J Rehabil Med 2020; 52: jrm00120

ORIGINAL REPORT

PERSONALIZED PREDICTIONS OF TREATMENT OUTCOME IN PATIENTS WITH POST-STROKE DEPRESSIVE SYMPTOMS

Johanne C. C. RAUWENHOFF, MSc[1,2], Suzanne C. VAN BRONSWIJK, MD[3], Frenk PEETERS, MD, PhD[3], Yvonne BOL, PhD[4], Alexander C. H. GEURTS, PhD[5] and Caroline M. VAN HEUGTEN, PhD[1,2,6]

From the [1]School for Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands [2]Limburg Brain Injury Centre, Maastricht, the Netherlands [3]Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands, [4]Department of Clinical and Medical Psychology, Zuyderland Medical Centre, Sittard-Geleen, [5]Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, the Netherlands and [6]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Objective: **Post-stroke depressive symptoms have a vast individual and societal impact. However, research into interventions for such symptoms show contradictory results; it is unclear what works for which patients. In addition, clinical prediction tools are lacking. This study aimed to develop a prognostic index model for treatment outcome in patients with post-stroke depressive symptoms.**

Methods: **Data from a randomized controlled trial ($n = 61$) evaluating 2 interventions for post-stroke depressive symptoms were used to predict post-treatment post-stroke depressive symptoms and participation. From 18 pre-treatment variables of patients and caregivers, predictors were selected using elastic net regression. Based on this selection, prognostic index scores (i.e. predictions) for both outcomes were computed for each individual patient.**

Results: **The depression model included all pre-treatment variables, explaining 44% of the variance. The strongest predictors were: lesion location, employment, participation, comorbidities, mobility, sex, and pre-treatment depression. Six predictors of post-treatment participation were identified, explaining 51% of the variance: mobility, pre-treatment participation, age, satisfaction with participation, caregiver strain, and psychological distress of the spouse. The cross-validated prognostic index scores correlated highly with the actual outcome scores (depression: correlation = 0.672; participation: correlation = 0.718).**

Conclusion: **Post-stroke depressive symptoms form a complex and multifactorial problem. Treatment outcome is influenced by the characteristics of the stroke, the patients, and their spouses. The results show that psychological distress is probably no obstacle to attempting to improve participation. The personalized predictions (prognostic index scores) of treatment outcome show promising results, which, after further replication and validation, could aid clinicians with treatment selection.**

Key words: post-stroke depressive symptoms; cognitive behavioural therapy; treatment outcome; prognostic index.

Accepted Sep 7, 2020; Epub ahead of print Sep 23, 2020

J Rehabil Med 2020; 52: jrm00120

LAY ABSTRACT

A stroke is a very dramatic event in a person's life. Patients may experience cognitive, emotional, and behavioural changes following a stroke, such as forgetfulness, mood changes, and lack of initiative. Therefore, returning to work and a busy social calendar might not be possible. One out of 3 patients who experience a stroke develops depressive symptoms. Unfortunately, these symptoms are difficult to treat. This study examined whether it is possible to predict the treatment outcome for individual stroke patients who have received psychological treatment for depressive symptoms. A statistical model was developed to predict the level of depressive symptoms and social participation for individual patients. With further development, this model could help psychologists decide which psychological treatment would be the best option for a particular patient. This might enable more patients to be provided with personalized treatment that could alleviate their depressive symptoms.

Correspondence address: Caroline van Heugten, Department of Psychiatry and Neuropsychology, Faculty of Health, Medicine and Life Sciences, PO Box 616, 6200 MD Maastricht, The Netherlands. E-mail: caroline.vanheugten@maastrichtuniversity.nl

Depressive symptoms are common following stroke and have a huge impact on functional and neurological outcome (1, 2), the rehabilitation process (3), and overall quality of life (4) of patients. Moreover, stroke places a high burden on society and the healthcare system (5). However, the treatment of post-stroke depressive symptoms (PSDS) is challenging. In clinical practice, treatment selection is often based on trial and error, with modest treatment efficacy. Furthermore, randomized controlled trials (RCTs) evaluating possible effective treatments, such as psychotherapy and pharmacological treatment, have contradictory results (6–8). A potential explanation is the heterogeneity and the multifactorial nature of the disorder (9). As a result, it is not clear what works for whom, and clinical prediction tools for treatment outcome are currently lacking.

There is increasing interest in the development of clinical prediction tools utilizing machine learning techniques (10). One such approach is the use of a large

doi: 10.2340/16501977-2744

pool of variables to develop predictive algorithms, which produce estimates of an individual's prognosis, otherwise known as prognostic index (PI) scores (11, 12). These predictive algorithms, or PI models, can predict the future symptom status of an individual patient; for instance, following a certain therapy, to determine the level of care that is needed in the future and therefore aid treatment selection. This technique is used within medical decision-making, for instance, to predict the effectiveness of different treatment options for breast cancer (13). Furthermore, the application is rapidly growing within research on depression (11, 12). The current study aimed to develop a PI model to predict the post-treatment outcome scores for patients with PSDS. Pre-treatment variables (such as clinical and injury-related variables) of a RCT, investigating 2 treatments for PSDS, were used to develop the PI model. This PI model predicted post-treatment outcome scores of depression and experienced participation restrictions for each participant. These predicted values are referred to as "PI scores".

