ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2024;105: 921–9

ORIGINAL RESEARCH

A Machine Learning Approach to Predict Post-stroke Fatigue. The Nor-COAST study

Geske Luzum, MSc,[a,*] Gyrd Thrane, PhD,[b,*] Stina Aam, PhD,[a,c] Rannveig Sakshaug Eldholm, PhD,[a,c] Ramune Grambaite, PhD,[d] Ragnhild Munthe-Kaas, PhD,[e,f] Pernille Thingstad, PhD,[a,g] Ingvild Saltvedt, PhD,[a,c] Torunn Askim, PhD[a]

From the [a]Department of Neuromedicine and Movement Science, NTNU-Norwegian University of Science and Technology, Trondheim, Norway; [b]Department of Health and Care Science, The Arctic University of Norway, Tromsø, Norway; [c]Department of Geriatric Medicine, Clinic of Medicine, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway; [d]Department of Psychology, NTNU-Norwegian University of Science and Technology, Trondheim, Norway; [e]Department of Medicine, Kongsberg Hospital, Vestre Viken Hospital Trust, Drammen, Norway; [f]Department of Medicine, Bærum Hospital, Vestre Viken Hospital Trust, Drammen, Norway; and [g]Department of Health and Welfare, Trondheim Municipality, Trondheim, Norway.

Abstract

Objective: This study aimed to predict fatigue 18 months post-stroke by utilizing comprehensive data from the acute and sub-acute phases after stroke in a machine-learning set-up.

Design: A prospective multicenter cohort-study with 18-month follow-up.

Setting: Outpatient clinics at 3 university hospitals and 2 local hospitals.

Participants: 474 participants with the diagnosis of acute stroke (mean ± SD age; 70.5 (11.3), 59% male; N=474).

Interventions: Not applicable.

Main Outcome Measures: The primary outcome, fatigue at 18 months, was assessed using the Fatigue Severity Scale (FSS-7). FSS-7≥5 was defined as fatigue. In total, 45 prediction variables were collected, at initial hospital-stay and 3-month post-stroke.

Results: The best performing model, random forest, predicted 69% of all subjects with fatigue correctly with a sensitivity of 0.69 (95% CI: 0.50, 0.86), a specificity of 0.74 (95% CI: 0.66, 0.83), and an Area under the Receiver Operator Characteristic curve of 0.79 (95% CI: 0.69, 0.87) in new unseen data. The proportion of subjects predicted to suffer from fatigue, who truly suffered from fatigue at 18-months was estimated to 0.41 (95% CI: 0.26, 0.57). The proportion of subjects predicted to be free from fatigue who truly did not have fatigue at 18-months was estimated to 0.90 (95% CI: 0.83, 0.96).

Conclusions: Our findings indicate that the model has satisfactory ability to predict fatigue in the chronic phase post-stroke and may be applicable in clinical settings.

Archives of Physical Medicine and Rehabilitation 2024;105:921–9

Fatigue is a multidimensional, subjective experience,[1] characterized by a sensation of early exhaustion, and aversion to effort.[2] The reported prevalence of post-stroke fatigue (PSF) ranges from 25% to 85%.[3] As much as 40% of people suffering from PSF consider it to be one of the worst consequences of stroke.[4]

The current body of literature suggests that the brain lesion, stroke-related inflammatory and neuroendocrine changes as well as impairment in attention and executive function can serve as a trigger for developing early fatigue.[5] The risk for persisting fatigue increases with the prevalence of pre-stroke fatigue and early PSF.[6-8] In line with this observation, most longitudinal studies report stable fatigue levels after the subacute phase.[8,9]

This study was supported by the Norwegian Health Association, and additional funding was provided by the Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Science, NTNU-Norwegian University of Science and Technology.

Clinical Trial Registration No.: NCT02650531

Disclosures: none.

* Luzum and Thrane contributed equally to this work.

0003-9993/$36 - see front matter © 2024 by the American Congress of Rehabilitation Medicine. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
https://doi.org/10.1016/j.apmr.2023.12.005

However, besides persistent forms of fatigue trajectories, one may recover from fatigue or develop late-onset fatigue.[5] Symptoms of fatigue seem to be linked to affective, physical, psychological, and behavioral factors, that also interact with each other, and may also coexist in combination with related symptoms.[5,10,11] In fact, fatigue may be more common among stroke survivors with depression, anxiety, pain, sleeping disturbances and attentional-executive impairment,[8,12-14] while conflicting evidence has been found for the relationship between age and fatigue.[8,13,15]

Because fatigue still seems to be an overlooked part in stroke care, the Canadian Stroke Best Practice Recommendations stress the importance of screening and evaluation procedures early after a stroke.[16] Identification of individuals at high risk of persisting fatigue may help to provide more targeted interventions, at the right time. Previous attempts to develop reliable prediction tools to identify individuals at risk of suffering from PSF have been found to have limited predictive value.[7,17] In addition, most clinical risk models have been based on simple regression models, including only a small number of predictors.[18] With regard to the increasing availability of large, comprehensive datasets and the risk of complex variable interactions, machine learning techniques may represent a viable solution to current prediction challenges.[19] Therefore, this study aimed to apply a machine-learning approach in predicting fatigue 18 months post-stroke by utilizing comprehensive data from the acute and sub-acute phases.

Methods

Study design and participants

This study utilized data from a national, multicenter, prospective, cohort study with consecutive inclusion during hospital admission with the diagnosis of acute stroke. Eligibility criteria of the main study (N=815) were: 1) Admittance to 1 of the 5 participation hospitals within 7 days after onset of symptoms; 2) acute stroke diagnosed according to the World Health Organization criteria or with findings of acute infarction or intra-cerebral hemorrhage on magnetic resonance imaging; 3) able to speak a Scandinavian language; 4) age≥18 years; and 5) living in the catchment area of the recruiting hospitals. An expected survival of less than 3 months led to exclusion from the study. The study protocol was approved by the regional committees for medical and health research. Details of the main study's protocol are described elsewhere.[20] For the present analyses we included subjects that had a valid measurement of fatigue 18 months post-stroke.

