

1 **Prediction of 7-year's Conversion from Subjective Cognitive Decline to**  
2 **Mild Cognitive Impairment**

3  
4 Ling Yue<sup>1,2†</sup>, Dan Hu<sup>3†</sup>, Han Zhang<sup>3†</sup>, Junhao Wen<sup>4,5</sup>, Ye Wu<sup>3</sup>, Wei Li<sup>1,2</sup>, Lin Sun<sup>1,2</sup>, Xia Li<sup>1,2</sup>,  
5 Jinghua Wang<sup>1,2</sup>, Guanjun Li<sup>1,2</sup>, Tao Wang<sup>1,2\*</sup>, Dinggang Shen<sup>3,6\*</sup>, Shifu Xiao<sup>1,2\*</sup>

6  
7 <sup>1</sup>Department of Geriatric Psychiatry, Shanghai Mental Health Center, Shanghai Jiaotong  
8 University School of Medicine, Shanghai, China

9 <sup>2</sup>Alzheimer's Disease and Related Disorders Center of Shanghai Jiaotong University,  
10 Shanghai, China

11 <sup>3</sup>Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel  
12 Hill, NC, USA

13 <sup>4</sup>ARAMIS Lab, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne University, Inria, Paris,  
14 France

15 <sup>5</sup>Department of Radiology, Perelman School of Medicine, University of Pennsylvania, PA,  
16 USA

17 <sup>6</sup>Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic  
18 of Korea

19  
20 †Joint first authorship

21 \*Joint senior authors

22  
23 \*Correspondence:

24 T.W. ([wtshhwy@163.com](mailto:wtshhwy@163.com)) or to D.S. ([dgshen@med.unc.edu](mailto:dgshen@med.unc.edu)) or to S.X.  
25 ([xiaoshifu@msn.com](mailto:xiaoshifu@msn.com))

26  
27 *Acknowledgements:*

28 This work was supported by the China Ministry of Science and Technology  
29 (2009BAI77B03), Shanghai Mental Health Center (CRC2017ZD02, 2018-FX-05),  
30 Shanghai Clinical Research Center for Mental Health (SCRC-MH, 19MC1911100), the  
31 National Natural Science Foundation of China (81830059), Shanghai Jiaotong University  
32 School of Medicine (CBXJ201815, YG2016MS38), and Shanghai Municipal Human  
33 Resources Development Program (2017BR054).

34  
35 *Competing interests:*

36 We declare no competing interests.

37

38 **Abstract**

39

40 Subjective cognitive decline (SCD) is a high-risk yet less understood status years before  
41 Alzheimer's disease (AD). This work included 76 SCD individuals with two (baseline and seven  
42 years later) neuropsychological evaluations and a baseline T1-MRI. A machine learning-based  
43 model was trained based on 198 baseline neuroimage features and a battery of 25 clinical  
44 measurements to discriminate 24 progressive SCDs from 52 stable SCDs. The SCD  
45 progression was satisfactorily predicted with combined features. A history of stroke, a low  
46 education level, a low baseline MoCA score, a shrunk left amygdala, and enlarged white matter  
47 at the banks of the right superior temporal sulcus favor the progression. This is to date the  
48 largest retrospective study of SCD-to-MCI conversion with the longest follow-up, suggesting  
49 predictable far-future cognitive decline for the risky populations with baseline measures only.  
50 These findings provide valuable knowledge to the future neuropathological studies of AD in its  
51 prodromal phase.

52

53

54

## 55 Introduction

56

57 Alzheimer's disease (AD) is a neurodegenerative disorder with a slow and lengthy  
58 progression. The neuropathological process of AD is initiated at least 15-20 years before the  
59 first symptom of cognitive impairment(Elliott, Masliah, Ryan, & Silverberg, 2018; Tondelli et al.,  
60 2012). Current consensus has emphasized the need for early detection of AD based on  
61 neuroimaging (PET and MRI) (Guo, Landau, & Jagust, 2020) or invasively acquired CSF  
62 biomarkers(Frisoni et al., 2017). Detection of amnesic mild cognitive impairment (MCI), a  
63 prodromal stage of AD, maybe still too late for early intervention as the massive neuron loss  
64 and irreversible cognitive impairment may have already incurred at this stage(Ronald C.  
65 Petersen). It is crucial to explore early markers to predict possible AD conversion at an even  
66 earlier stage. Subjective cognitive decline (SCD), defined as a subjectively experienced decline  
67 in cognitive capacities in the absence of objectively measurable neuropsychological deficits, is  
68 regarded as the first indicator of AD years before MCI(Tondelli et al., 2012). Recently, the  
69 National Institute on Aging and Alzheimer's Association (NIA-AA) updated the research  
70 guideline for AD and defined SCD as a stage 2 AD between asymptomatic (stage 1) and  
71 symptomatic MCI (stage 3)(Elliott et al.).

72

73 Accumulating epidemiologic studies support the inclusion of SCD as a pre-MCI stage and  
74 a high-risk cohort of future progression(Buckley et al., 2016; Frank Jessen et al., 2020; Mitchell,  
75 Beaumont, Ferguson, Yadegarfar, & Stubbs, 2014; Reisberg, Shulman, Torossian, Leng, & Zhu,  
76 2010; Snitz et al., 2018). A meta-analysis study showed that the annual conversion rate from  
77 SCD to MCI is 6.6%(Mitchell et al., 2014). However, SCD can manifest in the healthy elderly or  
78 people with other disease as well. Some studies regarded subjective cognitive complaints as a  
79 benign symptom that may not lead to severe consequences(Hessen et al., 2017), while others  
80 reported that a small part of SCD might develop to cerebrovascular disease, Parkinson's  
81 Disease, or non-Alzheimer degenerative dementia(Frank Jessen et al., 2020). The evidence  
82 altogether indicates that SCD could be a heterogeneous cohort, including those who will  
83 eventually become AD (progressive SCD or pSCD), stay stable (sSCD), and show related  
84 symptoms to other diseases. It is extremely challenging to differentiate pSCD from sSCD at the  
85 baseline. PET or CSF can detect  $\beta$ -amyloid or tau-related biomarkers for early AD diagnosis,  
86 but both are invasive and expensive, making it unsuitable for large-cohort screening(Elliott et  
87 al., 2018). Structural MRI (sMRI) is a non-invasive brain imaging technique that can detect AD-  
88 related morphological alterations (e.g., medial temporal lobe atrophy) in the subjects with SCD  
89 (Cherbuin, Sargent-Cox, Easteal, Sachdev, & Anstey, 2015; Dubois et al., 2018; F. Jessen et  
90 al., 2006; Meiberth et al., 2015; Peter et al., 2014; Stewart et al., 2011; Striepens et al., 2010;  
91 Tijms et al., 2018; Verfaillie et al., 2016; Yue et al., 2018). However, such findings are not  
92 conclusive for SCD compared to those for AD and MCI, possibly due to the confounding effects  
93 in the cross-sectional design used in most of the studies(F. Jessen et al., 2006; Meiberth et al.,  
94 2015; Striepens et al., 2010; Yue et al., 2018). Only a handful of longitudinal sMRI studies with  
95 only 2-5 years follow-up but such a short follow-up, which is, however, still insufficient for  
96 progression prediction(Cherbuin et al., 2015; Dubois et al., 2018; EY et al., 2019; Peter et al.,  
97 2014; Stewart et al., 2011; Tijms et al., 2018; Verfaillie et al., 2016). Identifying measurable and  
98 objective baseline markers from the SCD subjects based on non-invasive neuroimaging with a

99 longer follow-up time for individualized early AD prediction is of great clinical significance.

100

101 Besides neuroimaging markers, clinical and demographic features are also considered to  
102 have strong predictive abilities(Livingston et al., 2017). However, none of them has been  
103 investigated in any SCD progression study, and whether they are strong enough to  
104 independently differentiate pSCD from sSCD at the baseline is yet unclear. Different sMRI  
105 indicators and clinical/demographic features may likely have complex relationships, and they  
106 could be jointly used to predict SCD-to-MCI progression. Hence, in this study, we aimed to use  
107 machine learning, an advanced multivariate pattern analysis method, to not only detect the  
108 progression-related markers jointly and objectively but also conduct an individualized  
109 differentiation between pSCDs and sSCDs. To validate our model and further test the efficacy  
110 of the identified contributive markers, we further compared the detected markers among various  
111 normal control (NC)/SCD/MCI groups with independent datasets.

