espel-Huynh H, Zhang F, Thomas JG, et al. Prediction of eating disorder treatment response trajectories via machine learning does not improve performance versus a simpler regression approach. Int J Eat Disord. 2021 Jul;54(7):1250–1259.
24. Balthazar P, Harri P, Prater A, et al. Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. J Am Coll Radiol. 2018;15(3):580–586. doi: 10.1016/j.jacr.2017.11.035
25. Quintana DS. A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation. Elife. 2020;9:e53275.
•• **Valuable introductory paper to the topic of synthetic data for clinical research**
26. Nowok B, Raab GM, Dibben C. Synthpop: bespoke creation of synthetic data in R. J Stat Soft. 2016;74(11):1–26. doi: 10.18637/jss.v074.i11
27. Emam KE, Mosquera L, Zheng C. Optimizing the synthesis of clinical trial data using sequential trees. J Am Med Inform Assoc. 2021 Jan 15;28(1):3–13. doi: 10.1093/jamia/ocaa249
28. El Emam K, Mosquera L, Bass J. Evaluating identity disclosure risk in fully synthetic health data: model development and validation. J Med Internet Res. 2020;22(11):e23139. doi: 10.2196/23139
29. Kuhn M. A Short Introduction to the caret Package. R Found Stat Comput. 2013. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=768f3fead8440788ce69dedf67220de93ab2bacd
30. James G, Witten D, Hastie T, et al. An introduction to statistical learning: with applications in R. (NY): Springer New York; 2013. doi: 10.1007/978-1-4614-7138-7
31. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Stat Soc Ser B (Methodological). 1996;58(1):267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
32. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32. doi: 10.1023/A:1010933404324
33. De Vos JA, Netten C, Noordenbos G. Recovered eating disorder therapists using their experiential knowledge in therapy: a qualitative examination of the therapists' and the patients' view. Eat Disord. 2016;24(3):207–223. doi: 10.1080/10640266.2015.1090869
34. De Vos JA, Radstaak M, Bohlmeijer ET, et al. Having an eating disorder and still being able to flourish? Examination of pathological symptoms and well-being as two continua of mental health in a clinical sample. Front Psychol. 2018;9:2145. doi:10.3389/fpsyg.2018.02145
35. de VosJA, Radstaak M, Bohlmeijer ET, et al. The psychometric network structure of mental health in eating disorder patients. Eur Eat Disorders Rev. 2021;29(4):559–574. doi: 10.1002/erv.2832
36. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord. 1994;16(4):363–370.
37. Lamers SM, Glas CA, Westerhof GJ, et al. Longitudinal evaluation of the mental health continuum-short form (MHC-SF). Eur J Psychol Assess. 2012;28(4):290–296. doi: 10.1027/1015-5759/a000109
38. Lambert MJ, Burlingame GM, Umphress V, et al. The reliability and validity of the outcome questionnaire. Clin Psychol Psychother. 1996;3(4):249–258. doi: 10.1002/(SICI)1099-0879(199612)3:4<249:AID-CPP106>3.0.CO;2-S
39. Vander Zee KI, Sanderman R, Heyink JW, et al. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. Int J Behav Med. 1996;3(2):104–122. doi: 10.1207/s15327558ijbm0302_2
40. Van Dierendonck D. The construct validity of Ryff's scales of psychological well-being and its extension with spiritual well-being. Pers Individ Dif. 2004;36(3):629–643. doi: 10.1016/S0191-8869(03)00122-3
41. Gilbert P, Clarke M, Hempel S, et al. Criticizing and reassuring oneself: an exploration of forms, styles and reasons in female students. Br J Clin Psychol. 2004;43(1):31–50. doi: 10.1348/014466504772812959
42. Rosenberg M. Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy Measures package. Measures Package. 1965;61(52):18.
43. Krueger RF, Derringer J, Markon KE, et al. Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychol Med. 2012;42(9):1879. doi: 10.1017/S0033291711002674
44. de JongM, Spinhoven P, Korrelboom K, et al. Effectiveness of enhanced cognitive behavior therapy for eating disorders: a randomized controlled trial. Int J Eat Disord. 2020 May;53(5):447–457.
45. Serafini G, Parisi VM, Aguglia A, et al. A specific inflammatory profile underlying suicide risk? Systematic review of the main literature findings. Int J Environ Res Public Health. 2020 Apr 1;17(7):2393. doi: 10.3390/ijerph17072393
46. Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nature Mach Intell. 2020;2(7):369–375. doi: 10.1038/s42256-020-0197-y
47. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. N Engl J Med. 2017;376(26):2507. doi: 10.1056/NEJMp1702071
48. Tomba E, Tecuta L, Schumann R, et al. Does psychological well-being change following treatment? An exploratory study on outpatients with eating disorders. Compr Psychiatry. 2017;74:61–69. doi: 10.1016/j.comppsych.2017.01.001
49. Bardone-Cone AM, Hunt RA, Watson HJ. An overview of conceptualizations of eating disorder recovery, recent findings, and future directions. Curr Psychiatry Rep. 2018 8;20(9):79. doi: 10.1007/s11920-018-0932-9
50. Vall E, Wade TD. Predictors of treatment outcome in individuals with eating disorders: a systematic review and meta-analysis. Int J Eat Disord. 2015;48(7):946–971. doi: 10.1002/eat.22411
51. Kuhn M, Johnson K. Remedies for severe class imbalance. Applied predictive modeling. (NY): Springer New York; 2013. p. 419–443. doi: 10.1007/978-1-4614-6849-3_16