METHODS

Study design

Data used in this study came from a multicentre RCT investigating the effectiveness of cognitive behavioural therapy (CBT) and computerized cognitive training for PSDS (6). CBT was adapted for people with a stroke; for instance, 3 sessions of occupational therapy or movement therapy were added to the treatment to enable the application of pleasurable activities. A detailed description of the intervention is published elsewhere (14). During computerized cognitive training, patients could select a combination of 4 cognitive domains, such as memory, for training. The programme difficulty level was adjusted accordingly (6). Both interventions consisted of 13–16 sessions in a 4-month time period. Both interventions were effective in significantly improving depressive symptoms and quality of life. However, no significant differences between the interventions were found for any of the outcome measures. Further description of the study procedures and efficacy results can be found elsewhere (6). The trial was approved by the medical ethics committee of Nijmegen (the Netherlands). Trial registration: Dutch trial register (NL2857).

Participants

Participants met the following inclusion criteria: having sustained any type of clinically confirmed stroke at least 3 months earlier, scoring >7 on the depression subscale of the Hospital Anxiety and Depression Scale (HADS-D), being 18 years or older, having only mild cognitive impairments (Mini-Mental State Examination score (MMSE) >27 out of 30), scoring positively on the communication-related items of the National Institutes of Health Stroke Scale, and understanding the Dutch language. Exclusion criteria were: pre-stroke major depression requiring psychiatric care, premorbid disability as reflected in a Barthel Index (BI) score <19 (out of 20), stay in an inpatient setting, severe comorbidity that might affect mood (e.g. malignancies), and post-stroke major depression requiring treatment with antidepressants.

Outcomes

Both depression and participation restriction scores, assessed immediately post-treatment, were used as primary outcome measures for the current analysis. Post-treatment depression scores were measured using the HADS-D. Scores on the depression subscale range from 0 to 21, with higher scores indicating more depressive symptoms. Good internal consistency for the HADS-D (Cronbach's $\alpha = 0.81$) was found in a stroke population (15). In the current sample, patients showed a significant decrease in the HADS-D pre-treatment compared with post-treatment (mean difference –4.6; 95% CI –5.7 to –3.6) (6).

Participation restrictions were measured using the restrictions subscale of the Utrecht Scale for Evaluation of Rehabilitation-Participation (USER-P). This scale measures the experienced restrictions regarding vocational, leisure, and social participation. The 10 items are rated on a scale from 0 to 3 or a "not applicable option". The sum score is converted to a 0–100 scale based on the items deemed applicable, with a higher score indicating fewer participation restrictions. It is a valid and reliable measure for former rehabilitation outpatients; the internal consistency of the restriction subscale was found to be good (Cronbach's $\alpha = 0.91$) (16). In the current sample, patients did not show a significant difference pre-treatment compared with post-treatment on the USER-P restriction scale (mean difference 2.9; 95% CI –0.4 to 6.2). Despite this overall non-significant difference, inspection of raw data showed vast differences in pre-post change scores between participants. Because of this high variability, it is interesting to predict individual post-treatment scores in order to identify who might benefit from the treatments.

Pre-treatment variables

A correlation matrix was computed for all variables measured pre-treatment in the original study. Variables that were highly correlated ($r \geq 0.60$) were discussed between co-authors (JR, SB, FP, and CvH) and based on previous research and consensus, the variable that was considered redundant was removed from the data-set (see Table SI[1] and Table SII[1]). As a result, the 18 variables, described below, were selected as potential predictors in order to develop the PI model.

Demographic variables included sex, age, and employment status. Variables related to the stroke were: time since stroke, type of stroke (ischaemic stroke, haemorrhagic stroke, subarachnoidal haemorrhage, or combination), location (left hemisphere, right hemisphere, brainstem, subarachnoidal haemorrhage, or combination), cognitive impairments measured with the MMSE (17), activities of daily living measured with the BI (18), and stroke impact measured with the mobility subscale of the Stroke Impact Scale (SIS) (19). Variables related to the psychological characteristics of the patients were: symptoms of anxiety measured with the anxiety subscale of the HADS (HADS-A), depressive symptoms measured with the HADS-D, coping style measured with the Utrecht Proactive Coping Competency List (UPCC) (20), frequency of behaviour regarding social participation measured with the frequency subscale of the USER-P, participation restrictions measured with the restriction subscale of the USER-P, satisfactions regarding social participation measured with the satisfaction subscale of the USER-P, and comorbidities measured with the Cumulative Illness Rating Scale (CIRS) (21). Finally, since stroke places a high burden on the spouse, variables related to the psychological characteristics of the spouses were

[1]https://doi.org/10.2340/16501977-2744

considered as possible predictors. These included symptoms of anxiety and depression measured with the HADS and caregiver strain measured with the Caregiver Strain Index (CSI) (22).

Statistical analyses

Data pre-processing. Missing data (outcome variables and pre-treatment variables) were imputed using a non-parametric random forest approach (R-package "MissForest", (23)). This imputation method has been proven accurate with lower imputation errors compared with other methods (23, 24). The following data was used to inform the imputation procedure: (*i*) non-missing outcome variables; (*ii*) non-missing pre-treatment variables; (*iii*) post-treatment measures of the pre-treatment variables (available for HADS-A, UPCC, the USER-P subscales, and the spouses' HADS and CSI); and (*iv*) treatment condition (CBT or computerized cognitive training). The imputation method was tested by producing missing data in the complete (non-missing) data-set and then comparing the imputed data values with the actual data values. This comparison was performed using the normalized root mean squared error (NRMSE) for continuous data and the proportion of falsely classified entries (PFC) for categorical data (23).

