Supplementary material

Self-report and objective measures of swallowing capture different constructs in community-dwelling older people: a cross-classification study across two surveys

Contents

Supplementary Figure S1. Calibration plot for the temporal validation cohort

Supplementary Figure S2. Nomogram for the five-predictor objective model

Supplementary Figure S3. Exploratory decision curves

Supplementary Table S1. Full baseline characteristics

Supplementary Table S2. Comparison of modelling approaches

Supplementary Table S3. Univariable discrimination of candidate predictors

Supplementary Table S4. Decision-curve analysis

Supplementary Table S5. Agreement by cognitive stratum

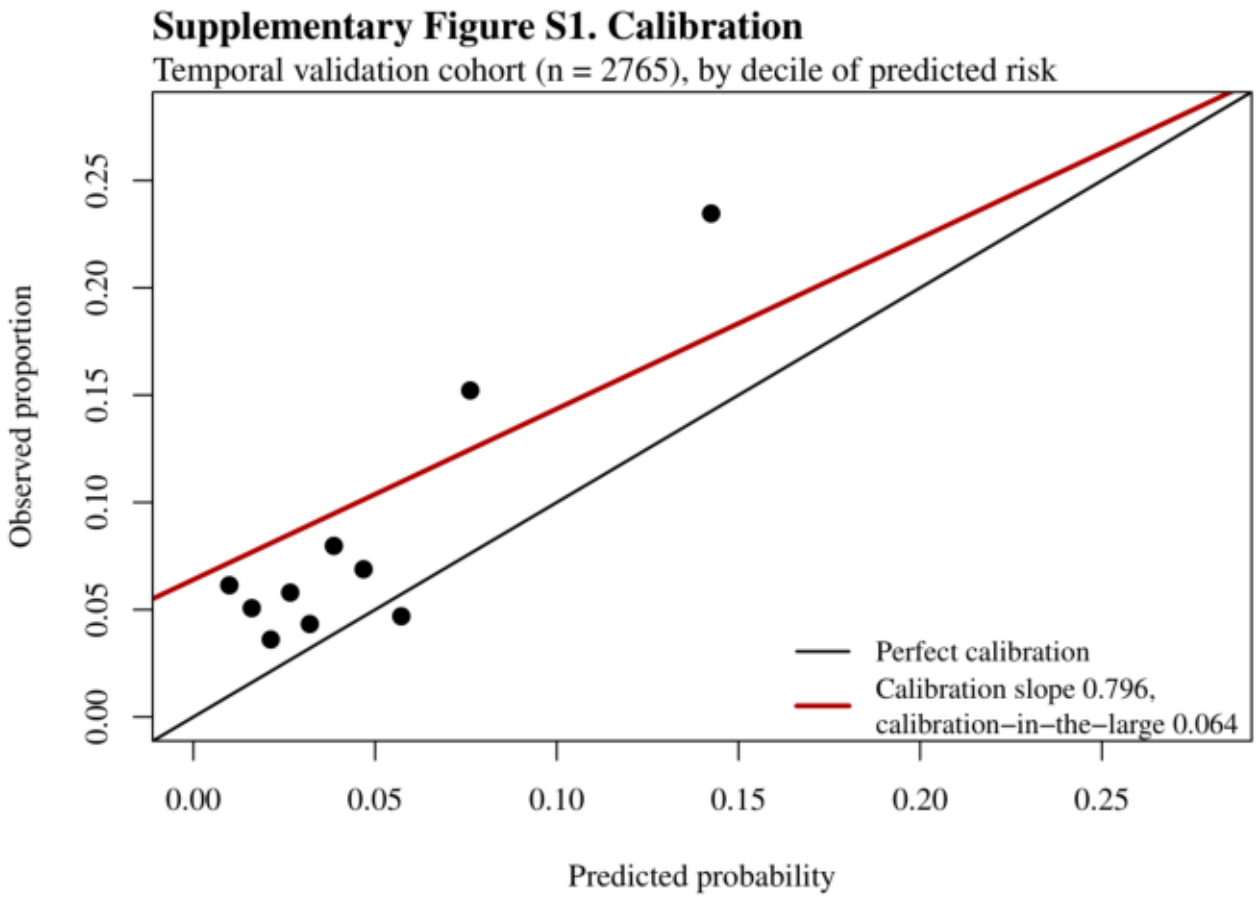
Supplementary Table S6. Sensitivity analysis: objective oral moisture value

