

Supplementary material has been published as submitted. It has not been copyedited, typeset or checked for scientific content by Journal of Rehabilitation Medicine

## Appendix S1

### Expanded Methods

*Objective one: To develop the s-MoCA-SWE*

Principal component analysis (PCA) and boosted regression tree (BRT) analyses were performed to identify a subset of the Montreal Cognitive Assessment (MoCA) items with the potential to form a short Swedish version of the MoCA (s-MoCA-SWE).

PCA with varimax rotation and the principal axis factoring method was performed. The PCA was performed in two steps.

- Step 1. A full model including ten variables of the MoCA was examined. Bartlett's and Kaiser–Meyer–Olkin values were evaluated (values  $<0.05$  and  $\geq 0.5$ , respectively). The correlation between components was studied to determine the type of rotation to be used. The communality values of ten variables were explored (if an item had a communality value  $\leq 0.3$ , it was excluded from the next step as such an item could cause noise in the model).
- Step 2. A weak correlation was found between the components; therefore, the varimax method was chosen for the rotation. The final number of components was determined based on the scree plot, and components having a value  $\geq 1$  were accepted.

BRT analysis was performed using the XGBoost library. This is a supervised machine learning algorithm. An outcome was a reference standard instrument, dichotomised MoCA score (impaired cognition,  $\leq 25$  points coded as 1), and explanatory variables were individual items of the MoCA. The model-building process was as follows:

- The outcome was dichotomised MoCA score (1 = impaired cognition,  $\leq 25$  points; 0 = normal cognition,  $\geq 26$  points). Ten MoCA items were treated as independent variables.
- The type of BRT analysis was specified as “classification”. The frame of the BRT model was created using 2000 trees and six parameters that could be adjusted. These six parameters were as follows:
  - Minimum number of data points in a node required for the node to be split further.
  - The depth of the tree.
  - Reduction in the loss function required to split a tree further.
  - Number of variables randomly sampled when building a tree.
  - Rate at which boosting algorithm adapted from iteration to iteration.
  - Amount of data exposed to the fitting process.

- The space-filling method was used to construct a grid for these six parameters.
- The data were split into training (75% of data used for hyperparameter tuning) and testing sets (25% of data used to evaluate model performance).
- The training dataset was prepared for a 10-fold cross-validation, which was repeated five times. An initial BRT model was estimated by assembling previously specified parameters that could be adjusted: the BRT model frame, space-filling grid, and data with 10-fold cross-validation. The predictions of the initial models were saved.
- The testing data set was used for validation of the model developed in the training set. Model predictions saved from training dataset were fitted in the testing data set. The final BRT model was selected by determining the model with the best area under the curve (AUC). Variable importance values were obtained from the training dataset. MoCA variables with a variable importance level  $\geq 5.0$  (relative influence value \*100) were considered to have the potential to form the index test s-MoCA-SWE/BRT. The total scores for the selected variables were calculated.

Two possible sets of items for the s-MoCA-SWE were available for final analysis: the s-MoCA-SWE/PCM and s-MoCA-SWE/BRT. Individual binary regression analyses were subsequently fitted for choosing the optimal subset of MoCA items. The dataset was divided into 10 folds. Each fold contained 90% of the testing data and 10% of the data was set as a holdout set. The regression model was fitted for a given 90% subset of data and identified threshold for best sensitivity. The threshold was further tested on the holdout set (10%). All abovementioned steps were performed for each fold separately and aggregated results were calculated.