Appendix 1: Psychometric Instruments

Name *Description*	Items	Baseline	3 Months	6 Months
Demographic variables *Background characteristics*	Age, gender, education level, living situation	•		
Anamnesis *Disease history*	Primary diagnosis, earlier ED diagnosis, age at onset of ED, duration ED, previous treatment type, personality disorder, comorbid psychiatry, BMI (factorized as healthy/unhealthy*) * According to CDC cut-offs https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html)	•		
Self-developed items	Regulation of emotion	•	•	
Self-developed Instruments developed for use in clinical practice	Motivation, engagement, capacity for self-reflection, hope, capacity for problem solving, mental resilience, family functioning, meaningful daily activities, therapeutic alliance, evaluation of choice of therapeutic method after 3 months, illness insight, social desirability		•	
The 36-Item Eating Disorder Examination Questionnaire (**EDE-Q**) (Fairburn & Beglin, 1994) *Measures the severity and range of features of eating disorders*	Global Score	•	•	•
	Eating Concern	•	•	•
	Weight Concern	•	•	•
	Shape Concern	•	•	•
	Restraint	•	•	•
	Behavioural Items (Binge eating, vomiting, laxative use, exercise)	•	•	•
The Mental Health Continuum-Short Form (**MHC-SF)** (Lamers et al., 2012) *Measures social, psychological, and emotional well-being*	Total Score	•	•	•
	Psychological, Emotional and Social Wellbeing	•	•	•
The Outcome Questionnaire (**OQ-45**) (Lambert et al., 1996) *Measures patient outcome from routine psychological treatment*	Total Psychopathology Score	•	•	•
	Anxiety Somatic Distress, Social Role, Interpersonal Functioning, Symptomatic Distress	•	•	•
	Nervousness, Depression, Anger, Feeling Weak, Suicidal Thoughts, Irritation, Fatigue, Feeling Worthless	•	•	
The RAND 36-Item Health Survey (**SF-36**) (Vander Zee et al., 1996) *A multidimensional measure of general health status*	All (9) scales	•		
Positieve Geestelijke Gezondheids Schaal (**PGGS**) (Van Dierendonck, 2004) *Measures multiple facets of psychological well-being*	All (6) scales	•		
Forms of Self-Criticism/Attacking and Self-Reassuring Scale (**FSCRS**) (Gilbert et al., 2004) *Measures reaction patterns to adversity*	All (3) scales	•	•	
Rosenberg Self-Esteem Scale (**RSE**) (Rosenberg, 1965) *Measures self-esteem*	Total Score	•	•	
The Personality Inventory for DSM-V (**PID-V**) (Krueger et al., 2012) Measures pathological personality trait facets and the five higher-order domains of criterion B of the DSM-5 alternative model of personality disorders	All (30) scales	•		

Appendix 2: Overview of variables included in final lasso model (variables reduced to zero not shown)