Supplementary Table S7. Resource use and yield of candidate screening strategies

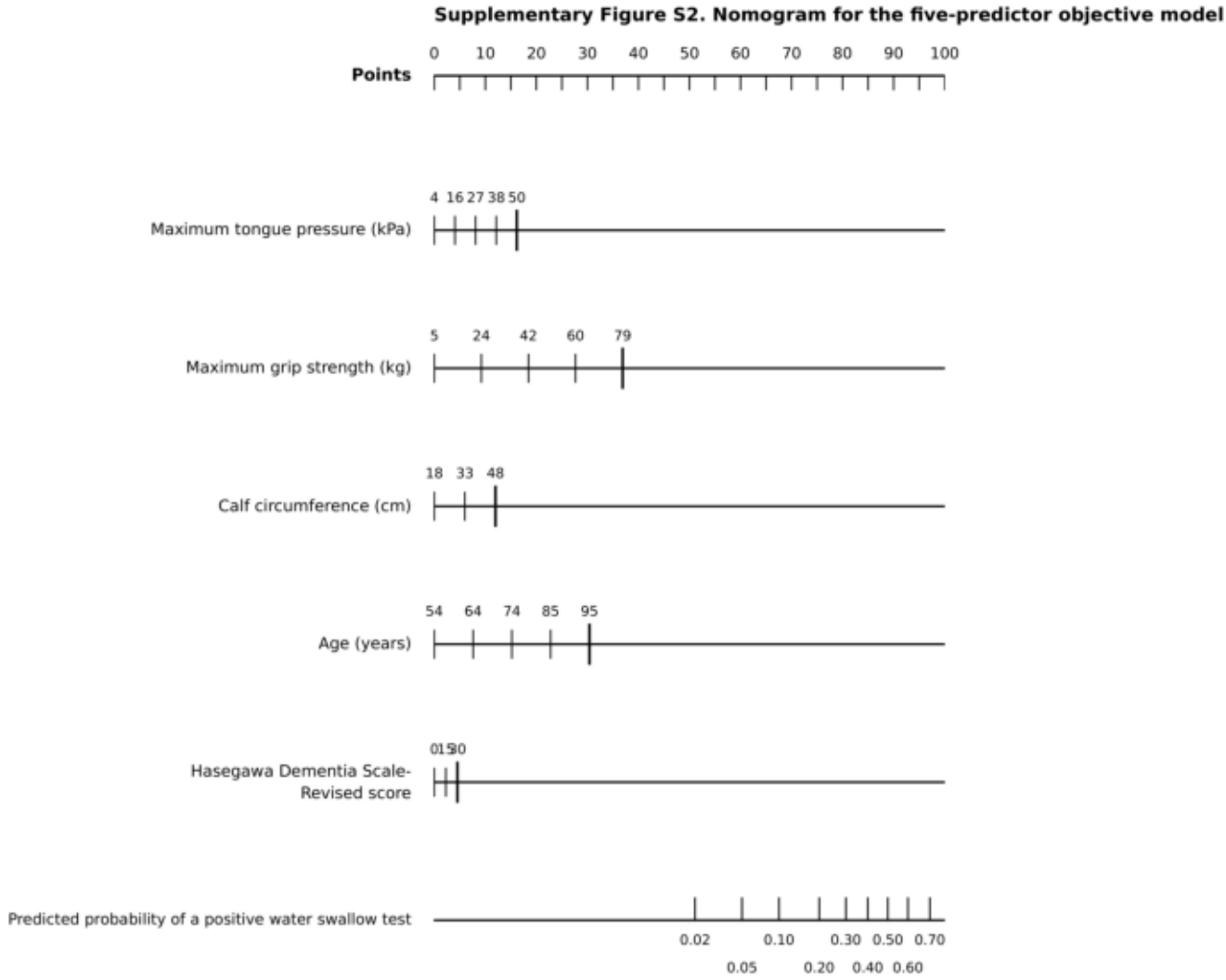
Supplementary Table S8. Characteristics of participants with and without a recorded cognitive score

Supplementary Table S9. Positive and negative percent agreement at candidate thresholds