112

113

## 114 **Results**

115

### 116 *Clinical information*

117 The demographic and clinical characteristics of the 76 SCD subjects are presented in Table  
118 1. There are 24 (31.5%) SCD subjects showing clinical progression to amnesic MCI. Compared  
119 to sSCD, the pSCDs were less educated ( $7.17\pm 4.15$  vs  $9.50\pm 3.06$  years,  $p = 0.019$ ) but with  
120 comparable baseline MoCA and MMSE scores. The two groups matched on age, gender, BMI,  
121 the status of physical diseases, lifestyle and GDS score (Table 1). At the follow-up that is seven  
122 years after the baseline scan, pSCDs developed to amnesic MCI with significant lower MoCA  
123 and MMSE scores than sSCDs but still comparable physical disease statuses and GDS score.

124

### 125 *Machine learning-based pSCD vs. sSCD classification*

126 With 233 clinical and MRI features, cost-sensitive support vector machines (CSVM)  
127 achieved a satisfactory pSCD vs. sSCD classification performance (accuracy, 69.74%,  
128 sensitivity, 62.50%, specificity, 73.08%, F1 score, 0.5660). Specifically, the algorithm  
129 successfully identified 15 pSCDs out of 24, and 38 sSCDs out of 52. Supplementary Table 1  
130 and Supplementary Table 2 lists all the features and the confusion matrix.

131

### 132 *The most contributive features in SCD progression prediction*

133 Five features were identified and consistently selected as contributive features in pSCD vs.  
134 sSCD classification (see details of the selection frequency in Supplementary Table 1). They  
135 are baseline stroke history, years of education, baseline MoCA score, baseline volume of the  
136 left amygdala, and baseline WM volume at the banks of the right superior temporal sulcus  
137 (wmSTSBanks). Figure 1A and Supplementary Table 3 show the group comparison results for  
138 these five features between pSCDs and sSCDs. In addition to education level, both MRI-based  
139 features (see Supplementary Figure 1 for locations) show significant group differences ( $p <$   
140  $0.05$ , uncorrected). The pSCD individuals show lower MoCA scores at baseline, more likely to  
141 have stroke history, lower education level, decreased left amygdala volume, and increased right  
142 wmSTSBanks volume, compared to sSCDs. With the five contributive features, we derived a

143 clinically feasible, much simpler, automatic prediction model for SCD progression based only  
144 on three clinical/demographic features and two MRI features at the baseline. Figure 1B shows  
145 the ROC curve of these five contributive features with a refined SCD progression prediction  
146 model and the AUC reached 0.7997. The distributions of the decision scores obtained by the  
147 scoring function learned by this prediction model for all the subjects are shown in Figure 1C,  
148 indicating a satisfactory separation of the two groups with only a few baseline features.

149

#### 150 *Relationship between MRI features and cognitive score*

151 The partial correlation analysis between the two important MRI features (volume of the left  
152 amygdala and the right wmSTSbanks) and the MoCA scores (baseline, follow-up, and the  
153 changes), after controlling education, stroke history, age, and gender, revealed a negative  
154 association between the baseline left amygdala volume and the follow-up MoCA score ( $r = -$   
155  $0.636$ ,  $p = 0.003$ , significant after Bonferroni corrections) for the pSCD subjects (Figure 2A). No  
156 significant association was found for the sSCD subjects. No correlation was found between the  
157 baseline volume of the right wmSTSbanks and the baseline MoCA scores or its changes.

158

#### 159 *Structural connections of the right wmSTSbanks*

160 For possible WM connections to the right wmSTSbanks (i.e., the only WM feature from the  
161 MRI detected by the machine learning-based pSCD vs. sSCD differentiation). dMRI  
162 tractography with an independent dMRI data from a healthy elderly cohort (Supplementary  
163 Experiment 1) showed that there are many WM fibers could pass through the right  
164 wmSTSbanks and linking to the right amygdala and the right hippocampus, two regions  
165 targeted by the AD pathology (Supplementary Figure 2). In addition, there are many other  
166 cortical regions, especially the right middle, superior, and inferior temporal gyri, as well as the  
167 right interior parietal and supramarginal areas and right putamen, could be reached by the WM  
168 fibers that passing through the right wmSTSbanks (Supplementary Figure 3).

169

#### 170 *Validation based on independent data sets*

171 From the CLAS database, we further identified sNC, pNC, sMCI subjects and extracted  
172 the same features from them. Together with the sSCD, all four groups were compared with the  
173 pSCD group, separately, with two-sample  $t$ -tests with Bonferroni correction. We found a similar  
174 trend that sNC, sSCD, and sMCI, whom were all potentially not affected by AD pathology, tend  
175 to have lower baseline wmSTSbanks volume compared with pSCD, with the pNC specifically  
176 sitting in-between the three stable groups and the sSCD group. Specifically, we found a  
177 significant difference ( $p_{\text{corrected}} < 0.05$ ) between sNC and pSCD in addition to the significant  
178 difference between sSCD and pSCD, as well as a trend-to-significant difference ( $p_{\text{corrected}} < 0.1$ )  
179 between sMCI and pSCD (Figure 2B, see details in Supplementary Experiment 2). Furthermore,  
180 the combined progressive group (pSCD+pNC) also showed a greater baseline right  
181 wmSTSbanks volume than the combined stable group (sSCD+sNC) (Supplementary Figure 4).  
182 From the ADNI data, although no significant difference was found from the comparison of the  
183 baseline volume of the right wmSTSbanks between sNCs and pNCs ( $p = 0.24$ ) due to lower  
184 statistical power caused by small sample size (due to short follow-up), we still spotted a similar  
185 trend to the main results, i.e., pNCs had a higher right wmSTSbanks volume than the sNCs  
186 ( $0.62 \pm 0.13$  vs.  $0.59 \pm 0.09$ , Figure 2C, see details in Supplementary Experiment 3).

## 187 Discussion

188

189 SCD is recognized as a high-risk status of AD with a normal cognitive level, which is earlier  
190 than the stage of MCI(Buckley et al., 2016; Mitchell et al., 2014; Wolfsgruber et al., 2015). In  
191 the current study, instead of predicting MCI progression, we focus on the detection of possible  
192 AD at an even earlier stage, i.e., predicting whether SCD converts to MCI. We were able to do  
193 so because of the unique CLAS dataset that focused on the risky population at even earlier  
194 stage compared to MCI and had an extremely long follow-up time (seven years). Based on a  
195 well-designed machine learning model with both comprehensive clinical features and MRI  
196 features, we demonstrated that the SCD individuals who would develop to amnesic MCI later  
197 might be identified and differentiated from the sSCDs with a combination of a few clinical  
198 information, psychometric scores, and an easily obtained, baseline grey-matter and white-  
199 matter volumetric data based on sMRI. Our results indicated that incident stroke, fewer years  
200 of education, lower baseline MoCA score, smaller left amygdala, and larger white matter at the  
201 banks of right superior temporal sulcus jointly favored amnesic MCI progression. We also  
202 showed that with a state-of-the-art, data-driven feature selection and classification, the sSCD  
203 could be individually separated from the pSCD in an automatic and objective way with  
204 satisfactory accuracy. We then proposed a clinically feasible, much-simplified prediction model  
205 with only five (three clinical and two sMRI-derived) baseline features with which an AUC of 0.8  
206 was reached. The strengths of this study include an unprecedentedly long-term follow-up, a  
207 targeted risk population (SCD) from well-characterized community samples, an advanced  
208 multivariate pattern recognition algorithm that jointly identified the associations among the  
209 clinical and brain MRI features towards an effective feature set, and a demonstrated feasibility  
210 of early computer-aided diagnosis of pSCD individuals.

211

212 While it is important to identify which individuals with SCD eventually develop clinically  
213 significant cognitive impairment years later, long term follow-up is crucial because it might take  
214 almost a decade for a subject with “compensatory normal cognition” to change from SCD to  
215 MCI (Molinuevo et al., 2017), especially for community-based cohorts(Snitz et al., 2018).  
216 However, to our best knowledge, only four SCD follow-up studies had investigated the link  
217 between baseline brain biomarkers and incident clinical progression, but all of them only  
218 followed up for 2-5 years(Dubois et al., 2018; EY et al., 2019; Tijms et al., 2018; Verfaillie et al.,  
219 2016). This might not be sufficient to fully investigate the conversion from SCD to MCI. One  
220 study only identified four out of 318 SCD individuals who progressed to prodromal AD within 30  
221 months(Dubois et al., 2018), while another study reported that community-based SCD subjects  
222 confer much lower risk of progression to MCI over three years compared to clinical  
223 cohorts(Snitz et al., 2018). Not only cutting off sample sizes, insufficient follow-up time could  
224 also misidentify many SCDs who could potentially convert to MCI, which could result in a highly  
225 mixed and heterogenous sSCD group, further reducing statistical power. Furthermore,  
226 compared to the subject recruitment from clinical sites, those recruited from communities can  
227 empower the AD early detection by enabling even earlier detection, even for those who have  
228 not developed cognitive impairment yet (on contrast, clinical cohorts could have more or less  
229 developed symptoms already at the baseline). Our study is an SCD conversion study with a  
230 long (> five years) follow-up, with advanced feature searching strategies from comprehensively

231 engineered features to identify pivotal factors that could provide early predictive value.