Variable transformation. All continuous variables were standardized and categorical variables were mean-centred to prevent potential errors in statistical inference (25). Variables with skewed distributions were transformed using a log transformation or a square root transformation based on visual inspection and normality tests. For variables that contained categories with limited observations, these categories were merged, since previous research recommends at least 10% of the sample in each category (26).

Prognostic index

Building the PI model. Two PI models were built to predict the study outcomes; 1 to predict post-treatment PSDS severity (HADS-D) and 1 to predict post-treatment participation restrictions (USER-P participation subscale). These PI models were constructed using elastic net regression (with R-package glmnet (27)). Elastic net regression is a combination of Lasso and Ridge regression. These are both linear regression models, which incorporate 2 penalty terms into the regression to prevent overfitting when many variables are included (28). Both penalty terms work by shrinking the regression coefficients of these variables. The Lasso (L1) penalty term can exclude variables by shrinking coefficients to 0; however, this has difficulties when handling highly correlated variables. The Ridge regression penalty (L2) is less affected by highly correlated variables, but shrinking coefficients to 0, and therefore selecting variables, is not allowed. Two tuning parameters are of importance in elastic net regression: (*i*) alpha that regulates the ratio between the L1 and L2 penalty terms (range between 0–1; 0 = Ridge/L2 penalty; 1 = LASSO/L1 penalty); (*ii*) lambda that regulates the overall degree of penalization. To determine the optimal alpha parameter, 25 iterations of 10-fold cross-validation were run with alpha values between 0 and 1 with 0.05 intervals. The optimal alpha was defined as the alpha that had the lowest cross-validation prediction error. With the resulting optimal alpha parameter, the optimal lambda parameter was determined using 1,000 iterations of 10-fold cross-validation. The optimal lambda was defined as the lambda with the lowest cross-validation prediction error.

Estimating the PI scores (i.e. predictions). Post-treatment depression severity and post-treatment participation restrictions were estimated for each individual using the final PI models. These individual estimates are also referred to as scores on the PI ("PI scores") since these predictions can be used to determine the level of future care that is needed (11, 12). To evaluate the predictive accuracy of the PI scores, the mean difference between the actual outcomes and the PI scores was calculated, and the association between these scores was examined using a correlation analysis. Finally, we determined whether outcomes varied between the 2 treatments for different levels on the PI: i.e. did individuals with certain prognoses benefit more from 1 of the 2 therapies? To test this, we examined the interactions between the PI scores and treatment condition in the following multiple regression analyses:

Evaluating the PI models. Predictors that were included in the PI models were categorized as important to less important depending on their parameters. The prediction accuracy of the PI models was evaluated using the adjusted R-square, i.e. the explained variance corrected for the number of included pre-treatment variables and the root mean squared error (RMSE), i.e. the root of the sum of the squared residuals, which are the observed values minus the model predictions.

Furthermore, the model performance was assessed using a re-sampling technique, namely 5-fold cross-validation (26). Therefore, the sample was (randomly) split into 5 equal groups. Then, for each of these groups, the PI scores of the individuals were predicted using the regression model based on information from the other 4 groups (the "training data-set" (29)). Model performance was then determined by evaluating the adjusted R-square, the RMSE, and the correlations between actual outcomes and PI scores.

RESULTS

Sample description, imputation of missing variables, and variable transformation

A total of 62 patients were included in the original study. For the current analyses, one participant was excluded due to drop-out before randomization. In total, 52 patients completed the post-treatment assessment. Table I shows the 18 pre-treatment variables grouped into 4 domains. No values were missing in the demographic variables. Of the injury-related variables, 18 values were missing (4.2%). Of the psychological variables of the patient, no values were missing. For the psychological variables of the spouse, 24 values were missing (19.4%) because not all spouses participated in the study (38 spouses participated). The data imputation was tested to be successful, with an estimated NRMSE of 0.27 and an estimated PFC of 0.26.

After standardization of the variables, 2 pre-treatment variables were found not to be normally distributed (time since stroke and CIRS score). These variables were both log-transformed.

Predictor selection

For the depression PI model, the alpha was estimated to be 0, indicating that it is based on a pure Ridge regression (including all pre-treatment variables). Therefore,

Table I. Sample description ($n = 62$)

Demographic variables		
Sex, *n*, women (%)	23 (37.1)	
Age, median (range)	61 (25–79)	
Active employment, *n* (%)	10 (16.1)	
Injury-related variables		
Time (months) since stroke, mean (SD)	41.9 (46.5)	
Type of stroke		
Ischaemic stroke, *n* (%)	45 (72.6)	
Haemorrhagic stroke, SAB, Combination, *n* (%)	11 (17.7)	
Unknown, *n* (%)	6 (9.7)	
Location of stroke		
Left, *n* (%)	24 (38.7)	
Right, *n* (%)	19 (30.6)	
Brainstem, cerebellum, combination, *n* (%)	11 (17.7)	
Unknown, *n* (%)	8 (12.9)	
MMSE score, mean (SD)	29.1 (1.4)	
BI, mean (SD)	19.5 (1.32)	
SIS score, mean (SD)	65.2 (20.6)	
Psychological variables	Pre-treatment	Post-treatment
HADS-D, mean (SD)	12.4 (3.3)	7.8 (3.6)
HADS-A, mean (SD)	9.9 (4.2)	7.4 (3.9)
UPCC, mean (SD)	2.5 (0.5)	2.5 (0.6)
USER-P frequency, mean (SD)	28.7 (9.9)	28.9 (9.2)
USER-P restriction, mean (SD)	72.2 (12.1)	75.3 (13.3)
USER-P satisfaction, mean (SD)	52.6 (16.9)	62.7 (16.8)
CIRS, mean (SD)	5 (3.9)	–
Psychological variables spouse		
HADS total score, mean (SD)	11.4 (6.6)	
CSI, mean (SD)	6.4 (3.1)	