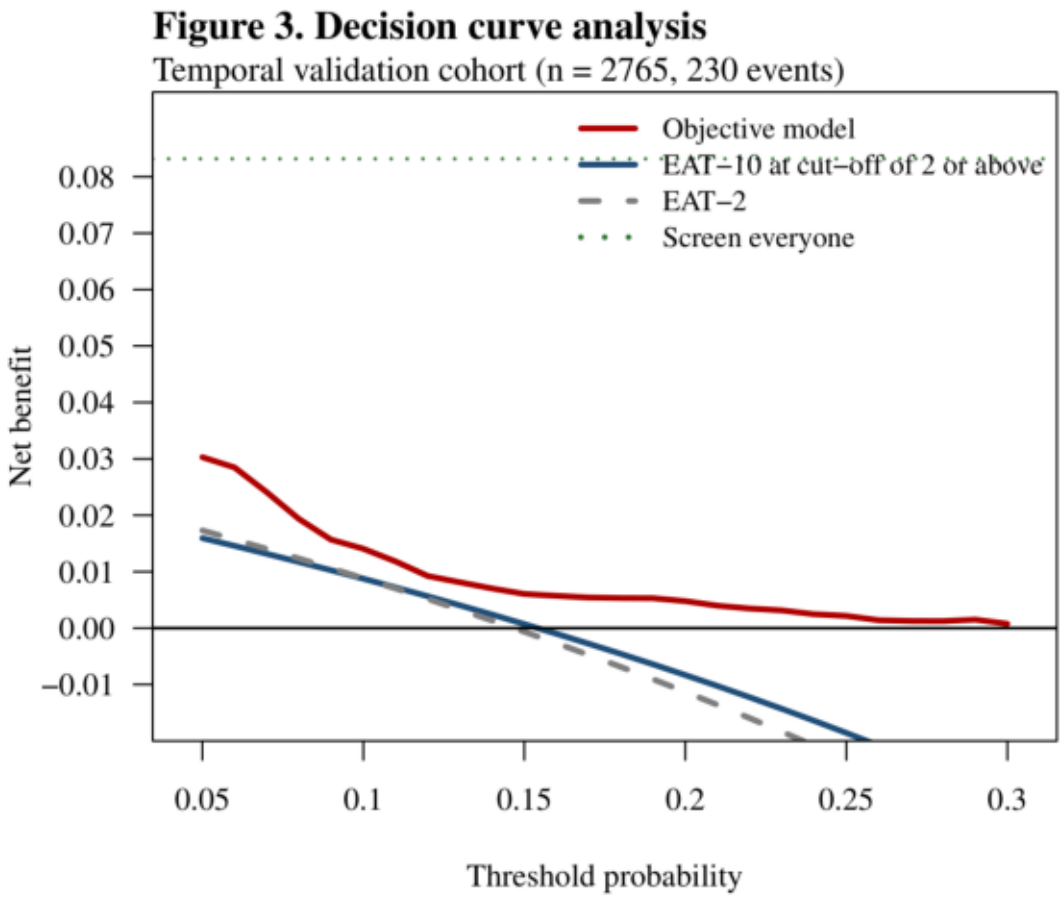
Supplementary Appendix. Model equation, package versions, random seed



Supplementary Figure S1. Calibration plot for the temporal validation cohort, by decile of predicted risk.



Supplementary Figure S2. Nomogram for the five-predictor objective model.



Supplementary Figure S3. Exploratory decision curves for the objective model, the Eating Assessment Tool at its cut-off and the two-item version, with reference strategies of testing everyone and testing nobody. Net benefit was computed with the water swallow test as the outcome; the horizontal line at the population prevalence represents testing everyone, and no strategy exceeded it at any threshold examined.

Supplementary Table S1. Full baseline characteristics of participants, by cognitive status and survey year

Characteristic	Presentation	All participants (n = 6,900)	Missing, %
Age, years	mean (SD)	73.4 (5.8)	0.0
Female sex	n (%)	47.2 (3258)	0.0
Hasegawa Dementia Scale-Revised score	mean (SD)	25.8 (3.6)	13.6
Maximum tongue pressure, kPa	mean (SD)	24.1 (7.8)	4.0
Maximum grip strength, kg	mean (SD)	25.5 (7.9)	1.3
Calf circumference, cm	mean (SD)	33.8 (3.1)	1.6
Number of natural teeth	mean (SD)	1.7 (0.4)	1.6
Subjective dry mouth	n (%)	50.9 (3488)	0.7
Unable to eat hard food	n (%)	30.5 (2079)	1.1
Pneumonia in the past year	n (%)	97.3 (6633)	1.2
Falls in the past year	n (%)	86.3 (5883)	1.2
Eating Assessment Tool total score	mean (SD)	0.52 (1.42)	0.8
Positive Eating Assessment Tool (>=2)	n (%)	13.6 (928)	0.8
Positive water swallow test	n (%)	7.5 (515)	0.0