232

233 Our most important finding is the early enlargement of the right wmSTSBanks in the pSCD  
234 vs. sSCD. This feature characterizes the relative WM volume beneath the cortical part of the  
235 superior temporal sulcus (STS). STS is a “chameleon” in the human brain and a high-order  
236 association cortex that receives connections from different sensory modalities(Hein & Knight,  
237 2008). It is believed to be involved in diverse cognitive functions, such as audiovisual integration,  
238 as well as motion, speech, and face processing(Hein & Knight, 2008). The WM under the STS  
239 may play a crucial role in efficient information transmission to support the multifunctionality of  
240 the STS. This is further proved by our fiber tracking result, which shows dense fiber connections  
241 with lateral and medial temporal cortices, an area for multimodal information integration and  
242 the language- and memory-related cognitive functions(R, MA, S, & G, 2011). Meanwhile, STS  
243 was consistently reported to be affected by AD neuropathology (e.g., neurofibrillary tangles and  
244 neuronal loss) in a very early stage. The neural loss in the STS was found to be correlated with  
245 duration of illness, ranging from no measurable loss in the early stage (with a duration of less  
246 than one year) to more than 75% neuron reductions in the severe AD stage(Gomez-Isla et al.,  
247 1997). Hence, we interpret our finding of early increment in the STS-related WM volume as a  
248 result from increased axonal connections as compensation of other affected regions to preserve  
249 normative cognitive ability in the SCDs who would convert to AD and such compensation could  
250 be enabled by the largely preserved STS neurons working with an elevated effort or load in the  
251 start-up phase of clinical AD. Our additional comparisons with sNC, sMCI, and pNC cohorts  
252 from the CLAS database further validate such a compensation hypothesis (Figure 2B). Of note,  
253 additional support came from the results with a similar trend of increased baseline  
254 wmSTSBanks volume in pNC compared to sNC from the ADNI data (Figure 2C), which might  
255 become significant providing sufficient follow-up time and more balanced sample size. From  
256 the functional study point of view, an 18FDG-PET study shows increased metabolism in the  
257 right WM adjacent with the inferior parietal lobe in SCD subjects compared to NCs(Scheef et  
258 al., 2012). Another study revealed increased regional cerebral blood flow (rCBF) along the  
259 default mode network regions in SCD compared to NCs(Wang et al., 2013). These areas are  
260 either closed to or include the wmSTSBanks, further supporting our compensation hypothesis.

261

262 Another potential reason for the greater WM underneath the STS could be grey-matter  
263 atrophy in the adjacent areas. A previous study reported decreased grey matter at the banks of  
264 STS in the MCI subjects who converted to AD within three years(Killiany et al., 2000). A very  
265 recent PET study with ADNI data also found that cognitively normal individuals with high A $\beta$   
266 burden in the cortical region of the banks of STS were at an increased risk of cognitive  
267 decline(Guo et al., 2020). Of note, we used a ratio of wmSTSBanks volume over the total white  
268 matter volume as it is a conventional way to exclude confounding effects (e.g., global WM  
269 atrophy)(Nordenskjold et al., 2013), there could be a possibility that it is the shrunk total white  
270 matter volume that led to such a result. However, we found that it is less likely because we  
271 found that the pSCD still showed larger volume than sSCD even without conducting such a  
272 global normalization (Supplementary table 4).

273

274 Another contributive sMRI feature is the decreased baseline volume of the left amygdala

275 in pSCD compared to sSCD. Amygdala atrophy in SCD was consistently reported(Hu et al.,  
276 2019; Striepens et al., 2010; Yue et al., 2018), especially by a recent study that revealed a  
277 smaller left amygdala in the SCDs with CSF biomarker (A $\beta$ 42) compared to those without(Hu  
278 et al., 2019). Another long-term follow-up study also suggests that atrophy of the amygdala in  
279 cognitively intact elderly people could predict dementia occurred 6 years later.(den Heijer et al.,  
280 2006) Also, from the cognitively intact elderly cohort, Tondelli et al. reported the reduced  
281 amygdala volume was detectable at least four years before any cognitive symptoms.(Tondelli  
282 et al., 2012) Moreover, decreased amygdala is also be associated with TDP43-related  
283 dementia(N et al., 2019). Hence, atrophy in this area in SCD may be an early marker for future  
284 cognitive decline. It is noted that we only found a relatively weak statistical difference ( $p = 0.011$ )  
285 in the amygdala by statistical analysis, which means that this region could be undetectable from  
286 a large number of candidate features if using traditional mass-univariate analysis. Instead, with  
287 advanced feature selection and multivariate pattern analysis in a unified framework of machine  
288 learning, we were able to detect such a contributive sMRI feature together with others.

289

290 While the pSCD subjects had a smaller baseline amygdala compared to sSCDs, we found  
291 a negative correlation between the baseline amygdala volume and the MoCA scores at the  
292 follow-up across the pSCD subjects while no such correlation was found from sSCD. This  
293 indicates that, for pSCD, larger amygdala could be associated with worse cognitive outcome.  
294 Such results seem contradictory as positive correlation was originally anticipated as sSCD (with  
295 good outcomes) was found to have a larger amygdala compared to pSCD (but we found that a  
296 larger amygdala in pSCD predicted worse outcome). Therefore, we think that the negative  
297 correlation could support a well-known hypothesis that higher cognitive reservation (as  
298 reflected by a larger amygdala at the baseline) is associated with a more rapid decline later  
299 (and therefore has much worse cognitive outcomes afterward) for the risky populations(Amieva  
300 et al., 2014). We further came into an updated hypothesis, that is, as long as the baseline  
301 amygdala size is below average, those with a larger amygdala may cover up baseline  
302 symptoms and, once they become clinically significant, the subjects could have more rapidly  
303 declined cognitive abilities. This finding might have helpful for developing possible preventative  
304 solutions or treatments in the future.

305

306 In addition to the sMRI features, the study also revealed three contributive clinical features  
307 in distinguishing the cognitive outcome of SCDs, including baseline MoCA, years of education,  
308 and history of stroke, all of which have been reported as indicators of poor cognitive prognosis  
309 in the previous studies(Amieva et al., 2014; K et al., 2019; Nasreddine et al., 2005). Again,  
310 traditional statistical analysis could not reveal significant differences in MoCA and stroke history  
311 (Supplementary Table 3). However, this does not contradict our findings, as our prediction  
312 model considered these features jointly instead of independently. Our method is quite effective,  
313 which is also because we utilized an effective feature selection algorithm using Relief, an  
314 algorithm that takes relationships among features into consideration. Furthermore, we derived  
315 a simplified model with only five features, all of which can be obtained at the baseline. This  
316 renders our model a merit of predicting SCD conversion at the baseline seven years before  
317 noticeable cognitive impairment, suitable for large-cohort screening in the future.

318

319           However, this study also has limitations. First, our study does not consist of annual follow-  
320 ups, which maybe helpful for timely identifying SCD's conversion to MCI. Second, although we  
321 followed up on our SCD subjects for a long time compared to the previous studies, it is still  
322 likely that more SCD will progress to MCI at later time(Kaup, Nettiksimmons, LeBlanc, & Yaffe,  
323 2015). Third, our data does not include FDG-PET, amyloid- $\beta$  marker, or APOE genotype,  
324 making it unclear whether the progression was due to AD or not.

325

326           Nevertheless, this is the first sMRI study of SCD subjects with unprecedentedly long follow-  
327 up time. We, with advanced machine learning from the comprehensively and thoroughly  
328 engineered feature sets, identified key baseline early AD markers comprising both clinical and  
329 neuroimaging features with an individualized SCD progression prediction task. We further  
330 proposed a much simpler prediction model with only five features and a satisfactory predictive  
331 accuracy. Our results provide a potentially feasible and objective early diagnostic tool and help  
332 better understanding of the pathology of AD in its prodromal phase.

