Relevance of Complaint Severity in Predicting the Progression of Subjective Cognitive Decline and Mild Cognitive Impairment: A Machine Learning Approach

Arturo Xosé Pereiro[a,*], Sonia Valladares-Rodríguez[b], Alba Felpete[a], Cristina Lojo-Seoane[a], María Campos-Magdaleno[a], Sabela Carme Mallo[a], David Facal[a], Luis Anido-Rifón[b], Sylvie Belleville[c,d] and Onésimo Juncos-Rabadán[a]

[a]*Department of Developmental Psychology, University of Santiago de Compostela, Santiago de Compostela, Spain*
[b]*School of Telecommunication Engineering, University of Vigo, Vigo, Spain*
[c]*Research Center of the Institut Universitaire de Gériatrie de Montréal, Montreal, Canada*
[d]*Université de Montréal, Montréal, Canada*

Abstract.

Background: The presence of subjective cognitive complaints (SCCs) is a core criterion for diagnosis of subjective cognitive decline (SCD); however, no standard procedure for distinguishing normative and non-normative SCCs has yet been established.

Objective: To determine whether differentiation of participants with SCD according to SCC severity improves the validity of the prediction of progression in SCD and MCI and to explore validity metrics for two extreme thresholds of the distribution in scores in a questionnaire on SCCs.

Methods: Two hundred and fifty-three older adults with SCCs participating in the Compostela Aging Study (CompAS) were classified as MCI or SCD at baseline. The participants underwent two follow-up assessments and were classified as cognitively stable or worsened. Severity of SCCs (low and high) in SCD was established by using two different percentiles of the questionnaire score distribution as cut-off points. The validity of these cut-off points for predicting progression using socio-demographic, health, and neuropsychological variables was tested by machine learning (ML) analysis.

Results: Severity of SCCs in SCD established considering the 5th percentile as a cut-off point proved to be the best metric for predicting progression. The variables with the main role in conforming the predictive algorithm were those related to memory, cognitive reserve, general health, and the stability of diagnosis over time.

Conclusion: Moderate to high complainers showed an increased probability of progression in cognitive decline, suggesting the clinical relevance of standard procedures to determine SCC severity. Our findings highlight the important role of the multimodal ML approach in predicting progression.

Keywords: Cognitive dysfunction, dementia, diagnosis, follow-up studies

*Correspondence to: Dr. Arturo X. Pereiro. Department of Developmental Psychology, Xosé María Suárez Núñez Street, Campus Sur. Santiago de Compostela, Galicia ES–15782, Spain. Tel.: +34 881813651; E-mail: arturoxose.pereiro@usc.es.

INTRODUCTION

Mild cognitive impairment (MCI) has been described as a prodromal stage of dementia in individuals with objective cognitive decline, cognitive

complaints, and well preserved functionally [1]. People with MCI have a considerably higher risk of conversion to dementia [2], particularly those subtypes with multiple impaired domains [3–5]. A new pre-symptomatic stage, subjective cognitive decline (SCD), has recently been proposed as falling between the normative age-related decline in cognition and MCI as a prodromal stage of dementia [6, 7]. Individuals diagnosed with SCD experience and express persistent subjective cognitive complaints (SCCs) of decline in cognitive ability, which cannot be explained by medical conditions or substance use, with no evidence of objective decline in cognition or functionality beyond the threshold suggesting subtle impairments. Although research on the progression of SCD is scarce, some results from studies carried out in community and clinical settings have suggested a significantly higher risk of progression to MCI than non-complainers [8, 9]. In addition, participants with SCCs (focusing particularly on memory) at baseline have shown an increased risk of non-normative cognitive decline and eventual progression to MCI and dementia [10–12].

Information about neuropathological progression across pre-symptomatic and prodromal stages of dementia, specifically Alzheimer's disease, is increasingly available, and quite a high consensus has been reached regarding the continuous nature of the disease across the stages that precede the diagnosis [13]. In accordance, similar socio-demographic (e.g., age, gender, education, professional attainment), health (e.g., comorbidity, functionality, depression) and cognitive (e.g., memory, executive functions) measures have been proposed as valid predictors of progression to dementia in both SCD [9–11, 14–18] and MCI [19–28]. Progression to dementia in SCD and MCI can potentially be predicted from a large number of sociodemographic, health, and cognitive variables. Machine learning (ML) is therefore a suitable approach for extracting complex patterns and analyzing the precision with which the data fit the progression observed in the different groups [29].

Despite the nature of SCD and MCI as respectively pre-symptomatic and prodromal stages of dementia, some participants with MCI [4] and to a greater extend those with SCD [11, 30] do not progress or convert to dementia during a standard follow-up period. The problem of the validity of predictors of conversion, particularly in SCD, must be addressed by improving the procedures used to measure the core diagnostic criteria and to establish the clinical significance of these [31, 32]. In both MCI (i.e., cognitive

performance) and SCD (i.e., complaints), compliance with the core criteria must be verified using indicators that already generally change in normative aging, making diagnosis of these conditions particularly challenging. Cut-off points ranging from 1 to 2 SD below mean are used to determine cognitive impairment in MCI [33]. However, in SCD the significance of complaints is evaluated and established in different ways across studies [11, 12]. No standard procedure for distinguishing normative and non-normative complaints has yet been established. As SCCs are common in normative aging [11], quantification of the complaints may be important for distinguishing participants undergoing normative aging from those with pre-symptomatic and prodromal stages of dementia. The frequency of perceived difficulties is often rated using items that assess cognition on a scale ranging from ability to disability, in order to quantify the importance of SCCs and to characterize SCD psychometrically [6, 31, 32]. Thus, exploration of a valid threshold that can differentiate participants by the importance of SCCs and successfully predict progression to dementia appears worthwhile.

Separate predictions of progression in SCD, particularly in MCI participants, have been studied in recent years by using neurocognitive markers. Despite the widespread idea that SCD and MCI are part of a continuum of deterioration that can lead to dementia, we are not aware of any study specifically aimed at predicting the progression of participants with these pre-symptomatic and prodromal stages of the disease.

The main objectives of this study were: 1) to determine whether differentiation of participants with SCD according to SCC severity improves the validity of the prediction of progression of pre-symptomatic and prodromal states of dementia using ML algorithms; and 2) to explore the validity of the quantifying of SCCs using two extreme cut-off points for the classification of low and high complainers in predicting progression to dementia, as a preliminary study of the clinical significance of complaints. For this purpose, we consider as extreme thresholds the scores corresponding to the 5th or 95th percentiles of the distribution of scores in a questionnaire on SCCs.

MATERIALS AND METHODS

Participants

A sample of 262 subjects of age 50–87 years already participating in the Compostela Aging Study (CompAS) and who completed clinical and

neuropsychological assessments at baseline and two follow-ups were included in the present study. Comp-AS is an ongoing longitudinal project [34] involving the detection of cognitive impairment in patients attending primary care centers in Galicia (an autonomous region in north-western Spain). All participants reported SCCs, including memory problems, to their general practitioners. None of the participants had prior diagnosis of dementia, psychiatric or neurological disorders, severe illness, deafness or blindness; they were not receiving chemotherapy, and they did not consume alcohol or other substances.

Trained psychologists carried out extensive neuropsychological and cognitive assessment of the participants, who were classified at baseline as MCI ($N = 54$) or SCD ($N = 199$) at a special meeting of the research team. MCI diagnosis was diagnosed according to the criteria established by Albert et al. [35] and participants were classified into four MCI subtypes following standard criteria [1]: single-domain amnestic MCI (sda-MCI); multiple-domain amnestic MCI (mda-MCI); single-domain non-amnestic MCI (sdna-MCI); and multiple-domain non-amnestic MCI (mdna-MCI). Participants who reported SCCs but did not exhibit objective cognitive impairment were considered to have SCD. Diagnosis at baseline was corrected conservatively when diagnosis at the first follow-up (18–24 months later) suggested some degree of recovery (i.e., from MCI to SCD, or from multiple-domain MCI to single-domain MCI).