Supplementary Table S2. Comparison of modelling approaches by ten-fold cross-validation in the derivation cohort

Approach	Ten-fold cross-validated AUC	Events per fold (mean)
Logistic regression	0.73	12.1
Lasso regression	0.726	12.1
Random forest	0.682	12.1
Gradient boosting	0.669	12.1
Support vector machine	0.607	12.1

Supplementary Table S3. Univariable discrimination of all candidate predictors in the derivation cohort

Variable	n	Univariable AUC (95% CI)
Maximum grip strength, kg	3353	0.697 (0.682 to 0.713)
Age, years	3402	0.694 (0.679 to 0.710)
Maximum tongue pressure, kPa	3328	0.686 (0.671 to 0.702)
Calf circumference, cm	3398	0.624 (0.608 to 0.641)
Hasegawa Dementia Scale-Revised score	2955	0.611 (0.593 to 0.629)
Unable to eat hard food	3373	0.571 (0.554 to 0.588)
Number of natural teeth	3371	0.569 (0.552 to 0.586)
Falls in the past year	3372	0.546 (0.529 to 0.563)
Objective oral moisture value	3393	0.542 (0.486 to 0.598)
Subjective dry mouth	3388	0.505 (0.488 to 0.522)
Pneumonia in the past year	3373	0.499 (0.482 to 0.516)
Female sex	3402	0.415 (0.399 to 0.432)

Supplementary Table S4. Exploratory decision analysis: net benefit, positive tests identified and participants referr

Threshold probability	Objective model: net benefit	Objective model: events detected	Objective model: referred for assessment	EAT-10 at cut-off of 2 or above: net benefit	EAT-10 at cut-off of 2 or above: events detected	EAT-10 at cut-off of 2 or above: referred for assessment	EAT-2: net benefit	EAT-2: events detected	Treat all: net benefit	Treat none: net benefit
0.05	0.0303	123	868	0.016	62	402	0.0173	69	0.0832	0
0.06	0.0285	112	633	0.0146	62	402	0.0157	69	0.0832	0
0.07	0.0241	95	472	0.0132	62	402	0.014	69	0.0832	0
0.08	0.0194	79	372	0.0117	62	402	0.0123	69	0.0832	0
0.09	0.0157	65	284	0.0103	62	402	0.0106	69	0.0832	0
0.1	0.0141	57	220	0.0088	62	402	0.0088	69	0.0832	0
0.11	0.0118	49	181	0.0072	62	402	0.007	69	0.0832	0
0.12	0.0092	40	146	0.0057	62	402	0.0052	69	0.0832	0
0.13	0.0081	35	119	0.004	62	402	0.0033	69	0.0832	0
0.14	0.007	31	102	0.0024	62	402	0.0013	69	0.0832	0
0.15	0.0061	27	85	7e-04	62	402	-6e-04	69	0.0832	0
0.16	0.0057	25	73	-0.001	62	402	-0.0027	69	0.0832	0
0.17	0.0054	23	62	-0.0028	62	402	-0.0047	69	0.0832	0
0.18	0.0053	22	55	-0.0046	62	402	-0.0069	69	0.0832	0
0.19	0.0053	21	48	-0.0064	62	402	-0.0091	69	0.0832	0
0.2	0.0048	18	37	-0.0083	62	402	-0.0113	69	0.0832	0
0.21	0.004	15	30	-0.0103	62	402	-0.0136	69	0.0832	0
0.22	0.0035	13	25	-0.0123	62	402	-0.016	69	0.0832	0
0.23	0.0032	12	23	-0.0143	62	402	-0.0184	69	0.0832	0
0.24	0.0025	10	20	-0.0164	62	402	-0.0208	69	0.0832	0
0.25	0.0022	9	18	-0.0186	62	402	-0.0234	69	0.0832	0
0.26	0.0014	6	12	-0.0208	62	402	-0.026	69	0.0832	0
0.27	0.0013	5	9	-0.0231	62	402	-0.0287	69	0.0832	0
0.28	0.0012	5	9	-0.0254	62	402	-0.0314	69	0.0832	0
0.29	0.0015	5	7	-0.0278	62	402	-0.0343	69	0.0832	0
0.3	8e-04	3	5	-0.0303	62	402	-0.0372	69	0.0832	0