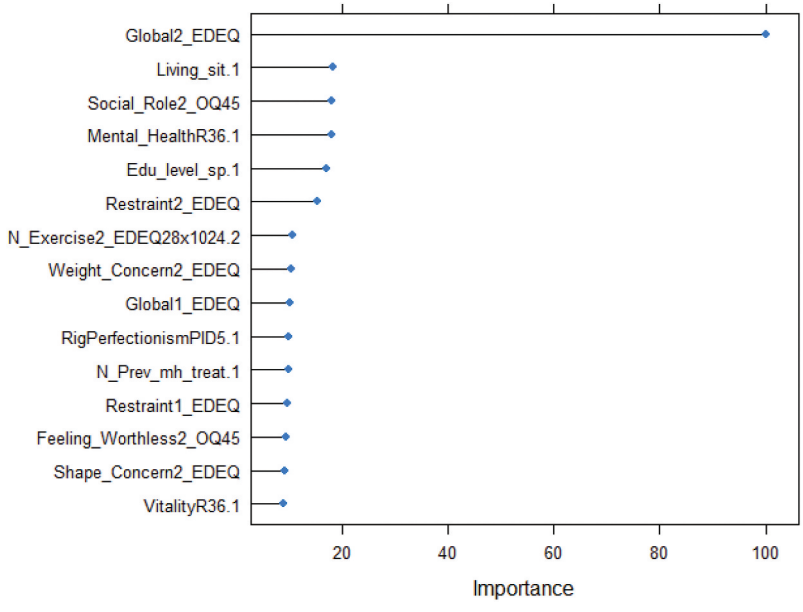
(Intercept)	5.411063e-01
Inadequate_Self_1_FSCRS	1.542612e-01
Fatigue1_OQ45.Always	−1.103295e-02
Suicidal_Thoughts1_OQ45.Always	5.055107e-02
Suicidal_Thoughts1_OQ45.Sometimes	−1.013331e-02
Feeling_Weak1_OQ45.Always	7.294666e-02
Feeling_Weak1_OQ45.Sometimes	−5.020752e-02
Feeling_Worthless1_OQ45.Rarely	−2.019199e-02
Nervousness1_OQ45.Always	−1.653661e-01
Nervousness1_OQ45.Never	1.234571e-01
Depression1_OQ45.Frequently	−9.742203e-03
Depression1_OQ45.Sometimes	5.152806e-02
Anger_WorkOrSchool1_OQ45.Always	7.225721e-02
Anger_WorkOrSchool1_OQ45.Sometimes	1.792683e-02
N_Objective_Binges1_EDEQ18×1014.1	−6.694530e-04
Subjective_Binges1_EDEQ19×1015.20	−1.111982e-02
Vomiting1_EDEQ21×1017.Unknown	−2.524368e-02
Laxative_Use1_EDEQ23×1019.No	2.978892e-02
Restraint1_EDEQ	8.933232e-02
Shape_Concern1_EDEQ	2.817078e-01
N_Prev_mh_treat.1.1	3.619566e-02
N_Prev_mh_treat.1.2	−6.967344e-02
N_Prev_mh_treat.1.3	−5.110160e-02
Duration_ED.1	−1.421482e-01
Personal_GrowthPGGS.1	−2.296848e-01
Social_FunctioningR36.1	5.370137e-03
Physical_RoleFunctioningR36.1	−1.285729e-01
Mental_HealthR36.1	−7.593206e-02
General_HealthR36.1	−7.681904e-02
AttentionseekPID5.1	−6.893865e-02
DepressivityPID5.1	−1.361861e-01
ImpulsivityPID5.1	−2.823948e-02
IntimityAvPID5.1	−7.239059e-02
PerceptDysregPID5.1	1.500116e-01
SubmissivenessPID5.1	−1.206813e-01
SuspiciousnessPID5.1	7.641579e-02
Gender.1.Famale	−3.306078e-02
Gender.1.Male	4.321643e-14
Edu_level.1.Unknown	1.307143e-01
Edu_level_sp.1.Hig_3rd	−1.241573e-01
Edu_level_sp.1.Sec_1st_H	−1.646539e-01
Edu_level_sp.1.Sec_1st_M	−1.338520e-01
Edu_level_sp.1.Sec_2nd_L	−1.015152e-01
Edu_level_sp.1.Sec_2nd_M	6.593350e-02
Living_sit.1.Children.No.Partner	−3.711511e-02
Living_sit.1.Mental.health.institution	1.587465e-03
Living_sit.1.No.children.Partner	2.259141e-01
Living_sit.1.Unknown	−2.007560e-02
physical_prob.1.No	2.131590e-01
Weight_anxiety.1.Unknown	1.507781e-01
Weight_anxiety.1.Yes	−1.025173e-13
Earlier_ED_diagnosis.1.Yes	6.273201e-02
Reassured_Self_2_FSCRS	3.303012e-01
Hated_Self_2_FSCRS	−8.108811e-02
Self_Esteem_2_RosenbergTot	−3.005930e-02
SocialWellbeing2_MHCSF	−7.227510e-03
Fatigue2_OQ45.Unknown	2.166253e-02