333

334

## 335 **Methods**

336

### 337 *Study cohorts*

338           The study is part of the China Longitudinal Ageing Study (CLAS) of Cognitive Impairment,  
339 which is a community-based study initiated in 2011(Xiao et al., 2016; Xiao et al., 2013). The  
340 current data constitute samples from Shanghai where all the subjects received a baseline T1-  
341 MRI scan. Some of the data was used in a cross-sectional study on SCD(Yue et al., 2018). In  
342 the current study, as the 7-year follow-up has recently finished, we could include all the subjects  
343 with both baseline SCD diagnosis and the follow-up visit. Of all the 111 SCD subjects defined  
344 at the baseline, the 7-year follow-up was completed in 92 (82.8%) subjects (Table 1). The SCDs  
345 with and without follow-up did not differ in gender (male/female: 44/48 vs. 9/10), age  
346 (69.23 $\pm$ 7.49 vs. 69.84 $\pm$ 6.90 years), years of education (9.01 $\pm$ 3.78 vs. 9.21 $\pm$ 4.44 years), the  
347 Mini-Mental State Examination (MMSE) score (27.55 $\pm$ 2.23 vs. 27.16 $\pm$ 2.59), or the Montreal  
348 Cognitive Assessment (MoCA) score (23.54 $\pm$ 4.56 vs. 23.00 $\pm$ 4.66). Of note, even the 7-year  
349 follow-up cannot guarantee the conversion from SCD to AD; therefore, we focused on on  
350 predicting SCD conversion to amnesic MCI (the pSCD is hereby defined accordingly from here  
351 on). Although we cannot guarantee that all the converted cases were due to AD, we still remove  
352 16 SCD subjects diagnosed with follow-up evaluations as other neurodegeneration diseases,  
353 psychogenic diseases, or organic etiologies to avoid confounding effects. Figure 3 shows a  
354 flowchart of participant selection. The study protocol was approved by the Ethical Committee  
355 of the Shanghai Mental Health Center. All participants signed written informed consent before  
356 enrollment.

357           All the SCD participants were assessed by self-report and diagnosed according to the  
358 following criteria(Frank Jessen et al., 2014; Molinuevo et al., 2017): 1) the onset age of memory  
359 decline is > 60 years old; 2) presence of gradual memory decline has persisted for  $\geq$  6 months;  
360 3) objective memory performance at baseline is within the normal range. More details on the  
361 SCD diagnosis procedure are listed elsewhere(Yue et al., 2018). MCI was clinically diagnosed  
362 according to the Peterson's criteria of MCI with consideration of comorbid conditions. At the

363 follow-up visit(R. C. Petersen et al., 1999), we had 24 pSCD subjects who had progressed to  
364 MCI and the 52 sSCD subjects who kept cognitive normal, none of which converted to AD at  
365 the follow-up visit.

366

### 367 *Subject assessment*

368 We performed a comprehensive battery of sociodemographic and physical health  
369 measurements for each subject. As potential risk factors of SCD progression(Livingston et al.,  
370 2017), sociodemographic data (gender, age, years of education, years of smoke, years of drink,  
371 and Body Mass Index (BMI)) and status of somatic diseases (years of hypertension,  
372 hyperlipidemia, diabetes mellitus, sleeping disorder, incident stroke, heart disease, and any  
373 type of surgery) at the baseline were collected, as well as the new onsets of the above somatic  
374 diseases and years of hearing loss acquired over the follow-up period. Three psychological  
375 assessments were carried out at the baseline, including MoCA (measuring overall cognitive  
376 performance), geriatric depression scale (GDS, also acquired at the follow-up due to the close  
377 association between depressive symptoms and cognitive performance), and social support  
378 questionnaire (measuring perceptions of social support and its satisfaction).

379

### 380 *MRI data acquisition and analysis*

381 T1-weighted 3D magnetization prepared rapid gradient echo (MPRAGE) sMRI were  
382 acquired from a 3.0T MRI scanner (Siemens MAGNETOM VERIO, Germany) with the following  
383 parameters: TR = 2300 ms, TE = 2.98 ms, flip angle, 9°, matrix size, 240 × 256, field of view  
384 (FOV), 240 × 256 mm, slice thickness, 1.2 mm, and number of slices, 176. All the sMRI data  
385 were processed using Clinica (Routier et al., 2018) ([www.clinica.run](http://www.clinica.run)) in FreeSurfer v6.0  
386 ([surfer.nmr.mgh.harvard.edu](http://surfer.nmr.mgh.harvard.edu)), including segmentation of the subcortical structures, extraction  
387 of cortical surfaces, cortical thickness estimation, spatial registration, and parcellation into 46  
388 global structures. Quality control was carefully conducted by overlapping the output  
389 parcellations on the FreeSurfer's template and visual assessment was carried out to ensure  
390 the registration and parcellation quality. Besides the 46 global volumetric measurements of  
391 different brain structures, we further derive 68 cortical, 68 white-matter (WM), and 16  
392 subcortical regions of interest (ROIs) based on the Desikan-Killiany atlas(Desikan et al., 2006).  
393 For each cortical ROI, mean cortical thickness was extracted without normalization(Westman,  
394 Aguilar, Muehlboeck, & Simmons, 2012). For each WM or subcortical ROI, the regional  
395 volumetric measure was extracted and normalized by the total volume of WM and subcortical  
396 regions, respectively. All the 198 (46 + 68 + 68 + 16) MRI features were listed in Supplementary  
397 Table 1.

398

### 399 *SCD progression prediction*

400 Due to the potential redundancy in the 223 (25 clinical + 198 MRI) features, it is necessary  
401 to select the most predictive features to train a concise model and avoid overfitting. We used a  
402 three-stage feature selection scheme. First, a Relief algorithm was employed to rank the  
403 features according to the extent of associations with the labels (i.e., pSCD and sSCD), which  
404 has been widely applied in machine learning tasks to improve classification performance, e.g.  
405 genetic analysis for diseases, Parkinson's disease diagnosis and contributive feature  
406 identification(Urbanowicz, Meeker, La Cava, Olson, & Moore, 2018). Considering contextual

407 information during ranking, Relief can properly handle strong dependency among features.  
408 Specifically, it generates a robust feature ranking by identifying the feature value differences  
409 between nearest-neighbor instance pairs. The number of neighbors was selected from a range  
410 between 1 and 20 at a step of 1. Therefore, we generated 20 ranking lists of features. Second,  
411 an overall ranking for all the features was obtained by integrating the 20 ranking lists and  
412 calculating the occurrence frequency of each feature in the top-50 of each ranking list. Finally,  
413 the classification performance based on CSVMs, which better handles the imbalanced samples  
414 than the traditional SVM) was derived to evaluate the ranked feature set and determine the  
415 predictive features(Park, Luo, Parhi, & Netoff, 2011). F1 score (defined as the harmonic  
416 average of the precision and sensitivity, i.e.,  $F1 \text{ score} = 2 \times (\text{precision} \times \text{sensitivity}) / (\text{precision} + \text{sensitivity})$ )  
417 was taken as the classification performance evaluation metric for further feature  
418 selection and parameter optimization of CSVM. A sequential forward selection strategy was  
419 adopted in this step, which sequentially added the features according to their ranks until adding  
420 a new feature did not improve the CSVMs' performance. The same CSVMs were used for SCD  
421 progression prediction. We set a higher misclassification penalty (with a freely estimable cost  
422 ratio, pSCD:sSCD) for the pSCD samples compared to those for the sSCD samples in the  
423 CSVMs to alleviate the problem caused by unbalanced samples, as pSCDs were much fewer  
424 than sSCDs.

425 The prediction model was validated by nested leave-one-out cross-validation (LOOCV).  
426 The feature set selected by the inner LOOCV according to the above method was used to train  
427 the classifier and the final prediction performance (including accuracy, sensitivity, specificity,  
428 and F1 score) was obtained by the outer LOOCV. In each iteration of the outer LOOCV, one  
429 subject was left out as a testing sample and the remaining 75 subjects were for training (these  
430 75 subjects were fed into the inner LOOCV where the predictive features were selected). The  
431 trained model was applied to the testing sample for generating a predicted label to compare  
432 with the ground truth. The outer LOOCV went through all the subjects. The frequency of each  
433 feature being selected across the 76 outer LOOCV iterations was used to assess feature  
434 importance. A feature was regarded as a contributive feature if this frequency is higher than 95%  
435 (i.e., selected from 72 out of 76 LOOCV iterations). Figure 4 shows the flowchart of the entire  
436 analysis.