In each successive follow-up assessment, participants were reclassified, using the same criteria as at baseline, as having SCD, sda-MCI, mda-MCI, sdna-MCI, mdna-MCI, or probable dementia (DSM-IV and NINCDS-ADRDA). The between-test interval was around 18–24 months. Nine participants whose diagnosis was reversed from MCI to SCD at the second or third follow up evaluation were excluded from the study, and the final sample therefore consisted of 253 participants. At the third evaluation (54–72 months), participants were classified into four groups according to changes in their cognitive status: participants with SCD at baseline who remained stable (SCD-stable group, $n = 170$, 67.20%); participants with SCD at baseline who progressed to MCI ($n = 20$) or dementia ($n = 9$) (SCD-worsened group, $n = 29$, 11.46%); participants diagnosed with MCI at baseline who remained stable (MCI-stable group, $n = 15$, 5.92%); and participants diagnosed as sda-MCI or sdna-MCI at baseline who progressed to mda-MCI and mdna-MCI ($n = 7$) or to dementia ($n = 32$) at follow-up 1 or 2 (MCI-worsened group, $n = 39$, 15.42%). We assumed, in accordance with Brambati et al. [3] and Campos-Magdaleno et al. [36], that the change from single-domain to multiple-domain indicates cognitive decline, as multiple domain MCI represents the most severely impaired of the MCI subtypes. The sample descriptive scores at baseline for sociodemographic, health, subjective complaints, cognitive status and functional measures are shown in Table 1.

SCD and MCI groups differed significantly in age, cognitive status, and informant-rated IADL variables, but were similar in regard to education, self-rated subjective complaints, gender, and comorbidity.

All participants gave their written informed consent prior to participation in the study. The research project was approved by the Galician Clinical Research Ethics Committee (Xunta de Galicia, Spain), and the study was performed in accordance with the ethical standards established in the Declaration of Helsinki as revised in Seoul 2008.

Materials and procedure

Socio-demographic variables such as age, gender, profession, and years of schooling were collected

Table 1
Descriptive scores at baseline (mean and SD, in parentheses) for sociodemographic, health,
subjective complaints, cognitive status and functional measures in SCD and MCI participants

	SCD group Mean (SD), range	MCI group Mean (SD), range
N (gender)	199 (64.84% women)	54 (61.11% women)
*Age	64.82 (8.63), 50–87	72.68 (7.52), 51–87
Years of schooling	10.17 (4.73), 1–22	9.74 (4.30), 3–25
CCI-health	0.78 (0.86), 0–3	1.00 (0.97), 0–4
QAM-participant	19.23 (4.47), 7–32	19.48 (4.43), 10–32
**Lawton-Brody-IADL-informant	7.55 (0.84), 5–8	6.51 (1.64), 2–8
*MMSE	28.15 (1.55), 21–30	24.60 (2.72), 19–30

CCI, Charlson comorbidity index; QAM, Questionnaire d'auto-évaluation de la Mémoire-short version; MMSE, Mini-Mental State Examination. $^*p < 0.05$. $^{**}p < 0.001$.

using an *ad hoc* questionnaire. The socio-cognitive complexity of tasks associated with professional attainment was ordered, according to the protocol outlined by Garibotto et al. [37] on a scale of 1 to 6 (where 1 = no occupation and 6 = high-ranking civil servant or director, university lecturer, self-employed with high level of responsibility).

With a diagnostic purpose, we used a short Spanish version of the Questionnaire d'auto-évaluation de la Mémoire (QAM) [38, 39] to evaluate the SCCs. The questionnaire consisted of seven items that rated the frequency of prospective and retrospective forgetfulness, distractions, and difficulties in lexical access and spatial orientation on a five-point Likert scale (from 'never' to 'always'). General cognitive status and functionality were assessed using respectively the Mini-Mental State Examination (MMSE) and the Lawton and Brody Index. Memory impairments were assessed using the Short delay free recall test in the Spanish version of the California Verbal Learning Test (CVLR) [40, 41]. Other cognitive domains such as language, attention-calculation, praxis, perception, and executive functioning were assessed using the corresponding subscales of Spanish version of the Cambridge Cognitive Examination (CAMCOG-R) [42].

With a predictive purpose, other measures of health, cognition, language, cognitive reserve, and neuropsychiatric symptoms were collected at baseline using measures from the Spanish versions of the following instruments: A) the Charlson Comorbidity Index (CCI), to quantify the individual's burden of disease from the patient's medical history; B) QAM informant subjective complaints; C) CAMCOG-R total score, as measure of general cognitive status; D) Logical memory I and II, Verbal paired associated I and II, and immediate and delayed auditory memory subtests of the Wechsler Memory Scale–Third Edition (WMS-III); E) CVLT measures for Long delay free recall, intrusions in free and cued recall, total perseverations, recognition accuracy, false positives in recognition, recall discriminability, semantic and serial strategies in long and short delay free recall, serial position effect (corrected score for primacy, middle and recency regions [36]; F) Working memory indexes from adapted versions of the reading and counting span tasks, with the partial-credit unit scoring (i.e., mean proportion of items successfully recalled, relative to serial order within a series); G) the Peabody Picture Vocabulary Test-Revised and the Vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS-III); H) Verbal fluency,

evaluated through the letter, semantic and ideational tests; I) Severity and stress measures of the Neuropsychiatric Inventory Questionnaire (NPI-Q); J) Total score of the Geriatric Depression Scale-15 items (GDS-15); K) Burden on the wellbeing of the informant-caregiver, from the Caregiver Burden Interview (CBI). Years of schooling, professional attainment, and vocabulary scores were also considered proxy measures of cognitive reserve. Total time from baseline to the last study-evaluation and time of stability (i.e., months without changes in diagnostic status) were also calculated for each participant.

Statistical analysis

Missing data in some measures were completed for 22 of the participants (8.69%) by estimating mean values from subsets of participants with similar characteristics (e.g., grouped by cognitive scores, gender, and five-year age intervals). Thus, original and imputed missing data sets were highly correlated according to intraclass correlation coefficients (ICC) ($\rho > 0.80$), showing good reliability values in the method of imputed missing values [43].

This work is a first approach to test the importance of: a) considering or not the severity of complaints in predicting progression to dementia; and b) comparing the prediction validity classifying participants in high and low complainers using two extreme distribution criteria (95th and 5th). Predictive validity for stable and worsening SCD and MCI groups was first analyzed in the four groups (4-group model), i.e., SCD-stable (SCD-s), SCD-worsened (SCD-w), MCI-stable (MCI-s), and MCI-worsened (MCI-w), by ignoring the severity of complaints made by the SCD participants. We subsequently took the severity of the complaints into account, differentiating them into low and high according to the percentile distribution of participants' total QAM scores calculated for each age-groups. We tested two different criteria: a) a stricter criterion considering the 95th percentile (i.e., 50–59 age-group = 21/22; 60–69 age-group = 20/21; 70–79 age-group = 20/21; >80 age-group = 19/20); and b) a less strict criterion considering the 5th percentile (i.e., 50–59 age-group = 18/19; 60–69 age-group = 17/18; 70–79 age-group = 16/17; >80 age-group = 16/17).