Main outcome variable

The primary response variable, fatigue 18 months post-stroke, was assessed using the Fatigue Severity Scale (FSS-7).[21,22] On a Likert Scale ranging from 1 (Strongly disagree) to 7 (Strongly agree), participants were asked to rate 7 fatigue-related statements in the last week. For the analysis, the mean score of FSS-7 was dichotomized in accordance with the recommendation to define high-level fatigue as a mean score of 5 or above.[23]

List of abbreviations:

FSS-7	**Fatigue Severity Scale**
HADS	**Hospital Anxiety and Depression Scale**
PA	**physical activity**
PPV	**Positive predictive value**
PSF	**Post-stroke fatigue**
ROC-AUC	**area under the receiver operator characteristic curve**

Prediction variables

Table 1 provides an overview over all prediction variables, obtained at baseline and at 3 months post-stroke, for participants with and without fatigue 18 months post-stroke.

Baseline

Measures of pre-stroke physical activity (PA) and pre-stroke fatigue were obtained using the standardized questions from the North-Trøndelag Health Study.[24,25] Pre-stroke global functioning was classified with the modified Rankin Scale.[26] The National Institutes of Health Stroke Scale was used to measure stroke symptoms and severity at the time of admission (day 1), and 7 days later (at discharge if discharged earlier than the seventh day).[27] The Oxfordshire Stroke Classification Project[28] and the Trial of Org 10172 in Acute Stroke Treatment classification were used to classify the stroke.[29]

Three months testing

The neuropsychological testing was based on a recommended test battery after stroke,[30,31] including the Trail Making Tests A and B,[32] the Verbal Fluency Test Letter,[33] and the Global Deterioration Scale.[34] The Global Detoriation Scale was scored by a trained research assistant, based on the combination of available information from tests and interviews. The Montreal Cognitive Assessment was also included in the assessment.[35]

Occurrence and severity of neuropsychiatric symptoms, like delusions, hallucinations and aggression, was assessed based on self-report or proxy information using the 12-item Neuropsychiatric Inventory questionnaire.[36] While, symptoms of anxiety and depression were screened using the Hospital Anxiety and Depression Scale (HADS).[37] If the participants were not able to respond, proxy information on depression was collected using the Cornell scale.[38] Post-stroke global functioning was measured with the modified Rankin Scale. Physical capacity was classified with the Short Physical Performance Battery (SPPB).[39] Grip-strength was measured using a Jamar handhold dynamometer. The Nine Hole Peg Test was used to assess dexterity.[40] Gait speed at preferred and fast speed, and during dual task-performance (counting backwards) were assessed on a 10-meter distance, with flying start. We also assessed fatigue at 3 months as a predictor variable. The FSS-7 served as continuous predictor variable at 3 months using the mean score and as dichotomized outcome variable at 18 months. Further, PA was measured over a course of 4-7 days, using a 3-axial accelerometer, activPAL[a] monitors.[41] We used a tailored MATLAB[a] script (available upon request) to extract mean weekly moderate PA (walking at intensities ≥3 metabolic equivalent of task, METs), mean weekly light PA (standing or walking at intensities <3 METs), mean weekly time upright, and mean weekly time standing, from the data. In addition, weekly time spent in moderate PA was split into short (<10 minutes) and long-bout activities (≥10 minutes). We only considered daytime PA between 8.00 AM and 11.30 PM.