437 To facilitate clinical application, after the contributive features were identified, we re-trained  
438 a refined SCD progression prediction model with only these contributive features as a  
439 representative prediction model. Based on this model, we drew a receiver operating  
440 characteristic (ROC) curve and used the area under the curve (AUC) to measure the  
441 discriminant ability in the prediction of SCD progression. We further derived an intuitive  
442 "decision score" for each subject according to the scoring function learned by such a  
443 representative prediction model.

444

#### 445 *Statistical analysis*

446 Group comparisons of the demographic and clinical data were conducted using two-tailed  
447 independent samples *t*-tests for continuous variables and  $\chi^2$ -tests for dichotomous variables  
448 with SPSS 19.0 (IBM Corp.). Differences in the global cognition (i.e., MoCA and MMSE)  
449 between groups were tested after controlling for age, gender and years of education. After  
450 detecting contributive features, a *t*-test or  $\chi^2$ -test was carried out for each of them to see if there

451 was any group difference between the pSCD and sSCD groups at the baseline. Specifically,  
452 we evaluated whether there were unique contributions between the identified MRI features and  
453 the cognitive measurement (as evaluated by MoCA) in the pSCD group and sSCD group,  
454 separately. This was achieved by partial correlation analysis to remove the possible influence  
455 of other non-MRI contributive features (education level and stroke history) as well as age and  
456 gender. For the partial correlation, MoCA scores at the baseline and follow-up, as well as 7  
457 years' MoCA changes were separately used to investigate how the identified MRI features were  
458 associated with cognitive abilities (either with the two terminal scores or their changes).  
459 Bonferroni correction was conducted for multiple comparison correction.

460

#### 461 *Result evaluations with independent datasets*

462 In addition to the main analysis, we analyzed various independent datasets to further validate  
463 and evaluate the main findings from the CLAS data. First, for the identified contributive MRI  
464 feature at the WM (i.e., the banks of the right superior temporal sulcus), we checked whether  
465 there were fiber connections between this region and amygdala and hippocampus, two regions  
466 specifically targeted by AD, based on an independent diffusion MRI (dMRI) tractography  
467 dataset that was used in our previous study (from 15 (10 females) healthy elderly subjects with  
468 mean age of  $70.6 \pm 6.2$  years, see Supplementary materials)(Li et al., 2016). We investigated  
469 which cortical areas could be reached by the tractography streams passing through the right  
470 wmSTSbanks (see details on fiber tracking in Supplementary Experiment 1). Second, we  
471 checked the volume of the wmSTSbanks in different NC cohorts (including a group of stable  
472 NCs or sNC and another group of progressive NCs or pNC) and a stable MCI (sMCI) cohort,  
473 also selected from the CLAS database (but with smaller sample size compared to the  
474 sSCD/pSCD in the main analysis), to see if there was any consistent (possibly much earlier, as  
475 progressive NCs were used) trend with the main results (see details in Supplementary  
476 Experiment 2). Third, we used sNC and pNC subjects from an independent ADNI2 dataset  
477 (<http://adni.loni.usc.edu/>), another widely adopted longitudinal dataset for tracking AD  
478 progression to investigate if any similar trend can be found (see details in Supplementary  
479 Experiment 3). The follow-up time for the ADNI data is  $45.39 \pm 7.89$  months ( $\sim 4$  years), smaller  
480 than the CLAS data.

481

482

483

484 **Table 1**  
 485 **Demographic and clinical data of the subjects with stable and progressive subjective**  
 486 **cognitive decline**

	sSCD (N = 52)	pSCD (N = 24)	t/λF	p
Age, year	68.56 (7.19)	71.29 (6.55)	2.505	0.118
Gender (male/female)	25/27	11/13	0.033	0.856
Education, year	9.50 (3.06)	7.17 (4.15)	7.584	0.019
Follow-up interval, month	83.97 (0.97)	84.16 (1.08)	0.578	0.450
<b>Baseline</b>				
MoCA*	24.04 (3.90)	22.96 (4.49)	0.502	0.481
MMSE*	27.71 (2.15)	26.75 (2.61)	0.038	0.846
BMI	24.29 (2.91)	23.59 (4.29)	0.725	0.473
Smoke, year	12.62 (17.81)	6.83 (15.78)	1.424	0.161
Drink, year	5.54 (11.71)	7.21 (16.54)	-0.505	0.615
Hypertension, year	6.45 (9.37)	7.79 (9.46)	-0.578	0.565
diabetes mellitus, year	1.84 (7.38)	2.29 (6.38)	-0.260	0.795
Hyperlipidemia, year	1.37 (4.68)	1.96 (5.41)	-0.489	0.627
Heart disease (yes/no)	16/36	8/16	0.050	0.823
Stroke history (yes/no)	3/49	4/20	1.211	0.271
Surgical history (yes/no)	29/21	10/14	1.736	0.188
Sleep disorder, year	1.04 (4.52)	2.75 (7.02)	-1.094	0.282
GDS score	3.37 (3.90)	2.29 (2.40)	1.241	0.218
Social support score	36.52 (9.59)	33.42 (10.67)	1.265	0.210
<b>7-year follow-up</b>				
MoCA*	23.12 (3.54)	17.33 (3.82)	29.358	<0.001
MMSE*	27.58 (1.95)	24.08 (3.87)	17.497	<0.001
BMI	25.01 (2.99)	23.69 (3.95)	1.607	0.112
Sleep disorder, year	0.96 (2.20)	2.25 (4.50)	-1.336	0.192
Hearing loss, year	0.43 (1.10)	0.79 (1.44)	-1.093	0.282
GDS score	4.52 (3.47)	5.29 (4.67)	-0.806	0.423
Hypertension, year <sup>†</sup>	0.06 (0.31)	0.25 (0.74)	-1.229	0.230
Diabetes mellitus, year <sup>†</sup>	0.500 (1.49)	0.46 (1.25)	0.107	0.915
Hyperlipidemia, year <sup>†</sup>	0.21 (0.87)	0.08 (0.41)	0.685	0.495
Heart disease (yes/no) <sup>†</sup>	6/46	4/20	0.062	0.803
Stroke history (yes/no) <sup>†</sup>	7/45	2/22	0.068	0.794
Surgical history (yes/no) <sup>†</sup>	9/43	8/16	2.429	0.119

487 Data presented are the means (standard deviations) or sample size (N). sSCD, stable subjective cognitive  
 488 decline, pSCD, progressive subjective cognitive decline, MoCA, Montreal Cognitive Assessment, MMSE,  
 489 Mini-Mental State Examination, BMI, Body Mass Index, GDS, Geriatric Depression Scale.

490 \*All the analyses of neuropsychological variables (i.e., MoCA and MMSE) were conducted after controlling  
 491 for age, gender, and education.

492 † Newly diagnosed after the baseline during the 7-year follow-up.

493

494 **FIGURE LEGENDS**

495

496 **Figure 1 | The five most contributive features.** (A) The error bar plots of the five most  
497 contributive features in z-scores identified by cost-sensitive support vector machines  
498 (CSVMs) in the sSCD vs. pSCD classification. Error bars indicate standard errors. (B) The  
499 ROC curve generated by using the five most contributive features to classify pSCDs and  
500 sSCDs (AUC = 0.7997). (C) The decision scores generated by the refined SCD  
501 progression prediction model with the five features. The white dot in the middle is the  
502 median value and the thick black bar in the center represents the interquartile range. The  
503 thin black line extended from it represents the upper (max) and lower (min) adjacent values  
504 in the data. sSCD, stable subjective cognitive decline, pSCD, progressive subjective  
505 cognitive decline, wmSTSBanks, baseline white matter volume at the banks of the right  
506 superior temporal sulcus, ROC, receiver operating characteristic, AUC, area under the  
507 curve, MoCA, Montreal Cognitive Assessment.

508

509 **Figure 2 | Clinical correlation, and validations based on CLAS study and ADNI.** (A)  
510 The scatter map shows a significant correlation (after removing the confounding effects)  
511 between the volume of the left amygdala and the MoCA score at the follow-up in the pSCD  
512 group ( $p < 0.05$ , after Bonferroni correction, where the uncorrected  $p$  threshold equals  
513  $0.05/(2 \text{ groups} \times 2 \text{ MRI features} \times 3 \text{ MoCA scores})$ , or 0.0042). (B) Comparison of the  
514 baseline right wmSTSBanks' volume among the groups with different cognitive functions  
515 at the baseline and the follow-up from the CLAS database, with additional cohorts of sNC,  
516 pNC, and sMCI compared to pSCD, separately (\* indicates  $p < 0.1$ , after Bonferroni  
517 correction with a threshold of  $0.1/4 = 0.025$  (as four different groups were compared with  
518 pSCD) to the original  $p$ -values; \*\* indicates  $p < 0.05$ , after Bonferroni correction with a  
519 threshold of  $0.05/4 = 0.0125$  to the original  $p$ -values). (C) Comparison of the baseline right  
520 wmSTSBanks volume between sNC and pNC groups from the ADNI data. Error bars  
521 indicate standard errors. The region volume was standardized (no unit). sNC, stable  
522 normal control, sSCD, stable subjective cognitive decline, pNC, progressive normal control,  
523 pSCD, progressive subjective cognitive decline, sMCI, stable mild cognitive impairment,  
524 MoCA, Montreal Cognitive Assessment.