SD: standard deviation; MMSE: Mini-Mental State Examination; BI: Barthel Index; SIS: Stroke Impact Scale; HADS: Hospital Anxiety and Depression Scale; UPCC: Coping Competency List; USER-P: Utrecht Scale for Evaluation of Rehabilitation-Participation; CIRS: Cumulative Illness Rating Scale; CSI: Caregiver Strain Index.

all 18 variables were included as predictors in the depression model (see Table II). The 7 variables with the highest parameters (higher than 0.3) were a lesion in the left hemisphere, being in employment, more social participation, fewer comorbidities, better mobility, male sex, and less severe depressive symptomatology pre-treatment. These variables were all associated with a lower HADS-D score post-treatment. For the participation restrictions PI model, the alpha was estimated to be 0.95, indicating that it is a combination of Lasso and Ridge regression. In this model, a total of 6 predictors were selected (see Table III). These were better mobility, fewer pre-treatment participation restrictions, older age, less caregiver strain, less anxiety and less depression of the spouse, and more satisfaction regarding participation pre-treatment. These variables were all associated with fewer participation restrictions post-treatment.

Table II. Predictors selected with elastic net regression for the depression model

Predictor	Coefficient
Location of stroke (right vs left)	0.530
Active employment	−0.445
Utrecht Scale for Evaluation of Rehabilitation-Participation frequency	−0.442
Cumulative Illness Rating Scale	0.405
Stroke Impact Scale score	−0.382
Sex	0.338
The depression subscale of the Hospital Anxiety and Depression Scale	0.330
Utrecht Scale for Evaluation of Rehabilitation-Participation satisfaction	−0.296
Mini-Mental State Examination score	0.203
Coping Competency List	−0.198
Caregiver Strain Index	0.171
The anxiety subscale of the Hospital Anxiety and Depression Scale	−0.159
Utrecht Scale for Evaluation of Rehabilitation-Participation restriction	−0.153
Time since stroke	−0.121
Location of stroke (left vs brainstem, cerebellum, combination)	0.093
Barthel Index	−0.087
Type of stroke	−0.642
Hospital Anxiety and Depression Scale spouse	−0.0004
Age	0.00002

Table III. Predictors selected with elastic net regression for the participation restrictions model

Predictor	Coefficient
Stroke Impact Scale score	4.105
USER-P restriction	1.966
Age	1.946
Caregiver Strain Index	1.491
Hospital Anxiety and Depression Scale spouse	−0.430
USER-P satisfaction	0.049

USER-P: Utrecht Scale for Evaluation of Rehabilitation-Participation.

Estimating and evaluating the prognostic index models

Table IV shows the model performance of the PI models of Depression and Participation Restrictions based on the complete data-set and based on a hold-out data-set. The R^2 of the depression model was 0.442, meaning that the model explains 44.2% of the variance. The mean difference between the actual post-treatment HADS-D scores and the PI scores was 2.162 (SD 1.562). The RMSE was 2.66, indicating that the mean of the model residuals (actual scores minus the model predictions) was 2.66 points on the HADS-D scale. Furthermore, the correlation between the actual and predicted values was significant and strong (correlation = 0.672, $p < 0.001$, see Fig. 1).

The R^2 of the depression PI model that was developed and fitted on a hold-out data-set (5-fold cross-validation) was 0.134 and the RMSE was 3.17. When examining the association between actual post-treatment HADS-D scores and the PI scores based on this cross-validation model, a moderate and significant correlation was found (correlation = 0.366, $p = 0.004$, see Fig. 1).

Table IV. Model performance of prognostic index (PI) models of depression and participation restrictions

	Depression PI model	Participation restrictions PI model
Model performance based on models fitted on the complete data-set		
R^2	0.451	0.516
RMSE	2.660	8.927
Model performance based on models fitted on a hold-out data-set (5-fold cross-validation)		
R^2	0.134	0.316
RMSE	3.169	10.321

RMSE; root mean squared error.