Table 1 Participant characteristics

Characteristics	Missing	Overall N=474	Fatigue (FSS-7>=5), 18 Months: No N=378 (79.7)	Fatigue (FSS-7>=5), 18 Months: Yes N=96 (20.3)
Baseline measurement				
Age at baseline, y, mean ± SD	0	70.49 (11.3)	70.67 (10.6)	69.79 (13.5)
Sex, n (%)	0			
Male		279 (58.9)	239 (63.2)	40 (41.7)
Female		195 (41.1)	139 (36.8)	56 (8.3)
Living conditions, n (%)	0			
Home without nursing care		455 (96.0)	365 (96.6)	90 (93.8)
Home with care/ nursing home		19 (4.0)	13 (3.4)	6 (6.2)
Marital status, n (%)	2			
Married or cohabitant		323 (68.1)	264 (69.8)	59 (61.5)
Single		73 (15.4)	54 (14.3)	19 (19.8)
Widow		76 (16.0)	58 (15.3)	18 (18.8)
Living situation, n (%)	0			
Living alone		139 (29.3)	109 (28.8)	30 (31.2)
Pre-stroke Activity score, mean ± SD	64	2.0 (0.5)	2.0 (0.5)	1.9 (0.6)
Pre-stroke mRS, mean ± SD	3	0.6 (0.8)	0.6 (0.8)	0.7 (0.9)
Pre-stroke fatigue self-reported, n (%)	8			
Yes		119 (25.1)	74 (19.6)	45 (46.9)
Dominant hand, n (%)	6			
Right		433 (91.4)	345 (91.3)	88 (91.7)
Left		27 (5.7)	22 (5.8)	5 (5.2)
No dominant side		8 (1.7)	8 (2.1)	0 (0.0)
Years of education, mean ± SD	0	12.8 (3.8)	12.8 (3.7)	12.7 (3.9)
Educational degree, n (%)	16			
Unskilled		165 (34.8)	123 (32.5)	42 (43.8)
Skilled worker		108 (22.8)	91 (24.1)	17 (17.7)
University degree		185 (39.0)	150 (39.7)	35 (36.5)
Affected side, n (%)	4			
Right		194 (40.9)	157 (41.5)	37 (38.5)
Left		199 (42.0)	155 (41.0)	44 (45.8)
Bilateral		15 (3.2)	12 (3.2)	3 (3.1)
Not relevant		53 (11.2)	43 (11.4)	10 (10.4)
Unknown		9 (1.9)	9 (2.4)	0 (0.0)
Stroke severity, mean ± SD				
NHISS at admission	8	3.6 (4.8)	3.6 (5.02)	3.6 (4.1)
NHISS at day 1	12	2.7 (3.9)	2.51 (4.0)	3.23 (3.5)
NIHSS at day 7	17	1.7 (2.5)	1.6 (2.5)	2.1 (2.4)
Oxfordshire classification, n (%)	2			
TACI		14 (3.0)	13 (3.4)	1 (1.0)
PACI		159 (33.5)	129 (34.1)	30 (31.2)
LACI		127 (26.8)	103 (27.2)	24 (25.0)
POCI		95 (20.0)	72 (19.0)	23 (24.0)
Hemorrhagic		37 (7.8)	28 (7.4)	9 (9.4)
Not classifiable		40 (8.5)	31 (8.2)	9 (9.4)
TOAST classification, n (%)	51			
Atherosclerosis		46 (9.7)	36 (9.5)	10 (10.4)
Cardiac embolic		95 (20.0)	76 (20.1)	19 (19.8)
Small vessel disease		100 (21.1)	78 (20.6)	22 (22.9)
Other		12 (2.5)	9 (2.4)	3 (3.1)
Unknown		170 (35.9)	140 (37.0)	30 (31.2)
3-months measurement				
mRS, mean ± SD	11	1.4 (1.0)	1.3 (1.0)	1.8 (0.9)
MoCA, mean ± SD	30	24.5 (4.1)	24.4 (4.3)	24.9 (3.5)
Cornell, mean ± SD	326	2.8 (3.7)	2.4 (3.5)	4.1 (3.9)
GDS, mean ± SD	14	1.9 (1.0)	1.9 (1.0)	2.0 (1.0)

(*continued on next page*)

Table 1 (*Continued*)

Characteristics	Missing	Fatigue (FSS-7>=5), 18 Months		
		Overall N=474	No N=378 (79.7)	Yes N=96 (20.3)
Trail Making Test A, s, mean ± SD	39	58.9 (48.4)	57.0 (45.9)	66.5 (57.1)
Trail Making Test B, s, mean ± SD	47	146.2 (81.3)	143.0 (79.5)	159.0 (87.3)
FAS, mean ± SD	0	32.6 (21.0)	33.1 (20.8)	30.7 (21.6)
NPI-Q, mean ± SD	29	1.4 (2.3)	1.2 (2.2)	2.1 (2.3)
HADS Anxiety, mean ± SD	0	3.4 (3.5)	3.0 (3.4)	4.7 (3.7)
HADS Depression, mean ± SD	0	3.1 (3.2)	2.8 (2.9)	4.6 (3.8)
SPPB, mean ± SD	36	9.6 (2.9)	9.7 (2.9)	8.9 (3.0)
Gait speed, s, mean, (SD)				
preferred	65	8.9 (3.6)	8.7 (3.5)	9.8 (4.1)
maximum	68	6.7 (2.8)	6.6 (2.8)	7.4 (2.8)
dual task	73	10.6 (5.5)	10.3 (4.5)	11.9 (8.5)
Peg test (pegs pr min), mean ± SD				
best	66	0.5 (0.2)	0.5 (0.2)	0.5 (0.2)
lowest	66	0.4 (0.2)	0.4 (0.2)	0.4 (0.2)
difference	66	0.1 (0.2)	0.1 (0.2)	0.1 (0.1)
PA in min/week, mean ± SD				
Total MPA	113	265.2 (164.3)	272.6 (162.0)	235.8 (170.9)
Total LPA	113	244.5 (124.9)	250.2 (127.1)	222.0 (113.6)
Upright	113	1894.2 (775.1)	1929.2 (781.2)	1756.1 (739.5)
Standing	113	1384.6 (603.1)	1406.5 (616.9)	1298.4 (540.7)
Long-bout (≥10 min) MPA	113	18.0 (51.5)	19.4 (55.6)	12.6 (29.9)
Short-bout (<10 min) MPA	113	247.1 (143.4)	253.2 (139.7)	223.1 (156.1)
FSS-7-score, mean ± SD	46	3.2 (1.8)	2.9 (1.7)	4.5 (1.8)

NOTE: Table 1 presents characteristics of all participants measured in the acute phase (baseline) and subacute phase (three months after stroke). Pre-stroke characteristics were measured retrospectively at baseline. All variables were used to develop the prediction models.
Abbreviations: mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; TACI, total anterior circulation infarct; PACI, partial anterior circulation infarct; LACI, lacunar infarct; POCI, posterior circulation infarct; TOAST, Trial of Org 10172 in Acute Stroke Treatment; MoCA, Montreal Cognitive Assessment; GDS, Global Deterioration Scale; FAS, Verbal Fluency Test Letter; NPI-Q, Neuropsychiatric Inventory questionnaire; HADS, Hospital Anxiety and Depression Scale; MPA, moderate physical activity; LPA, light physical activity;