Combining diagnostic group, progression, and severity led to analysis of predictive validity in six groups (6-group models): i.e., low SCD-stable (l-SCD-s), high SCD-stable (h-SCD-s), MCI-stable (MCI-s), low SCD-worsened (l-SCD-w), high

Table 2

Descriptive scores (means, standard deviations in parentheses, minimum and maximum) at baseline for socio-demographic variables (gender, age and years of schooling), health (Comorbidity Index, CBI), subjective complaints (QAM-participant), functional status (Lawton-Brody IADL) and cognitive status (Mini-Mental State Examination, MMSE), for the SCD patients (N = 199) classified in low and high SCD groups according to the stricter criterion (95th percentile) and the less strict criterion (5th percentile) for complaint severity

	95th criteria *Mean (SD), range*		5th criteria *Mean (SD), range*	
	low SCD	high SCD	low SCD	high SCD
N (*gender)	175 (67.42% women*)	24 (83.33% women)	120 (64.16% women*)	79 (78.48% women)
Age	64.59 (8.65), 50–84	66.37 (8.41), 51–87	64.31 (8.71), 50–84	65.57 (8.49), 50–87
Years of schooling	10.51 (4.76), 1–22	7.55 (4.00), 2–17	10.74 (4.96), 1–22	9.37 (4.29), 2–20
Comorbidity Index (CBI)	0.74 (0.85), 0–3	0.90 (0.86), 0–3	0.81 (0.90), 0–3	0.71 (0.78), 0–3
**QAM-participant	18.41 (4.08) **, 7–31	24.45 (3.37), 17–32	17.08 (3.71) **, 7–30	22.10 (3.81), 17–32
Lawton-Brody IADL	7.53 (0.85), 5–8	7.63 (0.84), 5–8	7.48 (0.89), 5–8	7.63 (0.78), 5–8
MMSE	28.17 (1.59), 21–30	28.04 (1.39), 25–30	28.14 (1.63), 21–30	28.18 (1.47), 25–30

Low SCD versus high SCD comparison: $^*p < 0.05$, $^{**}p < 0.001$.

SCD-worsened (h-SCD-w), and MCI-worsened (MCI-w). Two 6-group models were tested using the stricter (95th) and the less strict (5th) criterion (see Table 2).

Mann-Whitney U and Chi-square tests were performed using SPSS, version 21.0 (SPSS, Chicago, IL, USA) for descriptive analysis and group comparisons. Supervised ML techniques were used to predict the accuracy of classification of participants who progressed (i.e., worsened) or who remained stable (i.e., stable) across follow-ups considering the four groups in the 4-group (i.e., SCD-s, SCD-w, MCI-s, MCI-w) or considering the six groups in the 6-group (i.e., l-SCD-s, h-SCD-s, MCI-s, l-SCD-w, h-SCD-w, MCI-w) models. For each model, socio-demographic, health, functional, cognitive performance, cognitive reserve, affective and behavioral symptoms measured at baseline were used as predictors of participants' progression.

The data set consisted of 253 triplets from participants and 41 input variables: age, gender, years of schooling, professional category, Peabody vocabulary, WAIS vocabulary, CBI comorbidity, IADL-Lawton informant's score, CBI-burden, NPI-Q severity, NPI-Q stress, complaints from informant QAM, depression GDS-15, semantic fluency, letter fluency, ideational fluency, total score CAMCOG-R, logical memory I WMS-III, logical memory II WMS-III, paired associated I WMS-III, paired associated II WMS-III, immediate memory WMS-III, delayed memory WMS-III, primacy region CVLT, middle region CVLT, recency region CVLT, recall discriminability CVLT, long delay free recall (LDFR) CVLT, correct recognition CVLT, false positive recognition CVLT, intrusions free recall CVLT, intrusions cued recall CVLT, perseveration CVLT, semantic strategies short delayed free recall CVLT, semantic strategies long delayed free recall CVLT, serial strategies short delayed free recall CVLT, serial strategies long delayed free recall CVLT, working memory (WM) reading span, WM counting span, duration of stability, and total time.

ML analysis was performed using Python scripts and the Scikit-Learn library, with the Google Colaboratory tool [44]. For the training phase, a set of three ML classifiers (previously widely used in health research) was selected. In particular, we use the Random Forest Classifier and Extra Trees Classifier, based on trees that minimize a loss function, and the Support Vector Machines, which use hyperplanes that maximize the margin or distance from the line to the nearest points. We also applied a 10-fold stratified cross-validation in order to avoid biased or highly optimistic classification and to manage accurate imbalanced classes [45].

After the training phase, we computed widely used performance metrics, including score-based and graph metrics, to evaluate the adequacy of the ML prediction models. The true positives (TP: counts all instances of a group that are classified in that group), false positives (FP: counts all non-group instances that are classified in that group), true negatives (TN: counts all non-group instances that are not classified in that group), and false negatives (FN: counts all group instances that are not classified in that group) were computed. The following score-based metrics were calculated according to the correction in the classification for each of the groups considered in the 4-group model (i.e., SCD-s, SCD-w, MCI-s, MCI-w) and in the two 6-group models (i.e., l-SCD-s, h-SCD-s, MCI-s, l-SCD-w, h-SCD-w, MCI-w): 1) Recall or Sensitivity corresponding to the proportion of true positives [TP/(TP + FN)]; 2) Specificity corresponding to the proportion of true

negatives [TN/(TN + FP)]; 3) Precision or Positive Predictive Value (PPV) corresponding to the ratio of correctly predicted positive observations to the total predicted positive observations [TP/(FP + TP)]; 4) Accuracy corresponding to the number of correct predictions divided by the total number of predictions [(TP + TN)/(TP + TN + FP + FN)]; 5) F1-score corresponding to the harmonic mean of Precision and Recall [2*(Recall*Precision)/(Recall + Precision)]; 6) Error-rate corresponding to the proportion of false positives and false negatives divided by the total number of predictions [(1-(TP + TN/TP + TN + FP + FN))*100]; and 7) Negative Predictive Value (NPV) corresponding to the ratio of correctly predicted negative observations to the total predicted negative observations [TN/(TN + FN)].

The decision about the applicability of the 4-group or 6-group models was made by comparing the score-based metrics and their respective confusion matrices. The most valid, informative, and parsimonious model for classifying stable and worsening SCD and MCI participants was selected for further predictive analysis.

The significance of the inter-group comparisons between stable and worsening l-SCD, h-SCD, and MCI groups for the predictive measures at baseline was established, and the p value was revised to avoid the type I error for multiple comparisons [46].

Predictive loadings associated with each measure were calculated for the global predictive model selected, and importance was established according to the predictive weights.

RESULTS

Descriptive scores at baseline for socio-demographic, health, subjective complaints, cognitive status, and functional measures for the SCD groups with low and high subjective complaints, and considering more and less strict criteria for determining complaint severity, are shown in Table 2.

We observed significant group differences in gender and QAM scores, but not in age, education, comorbidity, functionality, or cognitive status when comparing the strict and less strict criteria for the importance of complaints.

Total scores of the complaints questionnaire did not show normal distribution [Shapiro–Wilk's test: $W(253) = 0.973$, $p < 0.001$] and therefore the reference scores for the extreme reference values in the distribution (5th and 95th) were calculated considering percentiles. The distribution was positively skewed (*Skewness* = 0.537; *Standard error* = 0.153) and adopted a slight leptokurtic shape (*Kurtosis* = 0.372; *Standard error* = 0.305).

Most of the SCD participants (79.4%) remained stable across the follow-ups, and only 20.58% progressed either to MCI (11.76%) or dementia (8.82%). The proportion of MCI participants who worsened reached 68.62%, with 60.78% converting to dementia and 7.84% progressing to multiple-domain MCI.

Predictive validity ignoring or considering the severity of complains

Sensitivity (recall), Specificity, Precision, Accuracy, F1-score, Error-rate and Negative Predictive Value (NPV) were calculated ignoring the severity of complains (for the 4-group models) (see Table 3) and taking the severity into account (6-group models) (see Table 4). Random forest algorithm yielded the best results for predictive validity, producing the most accurate classification, considering the predicted progression.