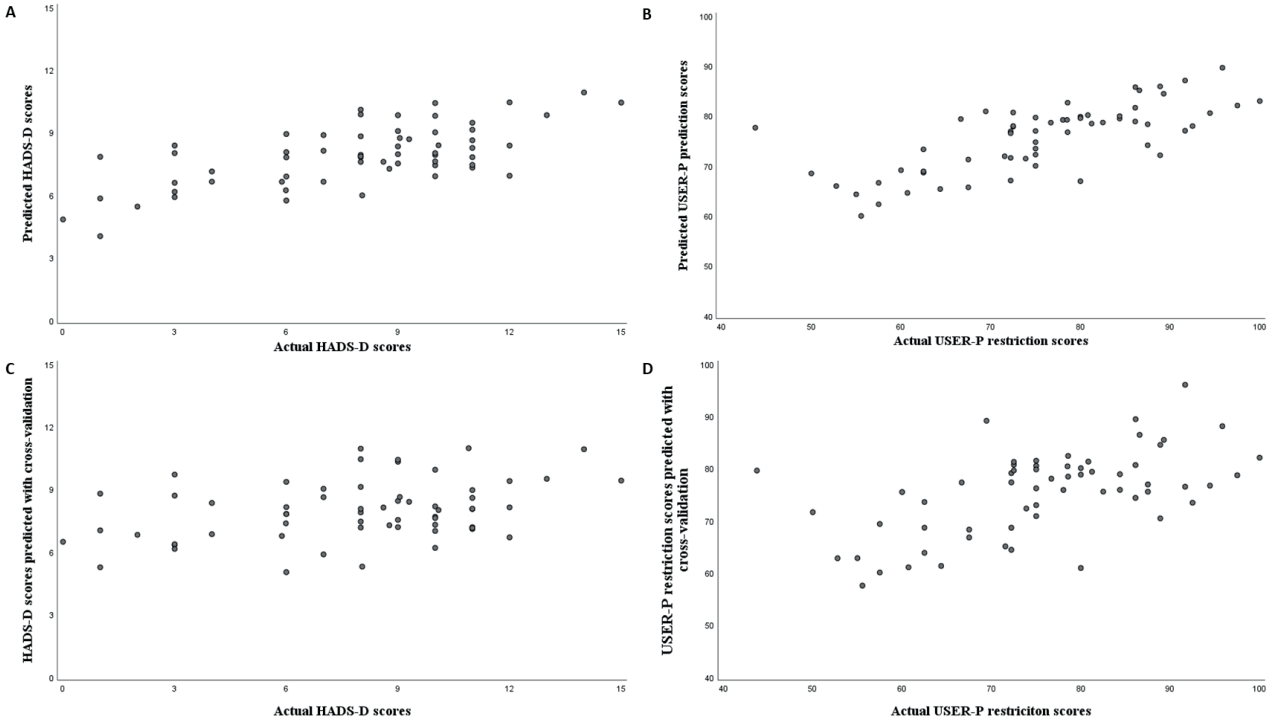

Fig. 1. Scatterplot of (A) actual and predicted Hospital Anxiety and Depression Scale (HADS)-D scores; (B) actual and predicted Utrecht Scale for Evaluation of Rehabilitation-Participation (USER-P) restriction scores; (C) actual HADS-D scores and HADS-D scores predicted with cross-validation; and (D) actual USER-P restriction scores and USER-P restriction scores predicted with cross-validation.

The R^2 of the participation PI model was 0.507, meaning that the model explains 50.7% of the variance. The mean difference between the actual post-treatment USER-P restriction scores and the PI scores was 6.563 (SD 6.102). The RMSE was 8.93, indicating that the mean of the model residuals (observed values minus the model predictions) was 8.93 points on the USER-P restriction scale. The correlation between the observed and predicted values was again strong and significant (correlation = 0.718, $p < 0.001$, see Fig. 1).

For the participation PI model that was developed and fitted on a hold-out data-set (5-fold cross-validation), the R^2 was 0.335 and the RMSE was 10.15. When examining the association between actual post-treatment USER-P restriction scores and the PI scores based on this cross-validation model, a moderate and significant correlation was found (correlation = 0.562, $p < 0.001$, see Fig. 1).

Multiple regression analyses were carried out to test whether depression and restriction outcomes varied between the 2 treatments for different levels on the PI. The models indicated no distinct treatment effects on different PI levels.

DISCUSSION

The goal of this study was to develop a PI model for treatment outcome in patients with PSDS. The post-treatment outcome scores for depression and experienced participation restrictions were predicted based on pre-treatment variables of both the patient and the spouse. The depression model explained 44.2% of the variance and the actual depression scores correlated highly with the predicted scores. The participation model explained 50.7% of the variance and predicted scores again correlated highly with actual scores.

The range of the HADS-D is 0–21. The RMSE, the squared root of the mean of squared differences between predictions and actual outcomes, was 2.66 for the HADS-D, which is probably not a clinically significant difference (the minimum difference on the HADS-D to be a clinically significant difference ranges from 0.5 to 6 dependent on the population (30–32)). The range of the USER-P participation is 0–100. The RMSE was 8.927 for the USER-P, which is probably not a clinically significant difference either (although this has not been studied yet, in this situation Ringash et al. (33) advise that 10% of the instrument range can be considered the minimum important difference, which would be at least 10 points in this case).

Furthermore, the performance of both models was promising when assessed using a re-sampling technique. No interaction effect was found between predicted scores and the received intervention, meaning that there was not a group of patients with a certain (predicted) prognosis who benefited more from 1 of the 2 therapies.

A more favourable outcome of PSDS was predicted by a left hemispheric lesion, male sex, better mobility, less depressive symptoms pre-treatment, more social participation, fewer comorbidities, and being in employment. These variables showed to be most predictive of PSDS post-treatment of the 18 variables included in the depression model.