Analysis

In a machine learning set up, the data were randomly split into a training data set (75%) and a test data set (25%). The split was stratified to ensure a balanced proportion of fatigue in both data sets. Missing values were imputed with missForest function.[42,43] The training and test data set were imputed separately. The imputation model included baseline and 3-month data. Each scale was imputed on an item level before new sum-scores were calculated. The 18-month fatigue score was not part of the imputation model. As fatigue occurred in only 20% of the population, we used an oversampling method with the training data set. The Random Over-Sampling Examples were used to balance the proportion of people with and without fatigue to reduce bias towards majority classes in the training data.[44] Random Over-Sampling Examples is a bootstrap-driven method that generates synthetic examples from conditional density estimates of both classes, accommodating continuous and categorical data.[45]

For the classification modeling we performed a cross-validation procedure. The process of randomly splitting the training data in 2 parts was repeated 10 times. This gave 20 different training/validation sets that were used to evaluate model performance in the training data set. We explored the performance of recursive partitioning using the rpart package,[46] random forest modeling using the ranger package,[47] and extreme gradient boosting models using the xgboost package[48] in R (version 4.3.0).[b] The performance in all models was evaluated with the area under the Receiver Operator Characteristic curve (ROC-AUC). Recursive partitioning involves iteratively dividing data into subsets based on feature values to create a hierarchical structure that optimally separates the data. For the recursive partitioning models, we tuned minimum node size, tree depth and cost complexity parameters. Extreme Gradient Boosting (XG-boost) combines multiple weak predictive models into a robust ensemble. It iteratively refines predictions by minimizing a loss function through gradient descent and employs regularization techniques to prevent overfitting. For the XG-boost models, we tuned the learn rate, tree depth and sample size. Finally, the prediction performance was validated with the test data set and the training model was refitted with the 10 most influential variables. The positive predictive value (PPV) was calculated to estimate the proportion of patients with a positive fatigue test who were correctly identified by the model. The negative predictive value (NPV) was defined as the proportion of patients with a negative test result who were correctly identified by the model. A simplified outline of the model development process is presented in figure 1. The script is available in the supplemental material (available online only at http://www.archives-pmr.org/).

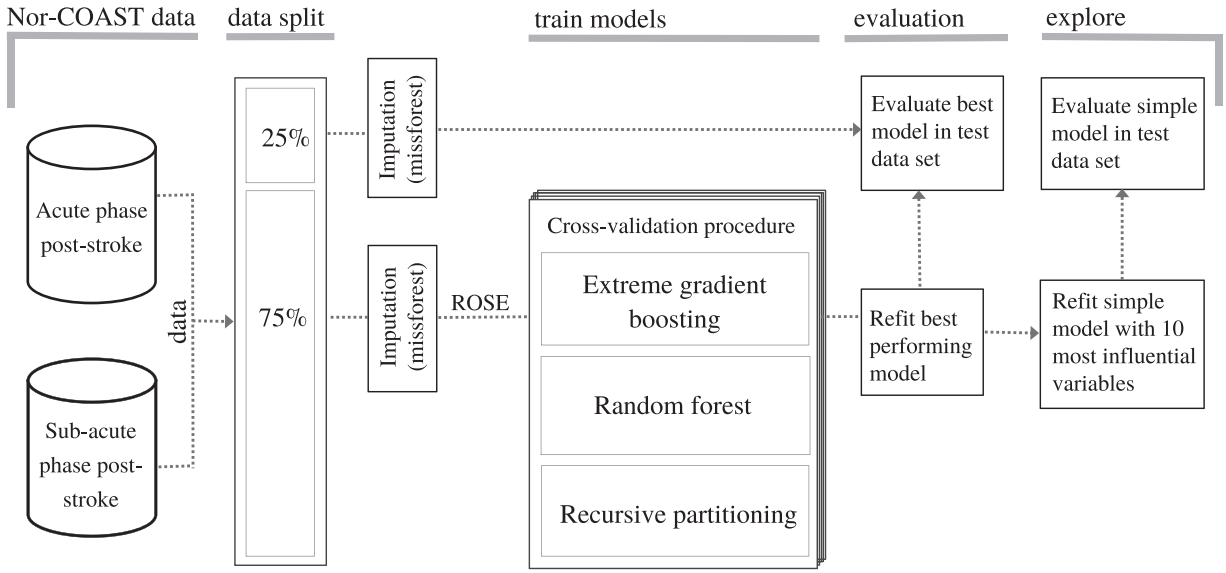

Fig 1 Model development strategy. The dataset from baseline and 3-months testing was divided into training data set (75%) and test data set (25%). Model development and model structure analysis were done using the training set. Model performance was validated using the test data set.

The Kuder Richardson (KR) coefficient was employed to assess the agreement between predicted and actual levels of fatigue, with the response scale of this variable categorized as no/yes. The KR coefficients can range from 0 to 1, where higher values indicate greater internal consistency.[49]

Results

Altogether, 474 participants (mean ± SD age; 70.5 (11.3), 59% male), mainly suffering from mild to moderate stroke, were included in the study. All participants were home-dwelling prior to the stroke. In total, 119 (25.1%) participants reported pre-stroke fatigue, and the mean ± SD FSS-7-scores at 3 and 18 months post-stroke were 3.18 (1.82) and 3.24 (1.78), respectively. The estimates of all 45 prediction variables are presented in table 1.

Selecting a prediction model for post-stroke fatigue

Fatigue was observed in 20.3% (N=96) of all subjects 18 months post-stroke. During the cross-validation procedure, the best tuned recursive partitioning model had a minimum node number of 40, a tree depth of 8, and cost complexity rate of 10-10. On average the recursive partitioning models correctly identified 68% of subjects with a mean FSS-7-score of 5 or higher (sensitivity) and correctly classified 60% of the subjects with a mean FSS-7-score lower than 5 (specificity). The ROC-AUC was estimated as 0.69.