Ignoring (4-groups model) or taking into account (6-groups models) the severity of complains did not showed differences in classifying stable SCD groups with lower severity regarding Sensitivity, Precision, and Accuracy. However, taking into account severity (6-group models) improved the global accuracy in the worsened SCD group, particularly identification of true negatives (i.e., correct classification of participants who do not belong to the low or high SCD-w group, in relation to both the total number

Table 3
Predictive validity values for the 4-group model: SCD-stable, MCI-stable, SCD-worsening, MCI-worsening

4-groups model	Sensitivity (Recall)	Specificity	Precision	Accuracy	F1-score	Error-rate	NPV
SCD-stable	0.98	0.52	0.93	0.92	0.92	7.17	0.86
MCI-stable	0.20	0.93	0.75	0.56	0.56	43.33	0.53
SCD-worsening	0.82	0.63	0.75	0.74	0.74	25.49	0.73
MCI-worsening	0.79	0.68	0.83	0.75	0.75	24.13	0.61

The validity metrics (Random Forest) correspond to the accuracy of the classification of the participants in each of the groups.

Table 4

Predictive validity values for the 6-group models: low-SCD-stable (l-SCD-s), high-SCD-stable (h-SCD-s), MCI-stable (MCI-s), low-SCD-worsening (l-SCD-w), high-SCD-worsening (h-SCD-w), and MCI-worsening (MCI-w) using the less strict criterion (5th percentile) and stricter criterion (95th percentile) for complaint severity

6-groups models (5th, 95th)	Sensitivity (Recall)	Specificity	Precision	Accuracy	F1-score	Error-rate	NPV
l-SCD-s (5th)	0.91	0.27	0.64	0.65	0.75	34.80	0.68
l-SCD-s (95th)	0.98	0.37	0.83	0.83	0.90	16.50	0.85
h-SCD-s (5th)	0.19	0.65	0.52	0.34	0.27	65.26	0.29
h-SCD-s (95th)	0.00	1.00	–	0.64	–	35.29	0.64
MCI-s	0.60	0.98	0.90	0.92	0.72	7.60	0.92
l-SCD-w (5th)	0.53	0.97	0.77	0.91	0.63	8.88	0.92
l-SCD-w (95th)	0.69	0.72	0.57	0.82	0.57	28.35	0.82
h-SCD-w (5th)	0.56	0.93	0.64	0.86	0.60	13.04	0.91
h-SCD-w (95th)	0.00	1.0	–	0.88	–	11.76	0.88
MCI-w	0.89	0.87	0.77	0.88	0.83	11.86	0.94

The validity metrics (Random Forest) correspond to the accuracy of the classification of the participants in each of the groups.

of true negatives, the specificity and the total of predicted negatives, NPV). The 6-group model was also globally better for simultaneously classifying the stable and worsening MCI participants. In summary, specificity, accuracy, and NPV were more favorable in the two 6-group models (considering both the more and less strict criteria for complaint importance) for predicting SCD-w, MCI-s, and MCI-w, regardless of the increased probability of mistaking the classification associated with the greater number of groups considered.

Strict or less strict criteria for the valid identification of participants' progression

Regarding application of the less strict or strict criteria for identifying high versus low SCD participants, the stricter criterion (95th percentile) only yielded better metrics for classifying stable SCD groups. The metrics were similar or more favorable for the less strict criterion (5th percentile) in predicting progression in the SCD-w, MCI-s, and MCI-w groups. Closer analysis using the confusion matrix (see Fig. 1) showed the following: a) both criteria, particularly the stricter one, failed to correctly classify the h-SCD-s participants who were mainly considered l-SCD-s; b) the less strict criterion was better in successfully classifying the l-SCD-w and h-SCD-w participants; and c) the less strict criterion was better in the valid prediction of the stable and worsening MCI participants.

When the less strict criterion of complaint severity was applied, the progression rate was 3.92% (1.96% to dementia) in the low SCD group, and 16.66% (6.86% to dementia) in the high SCD group. By contrast, when the stricter criterion was applied, the

progression rate was 14.70% (6.86% to dementia) in the low SCD group and 5.88% (1.96% to dementia) in the high SCD group. Our results showed that the highest progression rates were observed in the groups that included not only those who complained the most but also those who complained moderately, whatever the rigor of the criterion used (i.e., in the high SCD group for the 5th percentile and in the low SCD group for the 95th percentile).

Descriptive values for the predictive measures for the two 6-group models with the less strict criterion for complaint severity were calculated, and pairwise differences between groups were estimated (see Table 5). To deal with familywise error rates for multiple hypothesis tests, the significance p level at 0.05 was corrected using the Shaffer & Holm procedure suggested by Holland & Copenhaver [46] for improvement of the Holm-Bonferroni method ($p_{Holm-Bonferroni} < 0.00008$).

Group differences in predictive variables at baseline considering the less strict criterion (5th) of the 6-group model

Differences in predictive variables at baseline were mainly found when comparing the worsening MCI group and the stable SCD groups (i.e., both low SCD and high SCD). Thus, comparison of the worsening MCI group and the stable SCD groups showed that the worsening MCI participants were significantly older, remained stable for less time, and obtained lower scores in cognitive variables (semantic and ideational fluency tests, CAMCOG-R test, vocabulary WAIS subtest, episodic memory measures from the WMS, i.e., logical memory I and II, paired associates I and II, immediate and delayed memory, and the CVLT,
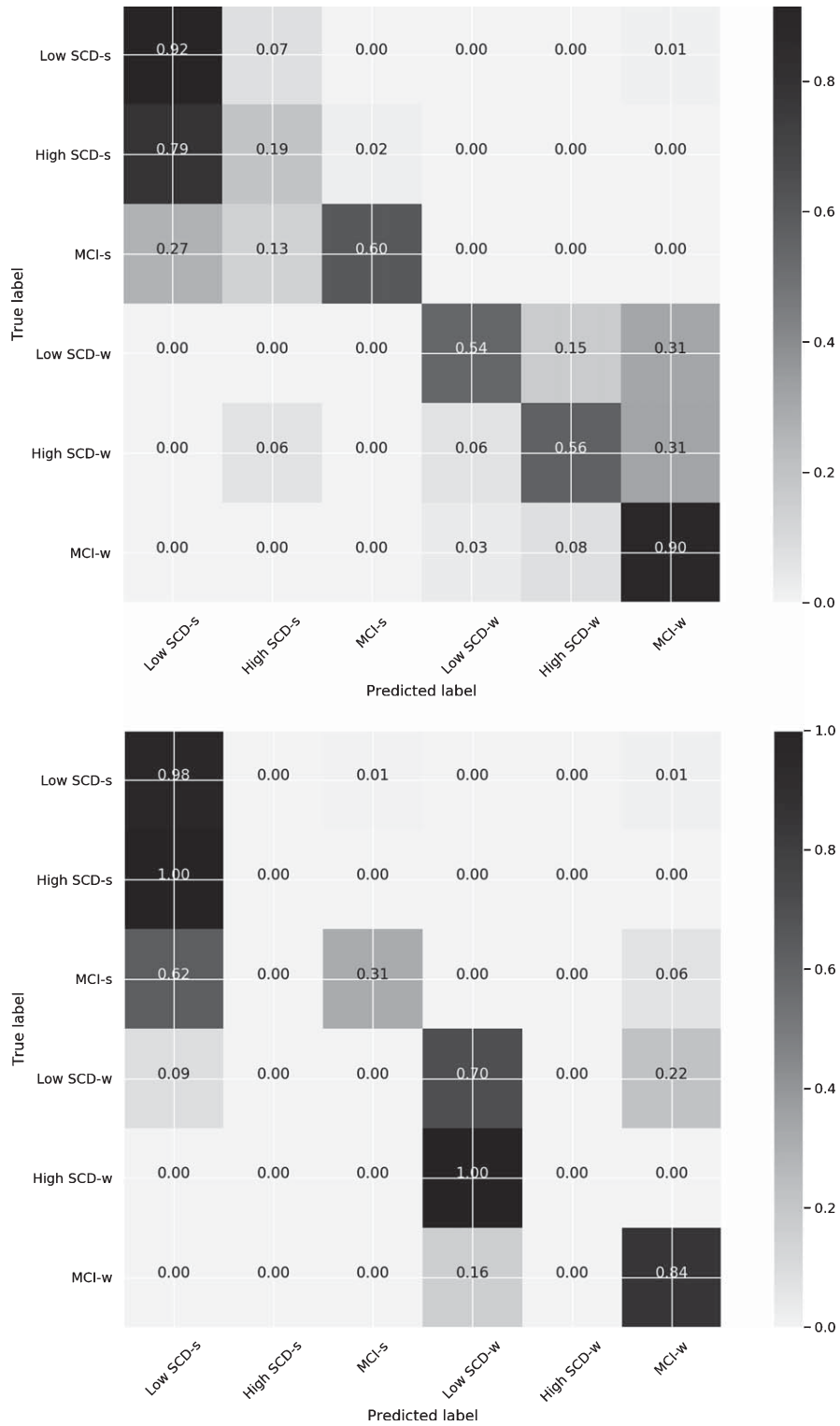
Fig. 1. Confusion matrices for the 6-group classification model with the less strict (5th percentile) criterion for complaint severity (above) and the corresponding stricter criterion (95th percentile) (below).