(*Continued*)

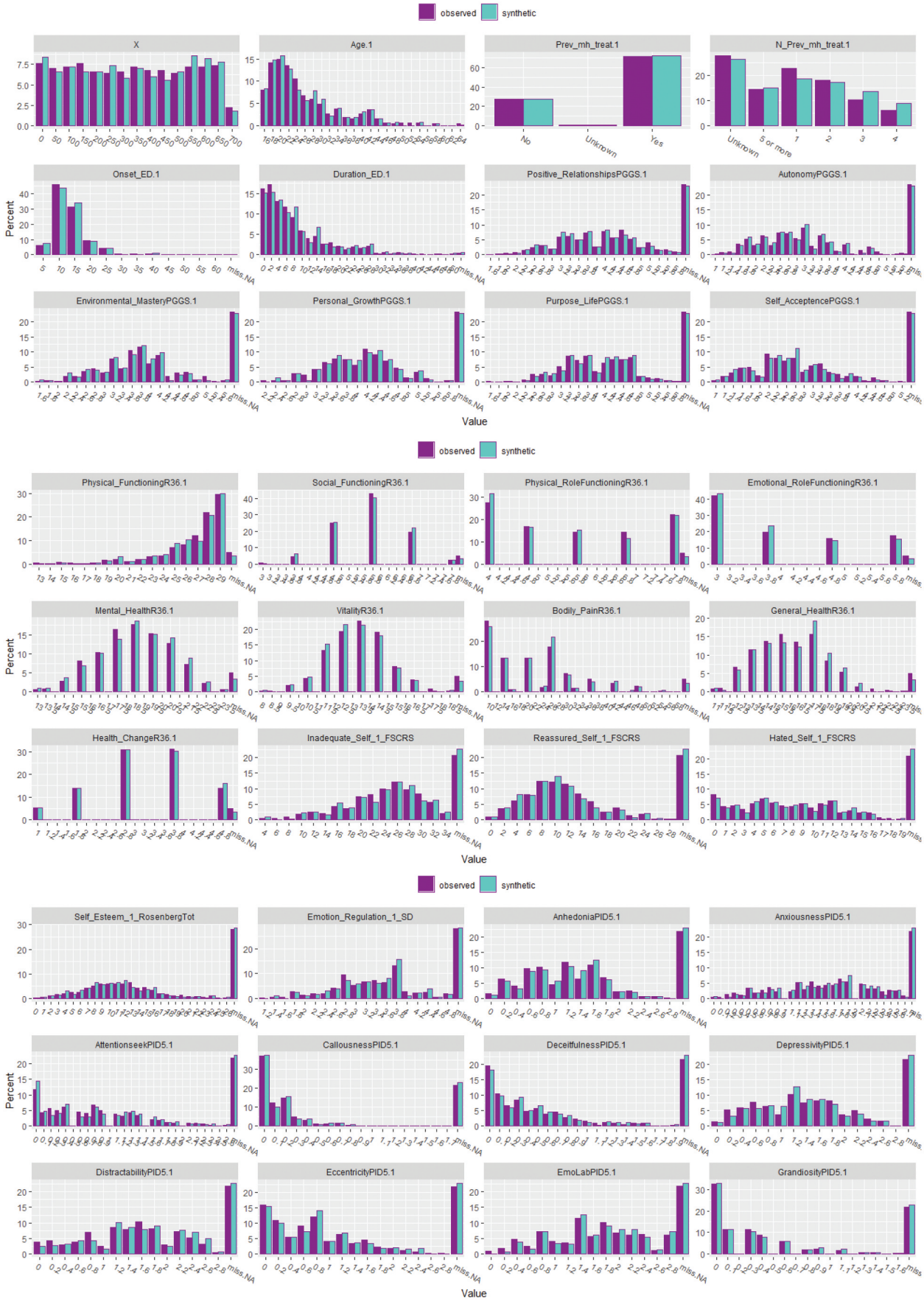
(Continued).