The most appropriate tuning of the extreme gradient boosting model was achieved at a tree depth of 15, a learn rate of 0.001 and an exposed sample size of 0.55. On average the XG-boost model achieved a ROC-AUC of 0.81, and correctly classified 74% of the subjects with fatigue and 72% of the subjects without fatigue.

The random forest model showed the best performance during the cross-validation procedure with a ROC-AUC of 0.82. This was slightly above the estimated ROC-AUC of the XG-boost model. On average the random forest model correctly identified 76% of fatigued subjects, and 72% of nonfatigued subjects. The overall performance measures for the 3 different models are given in table 2.

The ability to predict fatigue in new data

In the test data 26 participants were classified as having fatigue at 18 months post-stroke. The refitted random forest model identified 18 (sensitivity: 0.69, 95% CI: 0.50, 0.86) of these subjects. Of the 101 subjects that did not report fatigue 18 months post-stroke 75 were correctly identified (specificity: 0.74, 95% CI: 0.66, 0.83). The model performance as shown in figure 2 (ROC-AUC = 0.79, 95% CI: 0.69, 0.87) was fairly close to the performance seen throughout the cross-validation procedure. If the model predicted fatigue, we could expect that 41% of the subjects would truly have fatigue 18 months post-stroke (PPV=0.41). If the model predicted no fatigue, we could expect that 90% of the subjects would truly not have fatigue 18 months post-stroke (NPV=0.90). See table 3 for more details.

The KR reliability coefficient was 0.54.

Explorative simple model

The variable importance plot (fig 3) shows the 10 most influential variables identified in the random forest model. Those were

Table 2 Model selection: Performance of classification models

Model	Train Data Set		
	ROC-AUC	Sensitivity	Specificity
Recursive partitioning	0.69	0.68	0.60
Random forest	0.82	0.76	0.72
Extreme gradient boosting	0.81	0.74	0.72

NOTE: Table 2 presents the performance of each of the classification models, recursive partitioning, random forest, and extreme gradient boost modeling.

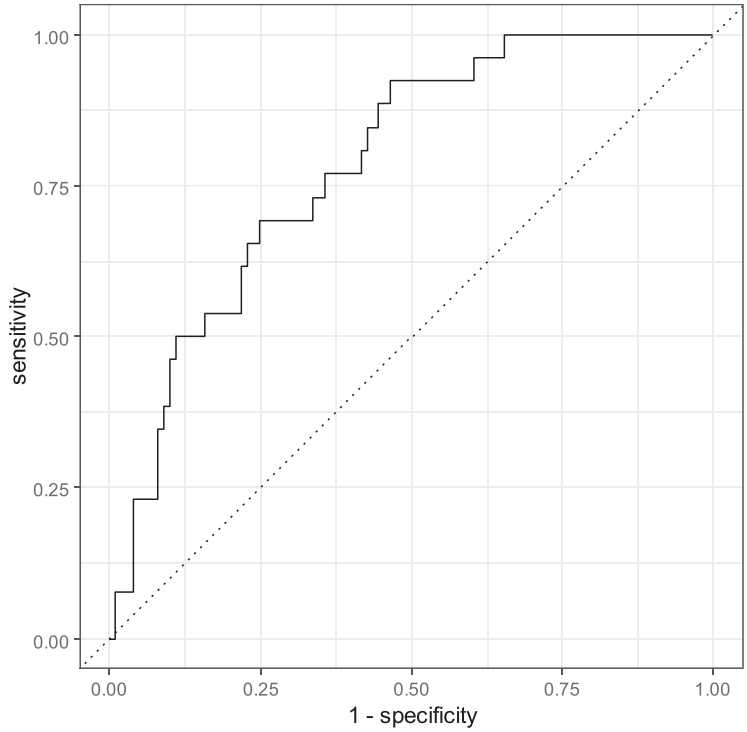

Fig 2 ROC-AUC of the best performing model, random forest.

Table 3 Performance of the random forest model in new data

		Predicted Fatigue						
		No	Yes	ROC-AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Fatigue	No	75	26	0.79 (0.69, 0.87)	0.69 (0.50, 0.86)	0.74 (0.66, 0.83)	0.41 (0.26, 0.57)	0.90 (0.83, 0.96)
	Yes	8	18					

NOTE: Table 3 represents the measures of prediction and prediction performance.