Table 5
Descriptive values (means, *M*, and standard deviations, *SD*) for the predictive variables and corrected Holm-Bonferroni estimation of the inter-group differences at baseline for the 6-group model (5th percentile)

	l-SCD-s (*M, SD*)	h-SCD-s (*M, SD*)	MCI-s (*M, SD*)	l-SCD-w (*M, SD*)	h-SCD-w (*M, SD*)	MCI-w (*M, SD*)
Age* (5)(9)	63.58 (8.53)	64.47 (8.86)	68.60 (6.38)	69.69 (8.42)	70.56 (4.61)	74.25 (7.39)
Gender	63.35% women	77.77% women	66.66% women	46.15% women	81.26% women	58.97 women
Years schooling	10.82 (5.05)	10.12 (4.28)	10.26 (3.78)	9.69 (4.17)	6.37 (2.62)	9.53 (4.52)
Professional category	3.13 (1.14)	3.25 (1.06)	2.86 (1.06)	3.15 (1.06)	2.75 (0.85)	2.94 (1.09)
Vocabulary-Peabody	66.13 (15.61)	67.26 (16.25)	54.60 (13.27)	58.50 (15.62)	52.18 (18.96)	56.57 (18.43)
Vocabulary-WAIS* (5)(9)	52.53 (11.85)	52.53 (13.00)	41.60 (12.21)	44.53 (14.97)	42.06 (10.27)	44.13 (13.92)
Comorbidity-CCI	0.79 (0.93)	0.82 (0.90)	1.26 (1.09)	0.69 (0.75)	0.68 (0.70)	0.74 (0.81)
IADL-Lawton-informant* (9)	7.48 (0.87)	7.66 (0.80)	7.00 (1.24)	7.41 (1.08)	7.56 (0.72)	6.35 (1.74)
Burden-CBI	6.87 (10.86)	6.67 (7.85)	7.00 (9.05)	8.41 (8.17)	13.62 (16.84)	13.25 (12.72)
Severity-NPI-Q	2.58 (3.68)	2.20 (2.38)	3.80 (2.69)	2.58 (3.02)	5.31 (6.61)	5.18 (5.07)
Estress-NPI-Q	2.69 (4.74)	2.26 (3.52)	2.80 (2.14)	2.66 (3.11)	7.06 (11.68)	6.50 (8.52)
Complaints informant-QAM	14.89 (4.29)	17.15 (5.06)	16.90 (3.28)	15.66 (3.44)	18.18 (4.08)	18.18 (4.20)
Depression-GDS	3.00 (2.51)	4.09 (3.13)	3.53 (2.97)	3.07 (2.32)	4.31 (3.13)	3.48 (3.00)
Semantic fluency* (5)(9)	18.40 (5.30)	17.00 (5.36)	13.66 (5.48)	14.23 (4.47)	14.25 (3.25)	12.10 (4.14)
Letter fluency	12.53 (4.99)	12.54 (4.75)	8.74 (2.12)	13.40 (3.91)	7.28 (3.98)	8.38 (6.07)
Ideational fluency* (5)(9)	5.59 (1.54)	5.23 (1.69)	4.46 (1.88)	4.61 (1.75)	4.18 (2.13)	3.51 (1.31)
CAMCOG-R total* (5)(9)	89.29 (7.08)	89.92 (6.09)	77.86 (11.70)	81.00 (7.05)	83.00 (7.50)	72.87 (10.48)
Logical mem I-WMS* (4)(5)(9)	24.32 (7.26)	21.88 (6.37)	16.73 (5.75)	17.00 (3.51)	15.18 (6.17)	11.66 (5.64)
Logical mem II-WMS* (2)(5)(9)	20.46 (7.63)	18.43 (6.72)	10.42 (6.30)	13.36 (4.31)	13.00 (7.87)	6.63 (5.42)
Paired assoc I-WMS* (5)(9)	14.72 (7.00)	16.49 (7.42)	8.80 (6.18)	10.30 (4.88)	11.87 (5.63)	5.72 (5.58)
Paired assoc II-WMS* (5)(9)	4.43 (2.22)	5.07 (2.36)	2.35 (1.73)	3.27 (1.79)	4.06 (2.12)	1.36 (1.79)
Immediate mem-WMS* (5)(9)	98.83 (12.27)	99.13 (12.74)	87.00 (9.41)	90.69 (8.08)	93.68 (11.33)	81.74 (14.29)
Delayed mem-WMS* (5)(9)	111.45 (14.64)	111.54 (16.25)	97.14 (10.95)	102.36 (9.98)	108.40 (16.72)	90.43 (13.62)
Primacy_Reg_CVLT	39.17 (6.04)	38.64 (5.96)	39.15 (10.06)	41.48 (9.32)	40.41 (9.67)	34.80 (10.66)
Middle_reg_CVLT* (5)(9)	28.21 (6.61)	28.93 (6.17)	21.04 (8.45)	28.10 (5.62)	28.18 (8.24)	16.24 (8.86)
Recency_reg_CVLT* (5)(9)	32.60 (7.36)	32.41 (6.28)	39.80 (9.15)	30.40 (10.32)	31.39 (9.79)	48.94 (13.38)
Recall discrim-CVLT* (2)(5)(6)(9)	94.36 (5.48)	93.93 (5.61)	83.60 (9.24)	84.46 (16.26)	89.56 (7.13)	76.03 (13.71)
Long Delay Free Recall-CVLT* (2)(5)(9)(14)(15)	11.46 (2.88)	11.41 (2.69)	5.66 (4.04)	8.15 (3.38)	10.06 (2.59)	3.02 (2.88)
Correct recogn-CVLT* (5)(9)	15.01 (1.41)	15.14 (1.13)	13.06 (2.73)	14.23 (2.48)	14.50 (1.67)	12.97 (2.79)
False posit recogn-CVLT* (2)(5)(9)	1.48 (1.89)	1.85 (1.96)	4.13 (2.35)	4.92 (6.88)	3.99 (2.87)	7.87 (5.81)
Intrusions Free recall CVLT	3.98 (4.23)	3.30 (3.66)	6.33 (5.36)	6.00 (4.02)	4.68 (3.70)	7.15 (8.85)
Intrusions Cued Recall CVLT *(5)(6)(9)	2.54 (2.87)	1.82 (2.26)	6.33 (3.35)	5.00 (3.51)	4.25 (4.02)	8.33 (9.17)
Perseverations-CVLT	7.08 (5.76)	6.93 (6.75)	4.66 (5.15)	4.46 (6.05)	5.75 (3.29)	1.82 (2.06)
Semantic strateg SDFR CVLT* (2)(5)(9)(14)(15)	3.53 (2.60)	3.46 (2.53)	0.93 (1.27)	1.92 (1.55)	3.06 (1.76)	0.30 (0.65)
Semantic strateg LDFR CVLT* (5)(9)(15)	4.61 (3.07)	4.55 (2.86)	2.00 (2.7)	2.07 (1.70)	3.12 (2.21)	0.71 (1.27)
Serial strateg SDFR CVLT *(5)(9)	0.72 (1.01)	0.76 (0.97)	0.13 (0.35)	0.61 (1.19)	1.00 (2.09)	0.17 (0.55)
Serial strateg LDFR CVLT	0.68 (1.02)	0.76 (1.17)	0.46 (0.51)	0.76 (1.09)	1.25 (1.34)	0.28 (0.72)
Reading span-WM *(5)(9)	70.49 (12.18)	68.41 (11.31)	51.00 (23.65)	56.19 (17.31)	57.80 (13.33)	44.00 (19.49)
Counting span-WM *(5)(9)	81.71 (8.76)	80.74 (10.24)	72.73 (10.34)	76.30 (8.76)	79.56 (5.63)	64.50 (15.37)
Time stable* (3)(4)(5)(7)(8)(9)(10)(11)(12)	48.68 (12.65)	49.42 (13.25)	52.26 (14.77)	19.92 (8.91)	23.62 (13.55)	14.07 (17.19)
Time total	48.68 (12.65)	49.42 (13.25)	52.26 (14.77)	43.53 (14.71)	54.25 (13.61)	37.61 (16.00)