525

526 **Figure 3 | Subject selection flowchart.** SCD, subjective cognitive decline, MCI, mild  
527 cognitive impairment.

528

529 **Figure 4 | Schematic flowchart of the SCD-to-MCI prediction framework based on the**  
530 **features extracted from sMRI and clinical measurement.** LOOCV, leave-one-out cross-  
531 validation, SVM, support vector machine, SCD, subjective cognitive decline, MCI, mild  
532 cognitive impairment.

533

534

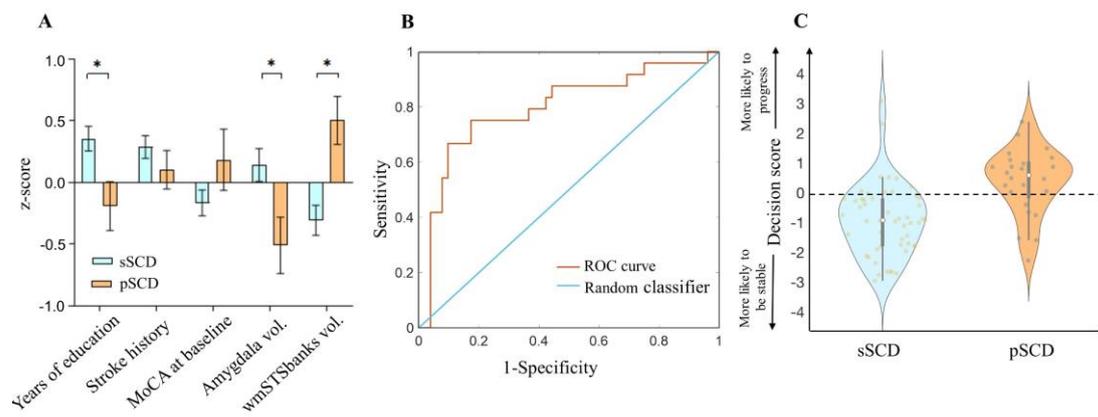
## 535 References

- 536 Amieva, H el ene, Mokri, Hind, Le Goff, M elanie, Meillon, C eline, Jacqmin-Gadda, H el ene, Foubert-  
537 Samier, Alexandra, . . . Dartigues, Jean-Fran ois. (2014). Compensatory mechanisms in  
538 higher-educated subjects with Alzheimer's disease: a study of 20 years of cognitive  
539 decline. *Brain*, *137*(4), 1167-1175.
- 540 Buckley, R. F., Maruff, P., Ames, D., Bourgeat, P., Martins, R. N., Masters, C. L., . . . study, Aibl. (2016).  
541 Subjective memory decline predicts greater rates of clinical progression in preclinical  
542 Alzheimer's disease. *Alzheimers Dement*, *12*(7), 796-804. doi:10.1016/j.jalz.2015.12.013
- 543 Cherbuin, Nicolas, Sargent-Cox, Kerry, Easta, Simon, Sachdev, Perminder, & Anstey, Kaarin J.  
544 (2015). Hippocampal atrophy is associated with subjective memory decline: The PATH  
545 Through Life study. *The American journal of geriatric psychiatry*, *23*(5), 446-455.
- 546 den Heijer, T., Geerlings, M. I., Hoebeek, F. E., Hofman, A., Koudstaal, P. J., & Breteler, M. B. (2006).  
547 Use of hippocampal and amygdalar volumes on magnetic resonance imaging to predict  
548 dementia in cognitively intact elderly people. *Archives of General Psychiatry*, *63*(1), 57-62.  
549 doi:10.1001/archpsyc.63.1.57
- 550 Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J.  
551 (2006). An automated labeling system for subdividing the human cerebral cortex on MRI  
552 scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980.  
553 doi:10.1016/j.neuroimage.2006.01.021
- 554 Dubois, Bruno, Epelbaum, Stephane, Nyasse, Francis, Bakardjian, Hovagim, Gagliardi, Geoffroy,  
555 Uspenskaya, Olga, . . . Potier, Marie-Claude. (2018). Cognitive and neuroimaging features  
556 and brain  $\beta$ -amyloidosis in individuals at risk of Alzheimer's disease (INSIGHT-preAD): a  
557 longitudinal observational study. *The Lancet Neurology*, *17*(4), 335-346.
- 558 Elliott, Cerise, Masliah, Eliezer, Ryan, Laurie, & Silverberg, Nina. (2018). NIA-AA Research  
559 Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers & Dementia*  
560 *the Journal of the Alzheimers Association*, *14*(4), 535.
- 561 EY, Lim, YS, Shim, YJ, Hong, SY, Ryu, AH, Cho, & DW, Yang. (2019). Different Cortical Thinning  
562 Patterns Depending on Their Prognosis in Individuals with Subjective Cognitive Decline.  
563 *Dementia and neurocognitive disorders*, *18*(4), 113-121. doi:10.12779/dnd.2019.18.4.113
- 564 Frisoni, Giovanni B, Boccardi, Marina, Barkhof, Frederik, Blennow, Kaj, Cappa, Stefano, Chiotis,  
565 Konstantinos, . . . Gietl, Anton. (2017). Strategic roadmap for an early diagnosis of  
566 Alzheimer's disease based on biomarkers. *The Lancet Neurology*, *16*(8), 661-676.
- 567 Gomez-Isla, T., Hollister, R., West, H., Mui, S., Growdon, J. H., Petersen, R. C., . . . Hyman, B. T. (1997).  
568 Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease.  
569 *Ann Neurol*, *41*(1), 17-24. doi:10.1002/ana.410410106
- 570 Guo, Tengfei, Landau, Susan M., & Jagust, William J. (2020). Detecting earlier stages of amyloid  
571 deposition using PET in cognitively normal elderly adults. *Neurology*, *94*(14), e1512-e1524.  
572 doi:10.1212/wnl.00000000000009216
- 573 Hein, G., & Knight, R. T. (2008). Superior temporal sulcus--It's my area: or is it? *J Cogn Neurosci*,  
574 *20*(12), 2125-2136. doi:10.1162/jocn.2008.20148
- 575 Hessen, E, Eckerstr m, M, Nordlund, A, Selseth, Almdahl I, St r m, J, Bjerke, M, . . . Reinvang,  
576 I. (2017). Subjective Cognitive Impairment Is a Predominantly Benign Condition in  
577 Memory Clinic Patients Followed for 6 Years: The Gothenburg-Oslo MCI Study. *Dement*  
578 *Geriatr Cogn Dis Extra*, *7*(1), 1-14.