A left hemispheric lesion was identified as an important predictor for less depressive symptoms post-treatment, which is somewhat surprising. To our knowledge, there is no earlier research on the association between lesion location and treatment outcome and studies identifying lesion laterality as a possible predictor related to post-stroke depression show inconsistent results (34, 35). Nevertheless, the meta-analyses of Wei et al. (36) did find an association between right-hemispheric lesion and risk of depression. However, this association was only apparent 1–6 months' post-stroke. The results are therefore probably not applicable to our sample, since there were only 3 patients who were less than 6 months post-stroke. Currently, the focus is shifting to damaged neuronal networks instead of brain regions as an underlying mechanism for PSDS (35), which, in future research, should also be considered in relation to treatment outcome for PSDS. The finding that pre-treatment depression severity is predictive of PSDS outcome is in line with previous studies, which show that pre-treatment depression levels play an important role in treatment outcome in patients irrespective of the presence of acquired brain injury (11, 37, 38). Likewise, the finding that being in employment is a predictive factor of a more favourable outcome is in line with earlier research in depressed patients without brain injury (11). The other predictors of post-treatment depression severity in the current study (including male sex, better mobility, more social participation, and fewer comorbidities) are all known protective factors against the development of PSDS (39–41). It is feasible that most of these resilience factors can also provide opportunities to better use and apply the competencies obtained during therapy.

In this study, the potential predictors for fewer restrictions regarding participation scores post-treatment in patients with PSDS were: better mobility, fewer pre-treatment participation restrictions, older age, less caregiver strain and psychological distress of the spouse, and more satisfaction regarding participation. The finding that mobility is the strongest predictive factor is not surprising. Social participation is associated with functional disability in the recovery process following a stroke (42). However, whereas one would hope that a treatment for stroke patients would decrease participation restrictions despite physical disabilities, the interventions in this study might not have achieved this. The finding that older age is predictive of a more favourable outcome is not in line with earlier research. Previous studies found that older age is often related to more experienced participation restrictions (43). However, in previous studies patients were relatively older and might experience, next to stroke-related restrictions, more restrictions due to older age. It seems probable that patients who are younger experience more participation restrictions because society expects a higher level of participation (i.e. going back to work, taking care of children). Furthermore, participation satisfaction and restrictions, but not participation frequency, were predictive factors. Earlier research found that change in frequency of vocational activities, but not social and leisure activities, are predictive of participation restrictions at 6 months post-stroke (44). Merely increasing participation frequency will, therefore, probably not lead to an improvement of participation restrictions and satisfaction.

Both the level of caregiver burden and psychological distress of the spouse were predictive of participation restrictions following the interventions. It seems plausible that spouses who are psychologically more resilient and experience less psychological distress and less caregiver strain can support and encourage their spouses better during treatment and help to change therapeutic intentions into practical therapeutic actions. Furthermore, earlier research found that spouses experience more participation restrictions themselves, when they have more depressive symptoms, are in employment, have a younger age, and support a stroke patient with more disabilities and lower participation levels (45). It seems that experienced participation restrictions reflect a close interplay between spouse and patient.

When comparing the 2 models in this study, it becomes apparent that improving PSDS and decreasing experienced participation restrictions might involve different processes. The treatment of depression seems complicated, with many factors influencing the outcome, while fewer factors influencing the outcome when decreasing participation restrictions. The results highlight the complexity and multifactorial nature of treating depressive symptoms following a stroke. The outcome of treatment for PSDS is influenced by characteristics of the patient, stroke, and well-being of the spouse, which should all be considered when treating a patient with PSDS. The process of decreasing experienced participation restrictions is, largely, influenced by the physical characteristics of the patient and psychological characteristics of the spouse. Interestingly, the levels of anxiety and depression of the patient him/herself were not predictive of restriction, implying that the experienced participation restrictions can be decreased regardless of experienced psychological distress. This is in line with third-generation cognitive behavioural therapies, including acceptance and commit-

ment therapy (ACT). The goal of ACT is not to decrease symptomatology, but to increase psychological flexibility and behaviours based on values despite the presence of, for instance, depressive thoughts and feelings (46).

Study strengths

This study has several strengths. First, the state-of-the-art variable selection approach used (i.e. elastic net regression) combines multiple predictors instead of examining individual predictors separately. To our knowledge, this is the first study to incorporate multiple predictors to develop personalized predictions for the outcome of treatment for PSDS. Secondly, elastic net regression is able to minimize the number of predictors and to categorize predictors from important to less important. Thirdly, we evaluated the performance of both PI models using a re-sampling technique (cross-validation). Fourthly, this study included a broad range of possible predictors. For instance, characteristics of the spouse were considered as predictors of outcome that have not been included in earlier research. Fifthly, this study predicted both depression and experienced participation restrictions, and therefore was able to show the different nature of these 2 outcomes.

Study limitations

This study has some limitations. The patients were relatively young compared with the mean stroke population (median age 61 years). Furthermore, patients with severe cognitive impairments, communication problems, major depression, or who were in need of inpatient care were excluded. This resulted in a sample of patients with less severe complaints. Both considerations should be taken into account when interpreting the results. Furthermore, the small sample size can be seen as a limitation, which is a common problem in studies using data from RCTs to develop prediction models (47). Due to the relatively small sample size, no separate training and testing data-sets were used. This could have led to overfitting of the models to the current data-set, which decreases the external validity of the PI models (48). However, machine learning methods have many advantages compared with more traditional models (such as linear models), because they have an increased model prediction accuracy by reducing overfitting (49). The external validity of both PI models was assessed with a re-sampling technique on a hold-out data-set, which showed promising results. In addition, elastic net regression includes 2 penalty terms to the regression function to prevent overfitting (28). Furthermore, although we tested whether depression and participation outcomes varied between the 2 treatments for different levels on the PI, we were not able to investigate predictors of differential treatment effects (i.e. moderators) specifically, due to the small sample size. The results apply to both interventions, which are very different in nature (i.e. behavioural therapy and cognitive training), although equally effective in the original RCT from which the data were drawn. It is clear that further research is needed to replicate and externally validate the current results.