*significance of the group comparison with Holm-Bonferroni correction ($p < 0.00008$) at baseline measurements; (1) l-SCD-s/h-SCD-s; (2) l-SCD-s/MCI-s; (3) l-SCD-s/l-SCD-w; (4) l-SCD-s/-h-SCD-w; (5) l-SCD-s/MCI-w; (6) h-SCD-s/MCI-s; (7) h-SCD-s/l-SCD-w; (8) h-SCD-s/h-SCD-w; (9) h-SCD-s/MCI-w; (10) MCI-s/l-SCD-w; (11) MCI-s/-h-SCD-w; (12) MCI-s/MCI-w; (13) l-SCD-w/h-SCD-w; (14) l-SCD-w/MCI-w; (15) h-SCD-w/MCI-w.

i.e., recall discriminability, long delay free recall, recognition accuracy, false positives in recognition, intrusions in cued recall, total perseverations, semantic strategies in long and short delay free recall, recall in the recency and middle region tests, and reading and counting span). In addition, the worsening MCI group obtained lower scores in IADL than the high SCD-stable group.

Some differences at baseline were also observed when comparing the two types of stable SCD (i.e., low and high SCD) groups and the stable MCI groups. Thus, scores were higher for low SCD-stable group than for the stable MCI group in several measures of episodic memory (i.e., Logical memory II of the WMS, and long delay free recall, recall discriminability, false positives in recognition, and semantic strategies in short delay free recall of the CVLT). Significant differences between high SCD-stable and stable MCI groups were observed for the measures of episodic memory recall of the CVLT (i.e., intrusions in cued recall and recall discriminability).

Significant differences were also observed when the two worsening SCD (low and high SCD) were compared with the worsening MCI group. The worsening SCD participants scored better in long delay free recall and used more semantic strategies in short delay free recall of the CVLT than the worsening MCI group. High SCD-w participants also used more semantic strategies in long delay free recall of the CVLT than the worsening MCI group.

The only significant difference between stable and worsening SCD groups was found between high SCD-worsened and low SCD-stable groups, only for the logical memory I subtest of the WMS.

Significant differences were found in the length of time participants remained stable, in all pairwise comparisons between stable and worsening, regardless of group (SCD or MCI). No significant differences between the SCD and MCI groups were observed either when both were stable or when they worsened.

Factors importance in predicting progression considering the less strict criterion (5th) of the 6-group model

The importance of the factors used in predicting progression of conditions with the less strict criterion (5th percentile) for this 6-group model was established by computing and ranking the standardized variable loadings of the algorithm that provided the best prediction metrics (i.e., Random Forest). The sum of the predictive loads associated with the 10 variables selected (i.e., 24.4% of the total data set) for the stable and worsening groups accounted for around 60% of the total importance of the predictive algorithm.

The algorithm showed that the following measures, ordered by loadings, mainly predicted progression: recall of words from the primacy portion (primacy reg CVLT: $w = -0.0434$), paired associates recall (paired assoc. I WMS: $w = 0.0176$; paired assoc. II WMS: $w = 0.0222$), professional category (Professional category: $w = 0.0208$), propensity to use of semantic strategies in the long delay free recall (semantic strategy LDFR CVLT: $w = 0.0170$), comorbidity (Comorbidity CCI: $w = 0.0164$), recognition accuracy (correct recogn. CVLT: $w = 0.0151$), education (Years of schooling: $w = 0.0131$), working memory counting span (counting WM: $w = 0.0127$), and vocabulary (Vocabulary WAIS: $w = 0.0113$).

DISCUSSION

Our main objectives were to analyze the validity of considering SCD subgroups according to SCCs' severity and to compare two extreme thresholds (5th and 95th percentiles) when applying ML algorithm in predicting progression of SCD and MCI participants.

Regarding progression, our findings indicated that most of the SCD participants (79.4%) remained stable across the follow-ups [11, 30] as only 20.58% progressed either to MCI (11.76%) or dementia (8.82%). By contrast, worsening of MCI participants reached 68.62%, with 60.78% converting to dementia and 7.84% progressing to multiple-domain MCI after 54–72 months of follow-up.

In line with the increased risk of progression and conversion estimated for SCD participants relative to non-complainer older adults [8, 9], we observed considerably higher rates of progression and conversion in participants with moderate to high complaints (above the 5th percentile) than in those with low levels of complaints (below the 5th percentile) as also observed in normative aging [11]. Similarly, even using the stricter criterion (95th), higher rates of progression and conversion were observed in the group in which participants with moderate to high complaints were classified together (i.e., low SCD).

Regarding quantification of SCC severity, our findings showed that joint consideration of the SCD participants classified according to stability, but without distinction by severity (i.e., the 4-group model), provided better predictive validity than the 6-group model for classifying stable SCD participants. However, when SCD participants were classified according to complaint severity, differentiating low and high SCCs (i.e., with the any of 6-group models), the predictive validity metrics improved overall for the worsening SCD and stable and worsening MCI participants. These results highlight the importance

of considering the severity of SCCs in predicting progression along the continuum of the pre-symptomatic and prodromal stages of dementia. This is particularly important for SCD participants who progress to MCI or dementia, because it enables more precise identification of participants who do not really belong to the l-SCD-w and h-SCD-w groups [6, 32]. Our findings give partial empirical support to the need to evaluate SCC severity by rating the frequency of perceived difficulties through items that assess complaints on a scale ranging from low to high [31, 32].

Comparison of two extreme thresholds for quantifying complaint severity indicated that the less strict criterion was better with almost all the metrics of predictive validity for predicting progression in participants with SCCs. Thus, even though both criteria failed to correctly classify the stable h-SCD participants (i.e., high SCD-stable were mostly classified as low SCD-stable), the less strict criterion was more successful in classifying the worsening SCD participants, and especially in accurately classifying stable and worsening MCI participants. Therefore, the predictive validity in our sample of participants with SCCs improved when the cognitively unimpaired participants declaring low frequency of cognitive difficulties (i.e., equal or below the 5th percentile) were distinguished from the moderate to high complainers (i.e., above the 5th percentile).