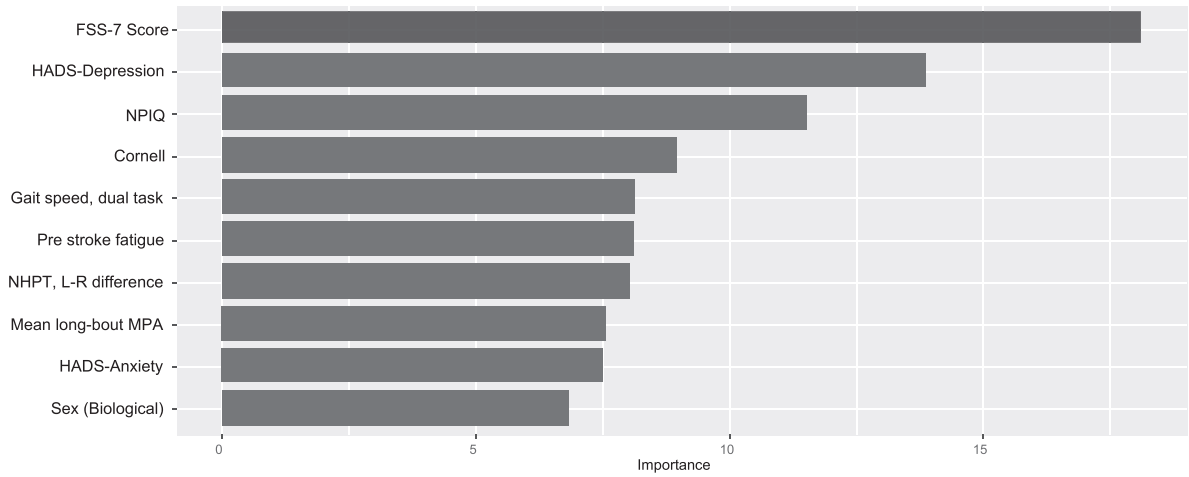

Fig 3 Variable importance plot. The 10 most influential prediction variables from the baseline and 3-months examinations are presented from descending importance. The top variables contribute more to the model than the bottom ones and have higher predictive power in classifying fatigue at 18 months post-stroke. Abbreviations: FSS-7, fatigue severity scale; HADS, hospital anxiety and depression scale; MPA, moderate physical activity; NPI-Q, neuropsychiatric inventory questionnaire; NHPT, nine hole peg test.

fatigue (FSS-7), symptoms of anxiety and depression (HADS), neuropsychiatric symptoms (Neuropsychiatric Inventory questionnaire and Cornell scale), gait speed during dual task, Nine Hole Peg Test (difference between hands), and mean weekly long-bout moderate PA, all measured 3 months post-stroke; in addition to self-reported pre-stroke fatigue, and sex, assessed at baseline. We were able to train an explorative simpler model that correctly predicted fatigue in 17 of the 26 participants with an FSS-7-score of 5 or above after refitting the train model with these 10 variables using the same random forest approach. This resulted in a sensitivity of 0.65 (95% CI 0.51, 0.78). The simple model successfully predicted absence of fatigue in 73 of the 101 patients with an FSS-7-score less than five, yielding a specificity of 0.72 (95% CI 0.67, 0.80). The explorative simple model's ROC-AUC was 0.79 (95% CI 0.73, 0.86), with a PPV of 0.38 (95% CI 0.31, 0.52), and NPV 0.89 (95% CI 0.85, 0.92).

Discussion

By employing a machine-learning methodology to analyze extensive data from individuals who have survived stroke, we have determined that the random forest model offers the most accurate predictive capability compared to other models. The model demonstrated satisfactory performance in predicting future fatigue in a new data set. The prevalence of fatigue doubled from around 20% in the general sample to 40% in subjects identified by the model being at risk of fatigue. Moreover, our results suggest that a streamlined approach utilizing a simple model, incorporating the 10 most influential variables identified, could suffice for identifying individuals at risk of fatigue during the long-term post-stroke chronic phase, while maintaining a satisfactory level of predictive performance. On the other hand, the model is even better in ruling out the risk of fatigue, as 90% of those with a negative test will not suffer from PSF at 18 months.

To our knowledge, this is the first study to use a machine learning approach to predict PSF. Previous attempts to predict fatigue symptoms post-stroke were based on regression modeling only.[7,17,50-52] In 2021, Su et al developed a prediction model including sex, pre-stroke sarcopenia, acute phase fatigue, dysphagia, and depression, which effectively predicted the risk of PSF at the discharge from hospital with good discrimination (concordance-index = 0.801, 95% CI: 0.700-0.902).[50] Another study found that Montreal Cognitive Assessment score as proxy for cognitive function in the acute stage post-stroke could not account for symptoms of fatigue.[51] Their hypothesis that the link between these 2 variables may only become evident in later stages after the stroke, aligns with our findings. Yet another regression-based prediction model was developed to identify predictors 3 months after stroke that may aid identify individuals with increasing symptoms of fatigue over the course of 24 weeks. Results showed that only FSS-7 score at discharge was an independent predictor, accurately identifying 7.9% of patients with increasing fatigue symptoms.[7] This finding, despite its limited predictive value, is in line with previous results[8] as well as those of our study, indicating that the level of fatigue experienced shortly after a stroke is a crucial predictor of fatigue in later stages post-stroke. However, it appears that the predictive value of early fatigue may diminish over even longer follow-up periods: a recent study found that fatigue assessed after acute stroke could not predict the level of fatigue 7 years later. Yet, acute stroke severity as well as female sex were substantiated as predictors of higher fatigue scores after 7 years.[17] In fact, sex has repetitively been linked to the prevalence of PSF,[15,17,50] and was also identified as 1 of the 10 most influential predictors of fatigue in our study. Previously suggested explanations of the observed link between female sex and PSF include potential sex-differences in expressing feelings of tiredness and effects of biological mechanisms related to sleep.[17] The rather low positive predictive value and high negative predictive value indicate that the model is better at ruling out fatigue at 18 months post-stroke than at confirming its presence. In general, this can be acceptable in situations where the benefit of identifying too many cases exceeds the potential disadvantages and where the risk of overtreatment is low. According to the moderate KR coefficient, caution should be exercised in interpreting fatigue scores.

Strengths

One of the key benefits of utilizing data from Nor-COAST is the prospective multicenter design, the large sample size, and the comprehensive test-battery, comprising widely used, reliable, and validated assessment tools.[20] Nor-COAST was designed to identify predictors for post-stroke cognitive impairment, with a test battery that is sensitive to mild cognitive impairment including nonamnestic deficits.[20] The latter is important, as many previous studies have excluded patients with cognitive deficit from their study populations.[15] Given the complexity of potential PSF prediction variables,[5] the use of machine learning is another strength.[19]

Study limitations

Firstly, we want to point out that fatigue is a poorly defined symptom, which can make quantification and classification difficult.[2] Also, the differentiation of fatigue symptoms from motoric impairment after stroke may be challenging for some individuals.[53]

Secondly, even though our study population is assumed to be representative for the majority of the national stroke population who suffers from mild to moderate impairment,[54] selection bias must be considered a threat. Particularly individuals with severe impairment may be underrepresented.