Conclusion

This proof of concept study shows that machine learning techniques, such as elastic net regression, can be used to compute personalized predictions of outcome following treatment for PSDS. However, the models developed in this study are not yet ready for implementation in clinical practice. The results demonstrate the complex and multifactorial nature of PSDS, and this should be considered in planning treatment approaches. Furthermore, this study shows that psychological factors are probably no obstacle to improving restrictions regarding social participation. In order to realize the use of these models in clinical practice, further research is needed to replicate and externally validate the current results. PI models have a great potential to aid clinicians and their patients with treatment selection and therefore increase the effectiveness of treatments for PSDS.

ACKNOWLEDGEMENTS

The authors would like to thank Joyce A. Kootker, PhD, for conducting the RCT from which data were used for this study, and to acknowledge the participants and the participating sites: Groot Klimmendaal, St Maartenskliniek, Adelante, Roessingh, VieCuri, Tolbrug, and ViaReva.

The study is funded by ZonMW. The funder had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript.

REFERENCES

1. Pohjasvaara T, Vataja R, Leppävuori A, Kaste M, Erkinjuntti T. Depression is an independent predictor of poor long-term functional outcome post-stroke. Eur J Neurol 2001; 8: 315–319.
2. Vuletic V, Sapina L, Lozert M, Lezaic Z, Morovic S. Anxiety and depressive symptoms in acute ischemic stroke. Acta Clin Croat 2012; 51: 243–246.
3. Gillen R, Tennen H, McKee TE, Gernert-Dott P, Affleck G. Depressive symptoms and history of depression predict rehabilitation efficiency in stroke patients. Arch Phys Med Rehabil 2001; 82: 1645–1649.
4. Kim ES, Kim JW, Kang HJ, Bae KY, Kim SW, Kim JT, et al. Longitudinal impact of depression on quality of life in stroke patients. Psychiatry Investig 2018; 15: 141–146.
5. Husaini B, Levine R, Sharp L, Cain V, Novotny M, Hull P, et al. Depression increases stroke hospitalization cost: an analysis of 17,010 stroke patients in 2008 by race and gender. Stroke Res Treat 2013; 2013: 846732.
6. Kootker JA, Rasquin SM, Lem FC, van Heugten CM, Fasotti L, Geurts AC. Augmented cognitive behavioral therapy for poststroke depressive symptoms: a randomized controlled