Concerning the baseline intergroup differences, our findings showed significant differences mainly between the worsening MCI and the stable SCD groups (i.e., l-SCD-s, h-SCD-s). Participants classified with worsening MCI were significantly older, progressed faster, and obtained lower scores at baseline in the vast majority of the cognitive measures (i.e., recall and recognition episodic memory, cognitive status, vocabulary, and working memory and executive measures) than the SCD-stable (both low and high) groups. The highly stable SCD participants even showed better functional status and use of semantic strategies than the MCI-worsened group at baseline. However, Stable SCD (i.e., l-SCD-s, h-SCD-s) and MCI-worsened groups did not differ significantly in gender, comorbidity, behavioral disturbance, depression, informant complaints, burden, cognitive reserve proxies (i.e., schooling, professional achievement) or total time in the study.

In contrast to the results reported by Bessi et al. [10] who did not implement corrections to avoid type I errors, we did not find any significant differences when comparing the stable and worsened MCI groups or the stable and worsened SCD groups at baseline.

Thus, the only significant difference between stable and worsening SCD groups were found in a measure of episodic memory recall (i.e., logical memory I subtest), indicating lower scores for high SCD-worsening participants than for low SCD-stable participants and enhancing the importance of episodic memory measures in discriminating SCD participants who progress along the continuum of cognitive decline [18, 47].

As expected, the stability of diagnosis over time indicated significantly faster progression in worsening than stable groups, both in MCI and SCD participants. Although previous studies reported faster progression rates in MCI participants with enhanced risk markers [5, 48] and SCD [11, 49, 50], we cannot reach relevant conclusions about this aspect, as we did not analyze time differences according to risk factors or severity of progression among those participants who worsened.

The variables with a major role in conforming the algorithm for predicting stable and worsening SCD and MCI participants were those related to working memory/executive functions (counting span, semantic strategies), episodic memory recall (primacy recall, paired associates' recall) and recognition, proxies for cognitive reserve (schooling, professional attainment, vocabulary), health (comorbidity) and duration of stability of diagnosis. Overall, our results support the important role of working memory and executive functions [17, 19, 23, 28, 51] and episodic memory [18, 20, 22, 24, 26, 27, 47] measures, as well as the modulating role of protective and risk factors such as cognitive reserve and health status in jointly predicting progression of SCD and MCI participants.

Our findings highlight the important contribution of the multimodal ML approach in predicting progression to Alzheimer's disease [29]. Thus, several of the most important measures in predictive algorithms (e.g., primacy recall, comorbidity, schooling, professional category, depressive symptomatology) were not significant in the comparison of stable and worsening groups at baseline after Holm-Bonferroni correction, even in the comparison of groups at the opposite extremes of the cognitive decline continuum (i.e., stable SCD groups and worsening MCI group). Moreover, calculation of factor importance in predicting progression in SCD and MCI, and the consequent simplification to the most relevant factors enables improvements in future protocols by focusing on these variables.

Further studies, preferably with a larger sample size, should be carried out to minimize the risk of

overfitting and to establish specific predictors for each stable and worsening SCD and MCI groups. The validity of the classification of SCD participants according to the severity of the complaints, indicated by more or less strict criteria, should be analyzed in community samples. Additional studies should be carried out to specifically search for significant predictive associations in each stable and worsening SCD and MCI groups.

ACKNOWLEDGMENTS

REFERENCES

[1] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.

[2] Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, Gronseth GS, Marson D, Pringsheim T, Day GS, Sager M, Steven J, Rae-Grant A (2018) Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* **90**, 126-135.

[3] Brambati SM, Belleville S, Kergoat MJ, Chayer C, Gauthier S, Joubert, S (2009) Single-and multiple-domain amnestic mild cognitive impairment: Two sides of the same coin. *Dement Geriatr Cogn Disord* **28**, 541-549.

[4] Facal D, Guárdia-Olmos J, Pereiro, AX, Lojo-Seoane C, Peró, M, Juncos-Rabadán O (2019) Using an overlapping time interval strategy to study diagnostic instability in mild cognitive impairment subtypes. *Brain Sci* **9**, 242.

[5] Forlenza OV, Diniz BS, Nunes PV, Memória CM, Yassuda MS, Gattaz, WF (2009) Diagnostic transitions in mild cognitive impairment subtypes. *Int Psychogeriatr* **21**, 1088-1095.

[6] Jessen F, Wiese B, Bachmann C, Eifflaender-Gorfer S, Haller F, Kölsch H, Luck T, Mösch E, van den Bussche H, Wagner M, Wollny A, Zimmermann T, Pentzek M, Riedel-Heller SG, Romberg HP, Weyerer S, Kaduszkiewicz H, Maier W, Bickel H, German Study on Aging, Cognition and Dementia in Primary Care Patients Study Group (2010) Prediction of dementia by subjective memory impairment: Effects of severity and temporal association with cognitive impairment. *Arch Gen Psychiatry* **67**, 414-422.

[7] Jessen F, Amariglio RE, Van Boxtel M, Breteler M, Ceccaldi M, Chételat G, Dubrois B, Dufouil C, Ellis KA, van der Flier WM, Glodzik L, van Harten AC, de Leon MJ, McHugh P, Mielke MM, Molinuevo JL, Mosconi L, Osorio RS, Perrotin A, Petersen RC, Rabin LA, Rami L, Reisber B, Rentz DM, Sachdev PS, de la Sayette V, Saykin AJ, Scheltens P, Shulman MB, Slavin MJ, Sperling RA, Stewart R, Uspenskaya O, Vellas B, Visser PJ, Wagner M, Subjective Cognitive Decline Initiative (SCD-I) Working Group (2014) A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement* **10**, 844-852.

[8] Abdulrab K, Heun R (2008) Subjective Memory Impairment. A review of its definitions indicates the need for a comprehensive set of standardised and validated criteria. *Eur Psychiatry* **23**, 321-330.

[9] Snitz BE, Wang T, Cloonan YK, Jacobsen E, Chang CCH, Hughes TF, Kamboh MI, Ganguli M (2018) Risk of progression from subjective cognitive decline to mild cognitive impairment: The role of study setting. *Alzheimers Dement* **14**, 734-742.

[10] Bessi V, Mazzeo S, Padiglioni S, Piccini C, Nacmias B, Sorbi, S, Bracco L (2018) From subjective cognitive decline to Alzheimer's disease: The predictive role of neuropsychological assessment, personality traits, and cognitive reserve. A 7-year follow-up study. *J Alzheimers Dis* **63**, 1523-1535.

[11] Mendonça MD, Alves L, Bugalho P (2016) From subjective cognitive complaints to dementia: Who is at risk? A systematic review. *Am J Alzheimers Dis Other Demen* **31**, 105-114.

[12] Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B (2014) Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatr Scand* **130**, 439-451.

[13] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman D, Jagust W, Jessen F, Karlawish J, Liu, E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemens E, Snyder HM, Sperling R (2018) NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562.

[14] Brigola AG, Manzini CSS, Oliveira GBS, Ottaviani AC, Sako MP, Vale, FAC (2015) Subjective memory complaints associated with depression and cognitive impairment in the elderly: A systematic review. *Dement Neuropsychol* **9**, 51-57.

[15] Cheng YW, Chen TF, Chiu MJ (2017) From mild cognitive impairment to subjective cognitive decline: Conceptual and methodological evolution. *Neuropsychiatr Dis Treat* **16**, 491-498.

[16] Fritsch T, McClendon MJ, Wallendal MS, Hyde TF, Larsen JD (2014) Prevalence and cognitive bases of subjective memory complaints in older adults: Evidence from a community sample. *J Neurodegener Dis* **2014**, 176843.

[17] Kurt P, Yener G, Oguz M (2011) Impaired digit span can predict further cognitive decline in older people with subjective memory complaint: A preliminary result. *Aging Mental Health* **15**, 364-369.

[18] Studart-Neto A, Nitrini R (2016) Subjective cognitive decline: The first clinical manifestation of Alzheimer's disease? *Dement Neuropsychol* **10**, 170-177.

[19] Artero S, Tierney MC, Touchon J, Ritchie K (2003) Prediction of transition from cognitive impairment to senile

dementia: A prospective, longitudinal study. *Acta Psychiatr Scand* **107**, 390-393.