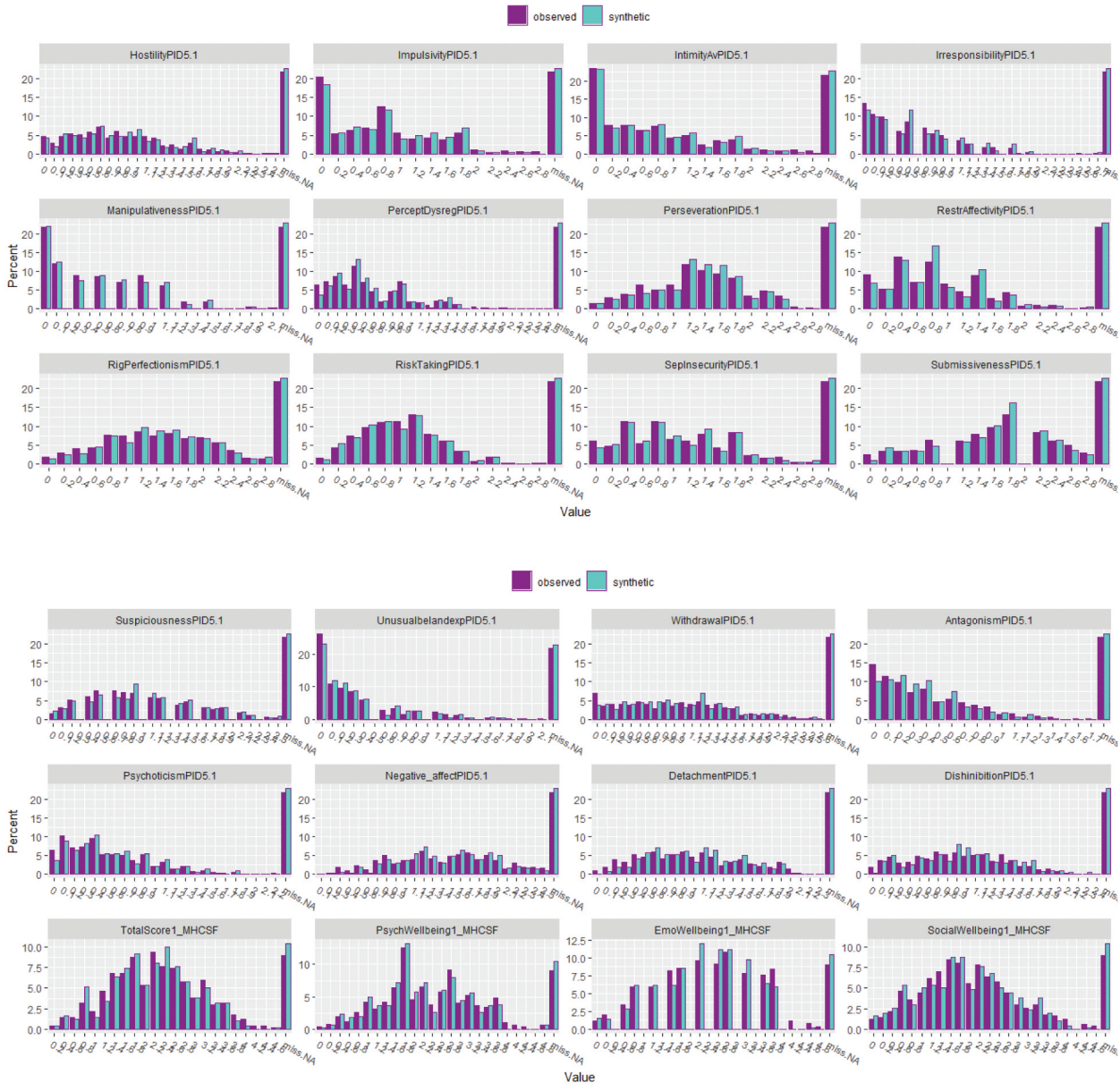
(Intercept)	5.411063e-01
Irritation2_OQ45.Never	−5.874316e-03
Suicidal_Thoughts2_OQ45.Rarely	1.664509e-02
Suicidal_Thoughts2_OQ45.Sometimes	−1.850063e-02
Suicidal_Thoughts2_OQ45.Unknown	7.470395e-04
Feeling_Weak2_OQ45.Always	−1.286647e-01
Feeling_Weak2_OQ45.Never	6.846640e-02
Feeling_Weak2_OQ45.Rarely	−2.680482e-01
Feeling_Weak2_OQ45.Unknown	3.218958e-05
Feeling_Worthless2_OQ45.Always	−1.214328e-01
Feeling_Worthless2_OQ45.Never	−2.942820e-01
Feeling_Worthless2_OQ45.Rarely	2.231465e-01
Feeling_Worthless2_OQ45.Sometimes	−4.327187e-02
Nervousness2_OQ45.Always	−3.180732e-03
Depression2_OQ45.Always	−9.364883e-02
Anger_WorkOrSchool2_OQ45.Never	−5.305011e-03
Interpersonal_Functioning2_OQ45	−1.508011e-02
Social_Role2_OQ45	−2.158416e-01
Subjective_Binges2_EDEQ19×1015.5	−2.622176e-02
Subjective_Binges2_EDEQ19×1015.No	−1.968185e-01
N_Vomiting2_EDEQ22×1018.2	−3.357886e-02
Laxative_Use2_EDEQ23×1019.Yes	−9.237342e-02
Shape_Concern2_EDEQ	−6.776423e-02
Global2_EDEQ	−8.350392e-01