- 579 Hu, X., Teunissen, C. E., Spottke, A., Heneka, M. T., Duzel, E., Peters, O., . . . Jessen, F. (2019). Smaller  
580 medial temporal lobe volumes in individuals with subjective cognitive decline and  
581 biomarker evidence of Alzheimer's disease-Data from three memory clinic studies.  
582 *Alzheimers Dement*, *15*(2), 185-193. doi:10.1016/j.jalz.2018.09.002
- 583 Jessen, F., Feyen, L., Freymann, K., Tepest, R., Maier, W., Heun, R., . . . Scheef, L. (2006). Volume  
584 reduction of the entorhinal cortex in subjective memory impairment. *Neurobiology of*  
585 *Aging*, *27*(12), 1751-1756. doi:10.1016/j.neurobiolaging.2005.10.010
- 586 Jessen, Frank, Amariglio, Rebecca E., Buckley, Rachel F., van der Flier, Wiesje M., Han, Ying,  
587 Molinuevo, José Luis, . . . Wagner, Michael. (2020). The characterisation of subjective  
588 cognitive decline. *The Lancet Neurology*, *19*(3), 271-278. doi:10.1016/s1474-  
589 4422(19)30368-0
- 590 Jessen, Frank, Amariglio, Rebecca E., van Boxtel, Martin, Breteler, Monique, Ceccaldi, Mathieu,  
591 Chételat, Gaël, . . . Subjective Cognitive Decline Initiative Working, Group. (2014). A  
592 conceptual framework for research on subjective cognitive decline in preclinical  
593 Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*,  
594 *10*(6), 844-852. doi:10.1016/j.jalz.2014.01.001
- 595 K, Blom, HL, Koek, MHT, Zwartbol, Y, van der Graaf, L, Kessler, GJ, Biessels, & MI, Geerlings. (2019).  
596 Subjective cognitive decline, brain imaging biomarkers, and cognitive functioning in  
597 patients with a history of vascular disease: the SMART-Medea study. *Neurobiology of*  
598 *Aging*, *84*, 33-40. doi:10.1016/j.neurobiolaging.2019.07.011
- 599 Kaup, Allison R, Nettiksimmons, Jasmine, LeBlanc, Erin S, & Yaffe, Kristine. (2015). Memory  
600 complaints and risk of cognitive impairment after nearly 2 decades among older women.  
601 *Neurology*, *85*(21), 1852-1858.
- 602 Killiany, R. J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., . . . Albert, M. S. (2000). Use  
603 of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Ann*  
604 *Neurol*, *47*(4), 430-439.
- 605 Li, Guanjun, Tang, Yingying, Yue, Ling, Wang, Jinghua, Zhang, Jianye, Xiao, Shifu, & Li, Huafang.  
606 (2016). Diffusion tensor imaging of patients with behavioral variant frontotemporal  
607 dementia: a controlled study. *Int J Clin Exp Med*, *9*(2), 1500-1509.
- 608 Livingston, Gill, Sommerlad, Andrew, Orgeta, Vasiliki, Costafreda, Sergi G, Huntley, Jonathan, Ames,  
609 David, . . . Cohen-Mansfield, Jiska. (2017). Dementia prevention, intervention, and care.  
610 *The Lancet*, *390*(10113), 2673-2734.
- 611 Meiberth, D., Scheef, L., Wolfsgruber, S., Boecker, H., Block, W., Traber, F., . . . Jessen, F. (2015).  
612 Cortical thinning in individuals with subjective memory impairment. *Journal of Alzheimers*  
613 *Disease*, *45*(1), 139-146. doi:10.3233/jad-142322
- 614 Mitchell, AJ, Beaumont, H, Ferguson, D, Yadegarfar, M, & Stubbs, B. (2014). Risk of dementia and  
615 mild cognitive impairment in older people with subjective memory complaints: meta-  
616 analysis. *Acta Psychiatrica Scandinavica*, *130*(6), 439-451.
- 617 Molinuevo, José L., Rabin, Laura A., Amariglio, Rebecca, Buckley, Rachel, Dubois, Bruno, Ellis,  
618 Kathryn A., . . . Jessen, Frank. (2017). Implementation of subjective cognitive decline criteria  
619 in research studies. *Alzheimer's & Dementia*, *13*(3), 296-311.  
620 doi:<https://doi.org/10.1016/j.jalz.2016.09.012>
- 621 N, Makkinejad, JA, Schneider, J, Yu, SE, Leurgans, A, Kotrotsou, AM, Evia, . . . K, Arfanakis. (2019).  
622 Associations of amygdala volume and shape with transactive response DNA-binding

- 623 protein 43 (TDP-43) pathology in a community cohort of older adults. *Neurobiology of*  
624 *Aging*, 77, 104-111. doi:10.1016/j.neurobiolaging.2019.01.022
- 625 Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow,  
626 H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild  
627 cognitive impairment. *J Am Geriatr Soc*, 53(4), 695-699. doi:10.1111/j.1532-  
628 5415.2005.53221.x
- 629 Nordenskjold, R., Malmberg, F., Larsson, E. M., Simmons, A., Brooks, S. J., Lind, L., . . . Kullberg, J.  
630 (2013). Intracranial volume estimated with commonly used methods could introduce bias  
631 in studies including brain volume measurements. *Neuroimage*, 83, 355-360.  
632 doi:10.1016/j.neuroimage.2013.06.068
- 633 Park, Yun, Luo, Lan, Parhi, Keshab K, & Netoff, Theoden. (2011). Seizure prediction with spectral  
634 power of EEG using cost-sensitive support vector machines. *Epilepsia*, 52(10), 1761-1770.
- 635 Peter, Jessica, Scheef, Lukas, Abdulkadir, Ahmed, Boecker, Henning, Heneka, Michael, Wagner,  
636 Michael, . . . Initiative, Alzheimer's Disease Neuroimaging. (2014). Gray matter atrophy  
637 pattern in elderly with subjective memory impairment. *Alzheimer's & Dementia*, 10(1), 99-  
638 108.
- 639 Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild  
640 cognitive impairment - Clinical characterization and outcome. *Archives of Neurology*,  
641 56(3), 303-308. doi:DOI 10.1001/archneur.56.3.303
- 642 Petersen, Ronald C. (2009). Early diagnosis of Alzheimer's disease: is MCI too late? *Current*  
643 *Alzheimer research*, 6(4), 324-330. doi:10.2174/156720509788929237
- 644 R, Ishibashi, MA, Lambon Ralph, S, Saito, & G, Pobric. (2011). Different roles of lateral anterior  
645 temporal lobe and inferior parietal lobule in coding function and manipulation tool  
646 knowledge: evidence from an rTMS study. *Neuropsychologia*, 49(5), 1128-1135.  
647 doi:10.1016/j.neuropsychologia.2011.01.004
- 648 Reisberg, B., Shulman, M. B., Torossian, C., Leng, L., & Zhu, W. (2010). Outcome over seven years  
649 of healthy adults with and without subjective cognitive impairment. *Alzheimers Dement*,  
650 6(1), 11-24. doi:10.1016/j.jalz.2009.10.002
- 651 Routier, Alexandre, Guillon, Jérémy, Burgos, Ninon, Samper-Gonzalez, Jorge, Wen, Junhao,  
652 Fontanella, Sabrina, . . . Gori, Pietro. (2018). *Clinica: an open source software platform for*  
653 *reproducible clinical neuroscience studies*. Paper presented at the Annual meeting of the  
654 Organization for Human Brain Mapping-OHBM 2018.
- 655 Scheef, L., Spottke, A., Daerr, M., Joe, A., Striepens, N., Kolsch, H., . . . Jessen, F. (2012). Glucose  
656 metabolism, gray matter structure, and memory decline in subjective memory impairment.  
657 *Neurology*, 79(13), 1332-1339. doi:10.1212/WNL.0b013e31826c1a8d
- 658 Snitz, Beth E., Wang, Tianxiu, Cloonan, Yona Keich, Jacobsen, Erin, Chang, Chung-Chou H., Hughes,  
659 Tiffany F., . . . Ganguli, Mary. (2018). Risk of progression from subjective cognitive decline  
660 to mild cognitive impairment: The role of study setting. *Alzheimer's & Dementia*, 14(6),  
661 734-742. doi:<https://doi.org/10.1016/j.jalz.2017.12.003>
- 662 Stewart, R., Godin, O., Crivello, F., Maillard, P., Mazoyer, B., Tzourio, C., & Dufouil, C. (2011).  
663 Longitudinal neuroimaging correlates of subjective memory impairment: 4-year  
664 prospective community study. *Br J Psychiatry*, 198(3), 199-205.  
665 doi:10.1192/bjp.bp.110.078683
- 666 Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., . . . Jessen, F. (2010).