trial. Arch Phys Med Rehabil 2017; 98: 687–694.
7. Lincoln NB, Flannaghan T. Cognitive behavioral psychotherapy for depression following stroke – a randomized controlled trial. Stroke 2003; 34: 111–115.
8. Villa RF, Ferrari F, Moretti A. Post-stroke depression: mechanisms and pharmacological treatment. Pharmacol Ther 2018; 184: 131–144.
9. Lorenzo-Luaces L. Heterogeneity in the prognosis of major depression: from the common cold to a highly debilitating and recurrent illness. Epidemiol Psychiatr Sci 2015; 24: 466–472.
10. Cohen ZD, DeRubeis RJ. Treatment selection in depression. Ann Rev Clin Psychol 2018; 14: 209–236.
11. Lorenzo-Luaces L, DeRubeis RJ, van Straten A, Tiemens B. A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. J Affect Disorders 2017; 213: 78–85.
12. van Bronswijk SC, Lemmens L, Keefe JR, Huibers MJH, DeRubeis RJ, Peeters F. A prognostic index for long-term outcome after successful acute phase cognitive therapy and interpersonal psychotherapy for major depressive disorder. Depress Anxiety 2019; 36: 252–261.
13. Chen AM, Meric-Bernstam F, Hunt KK, Thames HD, Outlaw ED, Strom EA, et al. Breast conservation after neoadjuvant chemotherapy: a prognostic index for clinical decision-making. Cancer 2005; 103: 689–695.
14. Kootker JA, Rasquin SM, Smits P, Geurts AC, van Heugten CM, Fasotti L. An augmented cognitive behavioural therapy for treating post-stroke depression: description of a treatment protocol. Clin Rehabil 2015; 29: 833–843.
15. Ayis SA, Ayerbe L, Ashworth M, Wolfe CDA. Evaluation of the Hospital Anxiety and Depression Scale (HADS) in screening stroke patients for symptoms: item response theory (IRT) analysis. J Affect Disorders 2018; 228: 33–40.
16. Post MWM, van der Zee CH, Hennink J, Schafrat CG, Visser-Meily JMA, van Berlekom SB. Validity of the Utrecht Scale for Evaluation of Rehabilitation-Participation. Disabil Rehabil 2012; 34: 478–485.
17. Burns A. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. M. Folstein, S. Folstein and P. McHugh, J Psychiatr Res (1975) 12, 189–198. Introduction. Int J Geriatr Psych 1998; 13: 285–285.
18. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988; 10: 61–63.
19. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. Stroke 1999; 30: 2131–2140.
20. Bode C, Thoolen B, de Ridder D. Measuring proactive coping. Psychometric characteristics of the Utrecht Proactive Coping Competence scale (UPCC). Psychol Gezondh 2008; 36: 81–91.
21. Giaquinto S. Comorbidity in post-stroke rehabilitation. Eur J Neurol 2003; 10: 235–238.
22. Robinson BC. Validation of a Caregiver Strain Index. J Gerontol 1983; 38: 344–348.
23. Stekhoven DJ, Buhlmann P. MissForest-non-parametric missing value imputation for mixed-type data. Bioinformatics 2012; 28: 112–118.
24. Waljee AK, Mukherjee A, Singal AG, Zhang YW, Warren J, Balis U, et al. Comparison of imputation methods for missing laboratory data in medicine. BMJ Open 2013; 3: e002847.
25. Kraemer HC, Blasey CM. Centring in regression analyses: a strategy to prevent errors in statistical inference. Int J Methods Psychiatr Res 2004; 13: 141–151.
26. Kuhn M, Johnson K. Applied predictive modeling. New York: Springer; 2013.
27. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33: 1–22.
28. Zou H, Hastie T. Regularization and variable selection via the elastic net (vol B 67, pg 301, 2005). J R Stat Soc B 2005; 67: 768–768.
29. Picard RR, Cook RD. Cross-validation of regression models. J Amer Statist Assoc 1984; 79: 575–583.
30. Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the minimal clinically important difference for the Hospital Anxiety and Depression Scale in patients with cardiovascular disease. J Cardiopulm Rehabil Prev 2019; 39: E6–E11.
31. Curtis M, Kon S, Canavan J, Jones S, Nolan C, Clark A, et al. The minimum important difference of the hospital anxiety and depression scale in COPD. Eur Respir J 2014; 44.
32. Corsaletti BF, Proença M-DGL, Bisca GKW, Leite JC, Bellinetti LM, Pitta F. Minimal important difference for anxiety and depression surveys after intervention to increase daily physical activity in smokers. Fisioterapia e Pesquisa 2014; 21: 359–364.
33. Ringash J, O'Sullivan B, Bezjak A, Redelmeier DA. Interpreting clinically significant changes in patient-reported outcomes. Cancer 2007; 110: 196–202.
34. Mitchell AJ, Sheth B, Gill J, Yadegarfar M, Stubbs B, Yadegarfar M, et al. Prevalence and predictors of post-stroke mood disorders: a meta-analysis and meta-regression of depression, anxiety and adjustment disorder. Gen Hosp Psychiatry 2017; 47: 48–60.
35. Nickel A, Thomalla G. Post-stroke depression: impact of lesion location and methodological limitations—a topical review. Front Neurol 2017; 8: 498.
36. Wei N, Yong W, Li X, Zhou Y, Deng M, Zhu H, et al. Post-stroke depression and lesion location: a systematic review. J Neurol 2015; 262: 81–90.
37. Anson K, Ponsford J. Who benefits? Outcome following a coping skills group intervention for traumatically brain injured individuals. Brain Inj 2006; 20: 1–13.
38. Cohen ZD, Kim TT, Van HL, Dekker JJ, Driessen E. A demonstration of a multi-method variable selection approach for treatment selection: Recommending cognitive–behavioral versus psychodynamic therapy for mild to moderate adult depression. Psychother Res 2020; 30: 137–150.
39. Cnossen MC, Scholten AC, Lingsma HF, Synnot A, Haagsma J, Steyerberg PEW, et al. Predictors of major depression and posttraumatic stress disorder following traumatic brain injury: a systematic review and meta-analysis. J Neuropsychiatry Clin Neurosci 2017; 29: 206–224.
40. Guiraud V, Gallarda T, Calvet D, Turc G, Oppenheim C, Rouillon F, et al. Depression predictors within six months of ischemic stroke: the DEPRESS Study. Int J Stroke 2016; 11: 519–525.
41. Ayerbe L, Ayis S, Rudd AG, Heuschmann PU, Wolfe CD. Natural history, predictors, and associations of depression 5 years after stroke: the South London Stroke Register. Stroke 2011; 42: 1907–1911.
42. D'alisa S, Baudo S, Mauro A, Miscio G. How does stroke restrict participation in long-term post-stroke survivors? Acta Neurologica Scandinavica 2005; 112: 157–162.
43. de Graaf JA, van Mierlo ML, Post MWM, Achterberg WP, Kappelle LJ, Visser-Meily JMA. Long-term restrictions in participation in stroke survivors under and over 70 years of age. Disabil Rehabil 2018; 40: 637–645.
44. Blömer A-MV, Van Mierlo ML, Visser-Meily JM, Van Heugten CM, Post MW. Does the frequency of participation change after stroke and is this change associated with the subjective experience of participation? Arch Phys Med Rehabil 2015; 96: 456–463.
45. Grigorovich A, Forde S, Levinson D, Bastawrous M, Cheung AM, Cameron JI. Restricted participation in stroke caregivers: who is at risk? Arch Phys Med Rehabil 2015; 96: 1284–1290.
46. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006; 44: 1–25.
47. DeRubeis RJ. The history, current status, and possible future of precision mental health. Behav Res Ther 2019; 123: 103506.
48. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: prediction, inference and data mining. New York: Springer-Verlag; 2009.
49. Steyerberg EW, Harrell Jr FE, Borsboom GJ, Eijkemans M, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001; 54: 774–781.