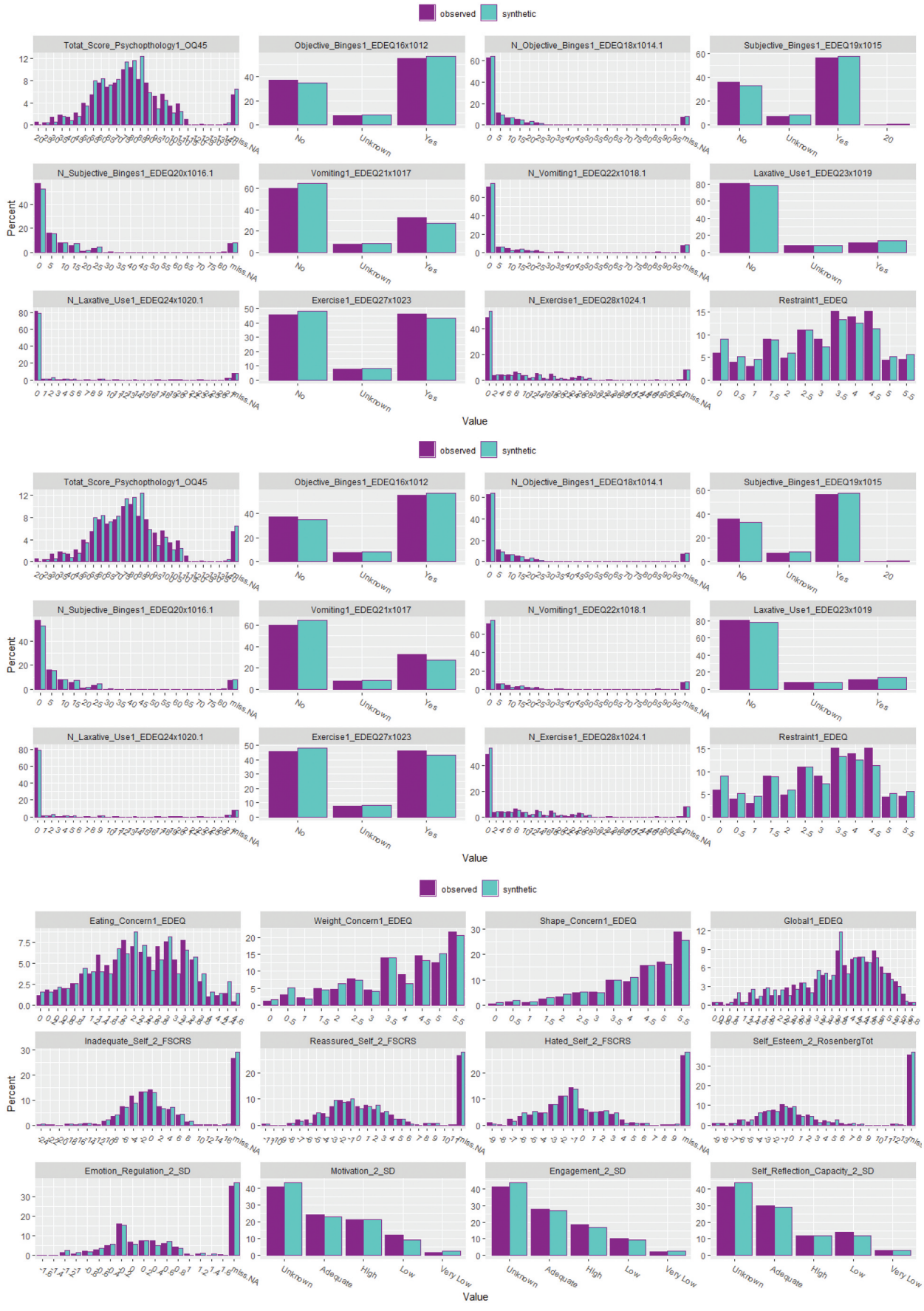
Appendix 3: Most important variables in final random forest model