- 667 Volume loss of the medial temporal lobe structures in subjective memory impairment.  
668 *Dement Geriatr Cogn Disord*, 29(1), 75-81. doi:10.1159/000264630
- 669 Tijms, Betty M., ten Kate, Mara, Gouw, Alida A., Borta, Andreas, Verfaillie, Sander, Teunissen,  
670 Charlotte E., . . . van der Flier, Wiesje M. (2018). Gray matter networks and clinical  
671 progression in subjects with predementia Alzheimer's disease. *Neurobiology of Aging*, 61,  
672 75-81. doi:<https://doi.org/10.1016/j.neurobiolaging.2017.09.011>
- 673 Tondelli, M., Wilcock, G. K., Nichelli, P., De Jager, C. A., Jenkinson, M., & Zamboni, G. (2012).  
674 Structural MRI changes detectable up to ten years before clinical Alzheimer's disease.  
675 *Neurobiology of Aging*, 33(4). doi:ARTN 825.e25  
676 10.1016/j.neurobiolaging.2011.05.018
- 677 Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature  
678 selection: Introduction and review. *J Biomed Inform*, 85, 189-203.  
679 doi:10.1016/j.jbi.2018.07.014
- 680 Verfaillie, Sander CJ, Tijms, Betty, Versteeg, Adriaan, Benedictus, Marije R, Bouwman, Femke H,  
681 Scheltens, Philip, . . . van der Flier, Wiesje M. (2016). Thinner temporal and parietal cortex  
682 is related to incident clinical progression to dementia in patients with subjective cognitive  
683 decline. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 5, 43-52.
- 684 Wang, Yang, Risacher, Shannon L, West, John D, McDonald, Brenna C, MaGee, Tamiko R, Farlow,  
685 Martin R, . . . Saykin, Andrew J. (2013). Altered default mode network connectivity in older  
686 adults with cognitive complaints and amnesic mild cognitive impairment. *Journal of*  
687 *Alzheimer's disease*, 35(4), 751-760.
- 688 Westman, Eric, Aguilar, Carlos, Muehlboeck, J. Sebastian, & Simmons, Andrew. (2012). Regional  
689 Magnetic Resonance Imaging Measures for Multivariate Analysis in Alzheimer's Disease  
690 and Mild Cognitive Impairment. *Brain Topogr*, 26(1), 9-23. doi:10.1007/s10548-012-  
691 0246-x
- 692 Wolfgruber, Steffen, Jessen, Frank, Koppara, Alexander, Kleineidam, Luca, Schmidtke, Klaus,  
693 Frölich, Lutz, . . . Heuser, Isabella. (2015). Subjective cognitive decline is related to CSF  
694 biomarkers of AD in patients with MCI. *Neurology*, 84(12), 1261-1268.
- 695 Xiao, S., Lewis, M., Mellor, D., McCabe, M., Byrne, L., Wang, T., . . . Dong, S. (2016). The China  
696 longitudinal ageing study: overview of the demographic, psychosocial and cognitive data  
697 of the Shanghai sample. *J Ment Health*, 25(2), 131-136.  
698 doi:10.3109/09638237.2015.1124385
- 699 Xiao, S., Li, J., Tang, M., Chen, W., Bao, F., Wang, H., . . . Zhang, M. (2013). Methodology of China's  
700 national study on the evaluation, early recognition, and treatment of psychological  
701 problems in the elderly: the China Longitudinal Aging Study (CLAS). *Shanghai Arch*  
702 *Psychiatry*, 25(2), 91-98. doi:10.3969/j.issn.1002-0829.2013.02.005
- 703 Yue, L., Wang, T., Wang, J., Li, G., Wang, J., Li, X., . . . Xiao, S. (2018). Asymmetry of Hippocampus  
704 and Amygdala Defect in Subjective Cognitive Decline Among the Community Dwelling  
705 Chinese. *Frontiers in psychiatry*, 9, 226.  
706  
707



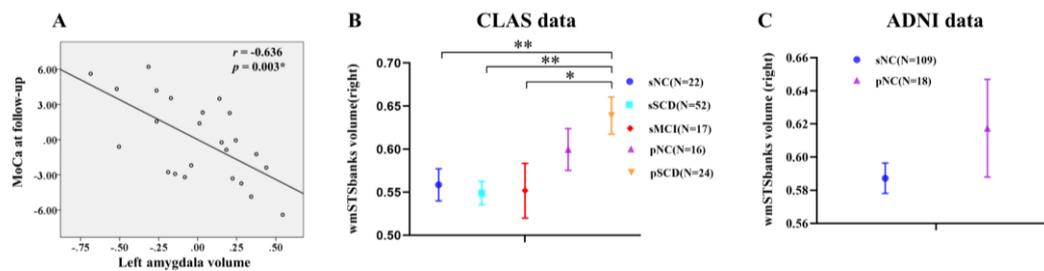
708

709

**Figure 1. The five most contributive features**

710

711

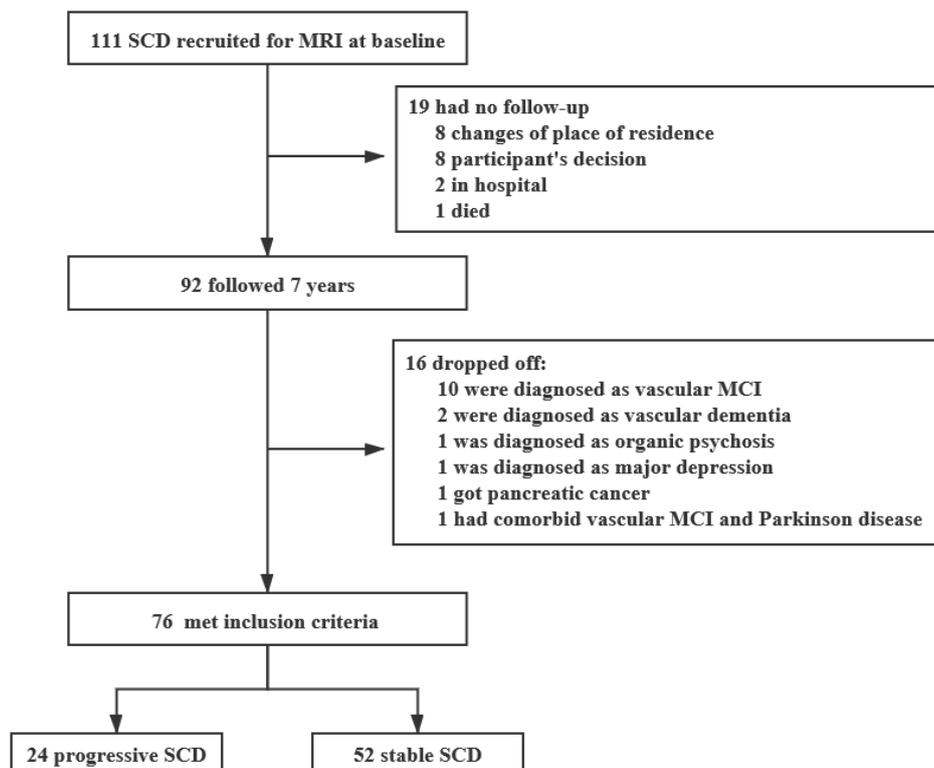


712

713

**Figure 2 | Clinical correlation, and validations based on CLAS study and ADNI**

714



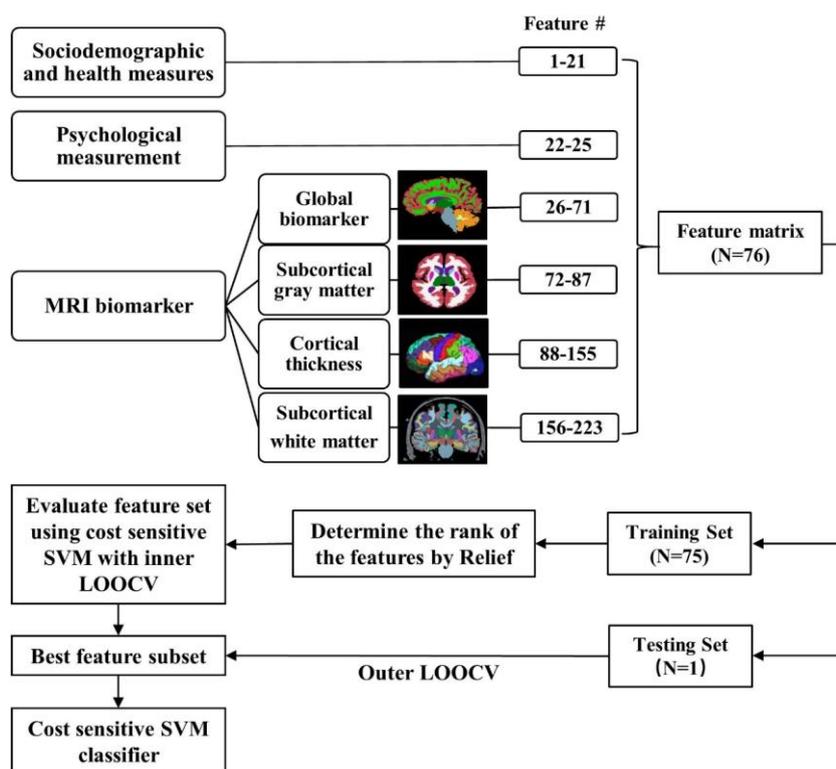
715

716

**Figure 3 | Subject selection flowchart**

717

718



719

720

**Figure 4 | Schematic flowchart of the SCD-to-MCI prediction framework based on the features extracted from sMRI and clinical measurement**

721

722