[20] Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J (2017) Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: A systematic review and meta-analysis. *Neuropsychol Rev* **27**, 328-353.

[21] Belleville S, Gauthier S, Lepage E, Kergoat MJ, Gilbert B (2014) Predicting decline in mild cognitive impairment: A prospective cognitive study. *Neuropsychology* **28**, 643-652.

[22] Egli SC, Beck IR, Berres M, Foldi NS, Monsch AU, Sollberger M (2014) Serial position effects are sensitive predictors of conversion from MCI to Alzheimer's disease dementia. *Alzheimers Dement* **10**, S420-S424.

[23] Emrani S, Libon DJ, Lamar M, Price CC, Jefferson AL, Gifford KA, Hohman TJ, Nation DA, Delano-Wood L, Jak A, Bangen KJ, Bondi MW, Brickman AM, Manly J, Swenson R, Au R (2018) Assessing working memory in mild cognitive impairment with serial order recall. *J Alzheimers Dis* **61**, 917-928.

[24] Gainotti G, Quaranta D, Vita MG, Marra C (2014) Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease. *J Alzheimers Dis* **38**, 481-495.

[25] Lee SJ, Ritchie CS, Yaffe K, Cenzer IS, Barnes DE (2014) A clinical index to predict progression from mild cognitive impairment to dementia due to Alzheimer's disease. *PloS One* **9**, e113535.

[26] Talamonti D, Koscik R, Johnson S, Bruno, D (2020) Predicting early mild cognitive impairment with free recall: The primacy of primacy. *Arch Clin Neuropsychol* **35**, 133-142.

[27] Tabert MH, Manly JJ, Liu X, Pelton GH, Rosenblum S, Jacobs M, Zamora D, Goodkind M, Bell K, Stern Y, Devanand DP (2006) Neuropsychological prediction of conversion to alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry* **63**, 916-924.

[28] Tierney MC, Yao C, Kiss A, McDowell I (2005) Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology* **64**, 1853-1859.

[29] Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M, Lin CT (2020) Machine learning techniques for the diagnosis of Alzheimer's disease: A review. *ACM Trans Multimed Comput Commun Appl* **16**, 30.

[30] Hessen E, Eckerström M, Nordlund A, Almdahl IS, Stålhammar J, Bjerke M, Eckerström C, Göthlin M, Fladby T, Reinvang I, Wallin A (2017) Subjective cognitive impairment is a predominantly benign condition in memory clinic patients followed for 6 years: The Gothenburg-Oslo MCI Study. *Dement Geriatr Cogn Dis Extra* **7**, 1-14.

[31] Ávila-Villanueva M, Fernández-Blázquez MA (2017) Subjective cognitive decline as a preclinical marker for Alzheimer's disease: The challenge of stability over time. *Front Aging Neurosci* **9**, 377.

[32] Rabin LA, Smart CM, Crane PK, Amariglio RE, Berman LM, Boada M, Buckley R, Chételat G, Dubois B, Ellis KA, Gifford, KA, Jefferson AL, Jessen F, Katz MJ, Lipton RB, Luck T, Maruff P, Mielke MM, Molinuevo JL, Naemm F, Perrotin A, Petersen RC, Rami L, Reisberg B, Rentz DM, Riedel-Heller SG, Risacher SL, Rodriguez O, Sachdev PS, Saykin AJ, Slavin MJ, Snitz BE, Sperling RA, Tandetnik C, van del Flier WM, Wagner M, Wolfsgruber S, Sikkes SAM (2015) Subjective cognitive decline in older adults: An overview of self-report measures used across 19 international research studies. *J Alzheimers Dis* **48**, S63-S86.

[33] Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, Nordberg A, Bäckman L, Albert M, Almkvist O, Arai H, Basun H, Blennow K, de Leon M, DeCarli C, Erkinjuntti T, Giacobini E, Graff C, Hardy J, Jack C, Jorm A, Ritchie K, van Duijn C, Visser P, Petersen RC (2004) Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med* **256**, 240-246.

[34] Juncos-Rabadan O, Pereiro AX, Facal D, Rodriguez N, Lojo C, Caamaño JA, Sueiro J, Boveda J, Eiroa P (2012) Prevalence and correlates of cognitive impairment in adults with subjective memory complaints in primary care centres. *Dement Geriatr Cogn Disord* **33**, 226-232.

[35] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 270-279.

[36] Campos-Magdaleno M, Facal D, Juncos-Rabadán O, Picón E, Pereiro AX (2016) Comparison between an empirically derived and a standard classification of amnestic mild cognitive impairment from a sample of adults with subjective cognitive complaints. *J Aging Health* **28**, 1105-1115.

[37] Garibotto V, Borroni B, Kalbe E, Herholz K, Salmon E, Holtoff V, Sorbi SF, Cappa F, Padovani A, Fazio F, Perani, D (2008) Education and occupation as proxies for reserve in aMCI converters and AD: FDG-PET evidence. *Neurology* **71**, 1342-1349.

[38] Van der Linden M, Wijns C, Von Frenkell R, Coyette F, Seron X (1989) *Un questionnaire d'auto-évaluation de la mémoire (QAM)*. Editest, Belgium.

[39] Benedet MJ, Seisdedos N (1996) *Evaluación Clínica de las Quejas de la Memoria en la Vida Cotidiana*, Editorial Médica Panamericana, Madrid.

[40] Benedet MJ, Alejandre MA (1998) *Test de aprendizaje verbal España-Complutense (TAVEC)*, TEA, Madrid.

[41] Delis DC, Kramer JH, Kaplan E, Ober BA (1987) *California Verbal Learning Test*. Psychological Corporation, San Antonio, TX.

[42] Roth M, Huppert FA, Mountjoy CQ, Tym E, López-Pousa S (2003) *CAMDEX-R: Prueba de Exploración Cambridge revisada para la valoración de los trastornos mentales en la vejez. Adaptación española*. TEA Ediciones, Madrid.

[43] Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N (2006) Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* **59**, 7-10.

[44] Bisong E (2019) Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Bisong E, ed Apress, Berkeley, CA. pp. 59-64.

[45] Schaffer C (1993) Overfitting avoidance as bias. *Mach Learn* **10**, 153-178.

[46] Holland BS, Copenhaver MD (1987) An improved sequentially rejective Bonferroni test procedure. *Biometrics* **43**, 417-423.

[47] Campos-Magdaleno M, Facal D, Juncos-Rabadán O, Braña T, Pereiro AX (2014) Cluster subtypes of the Spanish version of the California Verbal Learning Test in a sample of adults with subjective memory complaints. *J Clin Exp Neuropsychol* **36**, 317-333.

[48] Delor I, Charoin JE, Gieschke R, Retout S, Jacqmin P, Alzheimer's Disease Neuroimaging Initiative (2013)

Modeling Alzheimer's disease progression using disease onset time and disease trajectory concepts applied to CDR-SOB scores from ADNI. *CPT Pharmacometrics Syst Pharmacol* **2**, e78.

[49] Eckerström M, Göthlin M, Rolstad S, Hessen E, Eckerström C, Nordlund, A, Johansson B, Svensson J, Jonsson M, Sacuiu S, Wallin, A (2017) Longitudinal evaluation of criteria for subjective cognitive decline and preclinical Alzheimer's disease in a memory clinic sample. *Alzheimers Dement (Amst)* **8**, 96-107.

[50] Roehr S, Villringer A, Angermeyer MC, Luck T, Riedel-Heller SG (2016) Outcomes of stable and unstable patterns of subjective cognitive decline–results from the Leipzig longitudinal study of the aged (LEILA75+). *BMC Geriatr* **16**, 180.

[51] Viviano RP, Damoiseaux JS (2020) Functional neuroimaging in subjective cognitive decline: Current status and a research path forward. *Alzheimers Res Ther* **12**, 23.