Appendix 4: Comparison of score distributions of observed and synthetic dataset

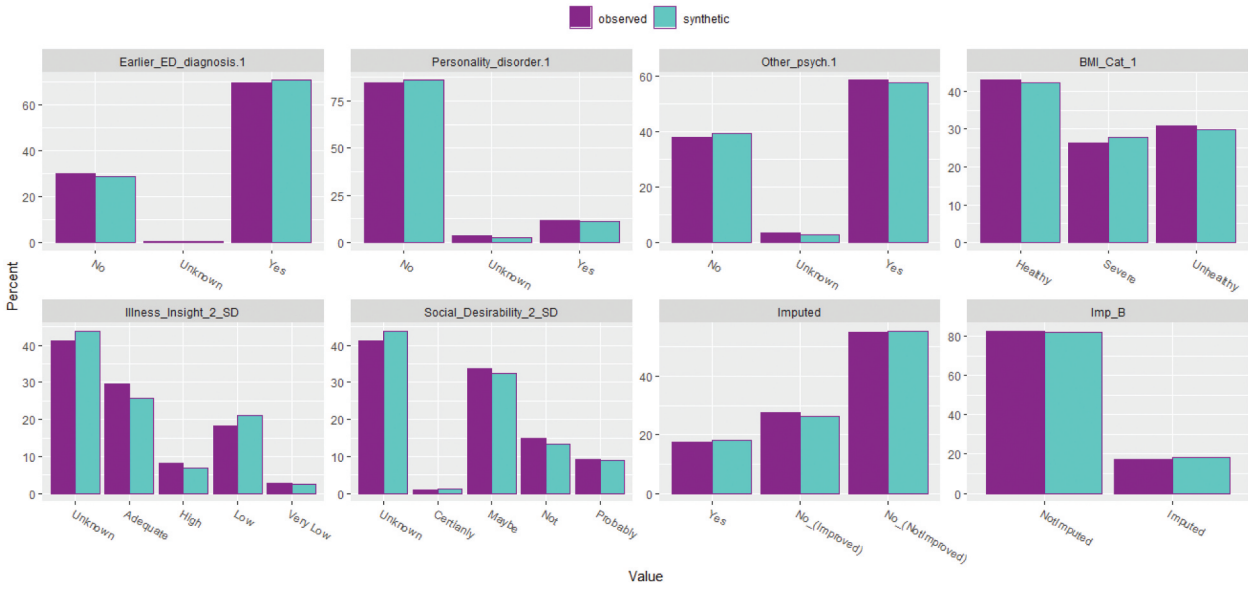



observed
synthetic
HostilityPID5.1
ImpulsivityPID5.1
IntimityAvPID5.1
IrresponsibilityPID5.1
ManipulativenessPID5.1
PerceptDysregPID5.1
PerseverationPID5.1
RestrAffectivityPID5.1
RigPerfectionismPID5.1
RiskTakingPID5.1
SepInsecurityPID5.1
SubmissivenessPID5.1
Percent
Value
observed
synthetic
SuspiciousnessPID5.1
UnusualbelandexpPID5.1
WithdrawalPID5.1
AntagonismPID5.1
PsychoticismPID5.1
Negative_affectPID5.1
DetachmentPID5.1
DishinibitionPID5.1
TotalScore1_MHCSF
PsychWellbeing1_MHCSF
EmoWellbeing1_MHCSF
SocialWellbeing1_MHCSF
Percent
Value