Thirdly, when modeling a large set of prediction variables especially when introducing synthetic data, we must assume that the variables identified as most important for prediction, or their order, may vary when the analyses are repeated. Further it is not possible to draw conclusions on mechanisms underlying the association of predictors and outcome but provide further observational information only. Even if the most influential variables may make sense in a clinical setting, it is important to keep in mind that we used mathematical models, that are solely suited for generating theoretical models about the relationship between predictors and dependent variables.

Fourthly, previous studies have suggested that other variables like stressful life-events,[55] sleep patterns,[15,56] serum and cytokines levels,[57] dysphagia and pre-stroke sarcopenia,[50] magnetic resonance imaging[58] and inflammation[59] may be important to consider as well. By adding some of these variables in future models the predictive values might improve even more.

Conclusions

In conclusion, our models demonstrate a promising capability in predicting 18-month PSF. The best performing model, the random forest, showed comparable prediction performance when only using the 10 most influential prediction variables. Our findings highlight the significance of monitoring fatigue early after a

stroke, to consider the prevalence of pre-stroke fatigue, and to assess levels of PA and function in addition to cognitive and neuropsychiatric symptoms.

Suppliers

a. MATLAB version: 9.4; The MathWorks Inc.
b. R version 4.3.0; R Foundation for Statistical Computing

Keywords

Stroke; long-term follow-up; fatigue; prediction; machine learning

Corresponding author

Torunn Askim, Bevegelsessenteret, 311.03.049, Øya, Olav Kyrres gate 13 *E-mail address:* torunn.askim@ntnu.no.

Acknowledgments

We thank all patients for their contributions as well as the Nor-COAST research group and the dedicated study staff at St. Olavs hospital, Haukeland University Hospital, Ålesund Hospital, Vestre Viken Hospital Trust, and Oslo University Hospital.

References

1. Annoni JM, Staub F, Bogousslavsky J, Brioschi A. Frequency, characterisation and therapies of fatigue after stroke. Neurol Sci 2008;29 (Suppl 2):S244–6.
2. Staub F, Bogousslavsky J. Fatigue after stroke: a major but neglected issue. Cerebrovasc Dis 2001;12:75–81.
3. Cumming TB, Packer M, Kramer SF, English C. The prevalence of fatigue after stroke: a systematic review and meta-analysis. Int J stroke 2016;11:968–77.
4. Ingles JL, Eskes GA, Phillips SJ. Fatigue after stroke. Arch Phys Med Rehabil 1999;80:173–8.
5. Wu S, Mead G, Macleod M, Chalder T. Model of understanding fatigue after stroke. Stroke 2015;46:893–8.
6. Lerdal A, Lee KA, Bakken LN, Finset A, Kim HS. The course of fatigue during the first 18 months after first-ever stroke: a longitudinal study. Stroke Res Treat 2012;2012:126275.
7. van Eijsden HM, van de Port IG, Visser-Meily JM, Kwakkel G. Post-stroke fatigue: who is at risk for an increase in fatigue? Stroke Res Treat 2012;2012:863978.
8. Snaphaan L, van der Werf S, de Leeuw FE. Time course and risk factors of post-stroke fatigue: a prospective cohort study. Eur J Neurol 2011;18:611–7.
9. Duncan F, Wu S, Mead GE. Frequency and natural history of fatigue after stroke: a systematic review of longitudinal studies. J Psychosom Res 2012;73:18–27.
10. Aarnes R, Stubberud J, Lerdal A. A literature review of factors associated with fatigue after stroke and a proposal for a framework for clinical utility. Neuropsychol Rehabil 2020;30:1449–76.
11. Kliem E, Labberton AS, Barra M, Olsen A, Thommessen B, Thomas O, et al. Self-reported cognitive and psychiatric symptoms at 3 months predict single-item measures of fatigue and daytime sleep 12 months after ischemic stroke. Front Neurol 2022;13:944586.
12. Radman N, Staub F, Aboulafia-Brakha T, Berney A, Bogousslavsky J, Annoni J-M. Poststroke fatigue following minor infarcts A prospective study. Neurology 2012;79:1422–7.
13. Parks NE, Eskes GA, Gubitz GJ, Reidy Y, Christian C, Phillips SJ. Fatigue impact scale demonstrates greater fatigue in younger stroke survivors. Can J Neurol Sci 2012;39:619–25.
14. Schepers VP, Visser-Meily AM, Ketelaar M, Lindeman E. Poststroke fatigue: course and its relation to personal and stroke-related factors. Arch Phys Med Rehabil 2006;87:184–8.
15. Zhang S, Cheng S, Zhang Z, Wang C, Wang A, Zhu W. Related risk factors associated with post-stroke fatigue: a systematic review and meta-analysis. Neurol Sci 2021;42:1463–71.
16. Lanctôt KL, Lindsay MP, Smith EE, Sahlas DJ, Foley N, Gubitz G, et al. Canadian stroke best practice recommendations: mood, cognition and fatigue following Stroke, 6th edition update 2019. Int J Stroke 2019;15:668–88.
17. Pedersen A, Almkvist E, Holmegaard L, Lagging C, Redfors P, Blomstrand C, et al. Fatigue 7 years post-stroke: Predictors and correlated features. Acta Neurol Scand 2022;146:295–303.
18. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J 2017;38:1805–14.
19. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: making sense of the differences. Knee Surg Sports Traumatol Arthrosc 2022;30:753–7.
20. Thingstad P, Askim T, Beyer MK, et al. The Norwegian Cognitive impairment after stroke study (Nor-COAST): study protocol of a multicentre, prospective cohort study. BMC Neurol 2018;18:193.
21. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121–3.
22. Lerdal A, Kottorp A. Psychometric properties of the Fatigue Severity Scale—Rasch analyses of individual responses in a Norwegian stroke cohort. Int J Nurs Stud 2011;48:1258–65.
23. Lerdal A, Wahl A, Rustøen T, Hanestad BR, Moum T. Fatigue in the general population: a translation and test of the psychometric properties of the Norwegian version of the fatigue severity scale. Scand J Public Health 2005;33:123–30.
24. Holmen J, Midthjell K, Forsén L, Skjerve K, Gorseth M, Oseland A. [A health survey in Nord-Trøndelag 1984-86. Participation and comparison of attendants and non-attendants]. Tidsskr Nor Laegeforen 1990;110:1973–7.
25. Kurtze N, Rangul V, Hustvedt BE, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study: HUNT 1. Scand J Public Health 2008;36:52–61.
26. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604–7.
27. Brott T, Adams HP, Olinger CP, Marle JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke 1989;20:864–70.
28. Bamford J, Sandercock P, Dennis M, Warlow C, Burn J. Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 1991;337:1521–6.
29. Adams Jr HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 1993;24:35–41.
30. Hachinski V, Iadecola C, Petersen RC, Breteler MM, Nyenhuis DL, Black SE, et al. National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. Stroke 2006;37:2220–41.
31. Munthe-Kaas R, Aam S, Ihle-Hansen H, Lydersen S, Knapskog AB, Wyller TB, et al. Impact of different methods defining post-stroke neurocognitive disorder: The Nor-COAST study. Alzheimers Dement (N Y) 2020;6:e12000.
32. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept Mot Skills 1958;8:271–6.
33. Bechtoldt HP, Benton AL, Fogel ML. An application of factor analysis in neuropsychology. Psychol Rec 1962;12:147–56.

34. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982;139:1136–9.
35. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–9.
36. Cummings J. The neuropsychiatric inventory: development and applications. J Geriatr Psychiatry Neurol 2020;33:73–84.
37. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.
38. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell scale for depression in dementia. Biol Psychiatry 1988;23:271–84.
39. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49:M85–94.
40. Heller A, Wade DT, Wood VA, Sunderland A, Hewer RL, Ward E. Arm function after stroke: measurement and recovery over the first three months. J Neurol Neurosurg Psychiatry 1987;50:714–9.
41. Taraldsen K, Askim T, Sletvold O, Einarsen EK, Bjåstad KG, Indredavik B, et al. Evaluation of a body-worn sensor system to measure physical activity in older people with impaired function. Phys Ther 2011;91:277–85.
42. missForest: Nonparametric Missing Value Imputation using Random Forest. R package version 1.4.; 2013.
43. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28:112–8.
44. Lunardon N, Menardi G, Torelli N. ROSE: a package for binary imbalanced learning. R J 2014;6.
45. Menardi G, Torelli N. Training and assessing classification rules with imbalanced data. Data Min Knowl Discov 2014;28:92–122.
46. rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15.; 2019.
47. Wright MN, Ziegler A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw 2017;77:1–17.
48. xgboost: Extreme Gradient Boosting. R package version 1.5.0.2.; 2021.
49. Kuder GF, Richardson MW. The theory of the estimation of test reliability. Psychometrika 1937;2:151–60.
50. Su Y, Yuki M, Hirayama K, Otsuki M. Development and internal validation of a nomogram to predict post-stroke fatigue after discharge. J Stroke Cerebrovasc Dis 2021;30:105484.
51. Holmberg J, Jondell B, Abzhandadze T, Sunnerhagen KS. Very early cognitive screening and self-reported feeling of fatigue three months after stroke. Front Hum Neurosci 2021;15:742105.
52. Braaten RS, Askim T, Gunnes M, Indredavik B. Fatigue and activity after stroke. Secondary results from the Life After Stroke study. Physiother Res Int 2020;25:e1851.
53. Mead G, Lynch J, Greig C, Young A, Lewis S, Sharpe M. Evaluation of fatigue scales in stroke patients. Stroke 2007;38:2090–5.
54. Kuvås KR, Saltvedt I, Aam S, Thingstad P, Ellekjær H, Askim T. The risk of selection bias in a clinical multi-center cohort study. Results from the Norwegian Cognitive Impairment After Stroke (Nor-COAST) Study. Clin Epidemiol 2020;12:1327–36.
55. Tao C, Gu M, Xu P, Wang J, Xiao L, Gui W, et al. Stressful life events can predict post-stroke fatigue in patients with ischemic stroke. Eur J Neurol 2021;28:3080–8.
56. White JH, Gray KR, Magin P, Attia J, Sturm J, Carter G, et al. Exploring the experience of post-stroke fatigue in community dwelling stroke survivors: a prospective qualitative study. Disabil Rehabil 2012;34:1376–84.
57. Ormstad H, Aass HCD, Amthor K-F, Lund-Sørensen N, Sandvik L. Serum cytokine and glucose levels as predictors of poststroke fatigue in acute ischemic stroke patients. J Neurol 2011;258:670–6.
58. Cotter G, Salah Khlif M, Bird L, M EH, Brodtmann A, Egorova-Brumley N. Post-stroke fatigue is associated with resting state posterior hypoactivity and prefrontal hyperactivity. Int J Stroke 2021:17474930211048323.
59. Liu X, Wang B, Wang X, Tian M, Wang X, Zhang Y. Elevated plasma high-sensitivity C-reactive protein at admission predicts the occurrence of post-stroke fatigue at 6 months after ischaemic stroke. Eur J Neurol 2020;27:2022–30.