observed
synthetic
Totat_Score_Psychopthology1_OQ45
Objective_Binges1_EDEQ16x1012
N_Objective_Binges1_EDEQ18x1014.1
Subjective_Binges1_EDEQ19x1015
N_Subjective_Binges1_EDEQ20x1016.1
Vomiting1_EDEQ21x1017
N_Vomiting1_EDEQ22x1018.1
Laxative_Use1_EDEQ23x1019
N_Laxative_Use1_EDEQ24x1020.1
Exercise1_EDEQ27x1023
N_Exercise1_EDEQ28x1024.1
Restraint1_EDEQ
Eating_Concern1_EDEQ
Weight_Concern1_EDEQ
Shape_Concern1_EDEQ
Global1_EDEQ
Inadequate_Self_2_FSCRS
Reassured_Self_2_FSCRS
Hated_Self_2_FSCRS
Self_Esteem_2_RosenbergTot
Emotion_Regulation_2_SD
Motivation_2_SD
Engagement_2_SD
Self_Reflection_Capacity_2_SD
Percent
Value
No
Unknown
Yes
miss.NA
Unknown
Adequate
High
Low
Very Low


observed
synthetic
Hope_2_SD
Problem_Solving_Capacity_2_SD
Mental_Resilience_2_SD
Familiy_Functioning_2_SD
Meaningfull_Daily_Activities_2_SD
Therapeutic_Alliance_2_SD
Appropriate_TherapeuticMeth_2_SD
TotalScore2_MHCSF
PsychWellbeing2_MHCSF
EmoWellbeing2_MHCSF
SocialWellbeing2_MHCSF
Fatigue2_OQ45
Percent
Value
Irritation2_OQ45
Suicidal_Thoughts2_OQ45
Feeling_Weak2_OQ45
Feeling_Worthless2_OQ45
Nervousness2_OQ45
Depression2_OQ45
Anger_WorkOrSchool2_OQ45
Symptomatic_Distress2_OQ45
Interpersonal_Functioning2_OQ45
Social_Role2_OQ45
Anxiety_Somatic_Distress2_OQ45
Totat_Score_Psychopthology2_OQ45
Objective_Binges2_EDEQ16x1012
N_Objective_Binges2_EDEQ18x1014.2
Subjective_Binges2_EDEQ19x1015
N_Subjective_Binges2_EDEQ20x1016.2
Vomiting2_EDEQ21x1017
N_Vomiting2_EDEQ22x1018.2
Laxative_Use2_EDEQ23x1019
N_Laxative_Use2_EDEQ24x1020.2
Exercise2_EDEQ27x1023
N_Exercise2_EDEQ28x1024.2
Restraint2_EDEQ
Eating_Concern2_EDEQ


observed
synthetic
Weight_Concern2_EDEQ
Shape_Concern2_EDEQ
Global2_EDEQ
TotalScore3_MHCSF
PsychWellbeing3_MHCSF
EmoWellbeing3_MHCSF
SocialWellbeing3_MHCSF
Symptomatic_Distress3_OQ45
Interpersonal_Functioning3_OQ45
Social_Role3_OQ45
Anxiety_Somatic_Distress3_OQ45
Totat_Score_Psychopthology3_OQ45
Percent
Value
Objective_Binges3_EDEQ16x1012
N_Objective_Binges3_EDEQ18x1014.3
Subjective_Binges3_EDEQ19x1015
N_Subjective_Binges3_EDEQ20x1016.3
Vomiting3_EDEQ21x1017
N_Vomiting3_EDEQ22x1018.3
Laxative_Use3_EDEQ23x1019
N_Laxative_Use3_EDEQ24x1020.3
Exercise3_EDEQ27x1023
N_Exercise3_EDEQ28x1024.3
Restraint3_EDEQ
Eating_Concern3_EDEQ
No
Unknown
Yes
Weight_Concern3_EDEQ
Shape_Concern3_EDEQ
Global3_EDEQ
Y_ED_Psychopathology_1year
Not Improved
Reliably Improved
NA
Y_ED_Psychopathology_1year_Imputed
Gender.1
Female
Male
Edu_level.1
Unknown
Higher
Primary
Secondary
Edu_level_sp.1
Living_sit.1
Alone
Children, No Partner
Children, Partner
Mental health institution
No children, Partner
Other
With both parents (adolescent)
With one parent (adolescent)
physical_prob.1
Weight_anxiety.1
Prim_diag.1
AN
BED
BN
OSFED


observed
synthetic
Earlier_ED_diagnosis.1
Personality_disorder.1
Other_psych.1
BMI_Cat_1
Illness_Insight_2_SD
Social_Desirability_2_SD
Imputed
Imp_B
Percent
Value
No
Unknown
Yes
Healthy
Severe
Unhealthy
Adequate
High
Low
Very Low
Certainly
Maybe
Not
Probably
No_(Improved)
No_(NotImproved)
NotImputed
Imputed