Supplementary Table S5. Agreement between the Eating Assessment Tool and the water swallow test by cognitive

Survey	Cognitive stratum (HDS-R score)	Participants, n	Positive comparator tests, n	Prevalence of positive water swallow test, %	Cohen's kappa	Observed agreement, %	Positive percent agreement, %	Negative percent agreement, %
2024	All participants	2886	121	4.2	0.022	85.4	15.7	88.4
2024	20 or below (cognitively impaired)	251	23	9.2	-0.011	78.9	13.0	85.5
2024	21 or above (intact cognition)	2635	98	3.7	0.025	86.0	16.3	88.7
2025	All participants	2765	230	8.3	0.101	81.6	27.0	86.6
2025	20 or below (cognitively impaired)	164	25	15.2	0.280	78.7	48.0	84.2
2025	21 or above (intact cognition)	2601	205	7.9	0.082	81.8	24.4	86.7
2025	20 or below	164	25	15.2	0.280	78.7	48.0	84.2
2025	21 to 26	1200	118	9.8	0.079	80.3	22.9	86.6
2025	27 or above	1401	87	6.2	0.084	83.1	26.4	86.8
2025	21 to 22	155	17	11.0	0.134	77.4	35.3	82.6
2025	23 to 24	282	34	12.1	0.089	81.9	17.6	90.7
2025	25 to 26	763	67	8.8	0.063	80.3	22.4	85.9

Supplementary Table S6. Sensitivity analysis adding the objective oral moisture value to the five-predictor model i

Item	Result
Cohort	2024 derivation cohort; participants with complete data on all six variables
Participants, n	2884
Outcome events, n	121 (4.2%)
Oral moisture value	Median 31.4 (interquartile range 30.0 to 32.9); observed range 8.5 to 49.7; missing in 9 of 2,893
Moisture, univariable AUC	0.542 (95% CI 0.486 to 0.598)
Five-predictor model, AUC	0.750 (95% CI 0.709 to 0.792)
Six-predictor model (five plus moisture), AUC	0.753 (95% CI 0.710 to 0.795)
Difference in AUC	0.003
DeLong test (paired)	D = -0.417, P = 0.677
Adjusted coefficient for moisture	-0.065 (SE 0.026), P = 0.013
Optimism-corrected AUC	0.729 (five-predictor); 0.729 (six-predictor)
Calibration	Not assessed; no improvement in discrimination to justify further examination
Temporal validation	Not possible; the oral moisture value was not measured in the 2025 survey
Interpretation	Adding the objective oral moisture value did not improve discrimination or optimism-corrected discrimination; the measurement was not retained

Supplementary Table S7. Resource use and yield of candidate screening strategies in the temporal validation coho

Strategy	Screened, %	Positive percent agreement, %	Screened per 1,000	Screened per positive test
All participants (universal screening)	100	100	1000	12
Questionnaire positive only (consensus pathway)	14.5	27	145	6.5
Either questionnaire positive	18.3	32.6	183	6.7
Model positive (threshold = prevalence)	12.5	32.6	125	4.6
Model or questionnaire positive (parallel)	24.7	49.1	247	6.1
Model threshold 0.02	77.3	86.1	773	10.8
Model threshold 0.04	43	61.3	430	8.4
Model threshold 0.06	22.9	48.7	229	5.7
Model threshold 0.08	13.5	34.3	135	4.7
Model threshold 0.10	8	24.8	80	3.9
Model threshold 0.12	5.3	17.4	53	3.6
Model threshold 0.14	3.7	13.5	37	3.3
Model threshold 0.16	2.6	10.9	26	2.9
Model threshold 0.18	2	9.6	20	2.5
Model threshold 0.20	1.3	7.8	13	2.1
Model threshold 0.22	0.9	5.7	9	1.9
Model threshold 0.24	0.7	4.3	7	2
Model threshold 0.26	0.4	2.6	4	2
Model threshold 0.28	0.3	2.2	3	1.8
Model threshold 0.30	0.2	1.3	2	1.7

Supplementary Table S8. Characteristics of participants with and without a recorded Hasegawa Dementia Scale-Re

hdsr_status	n	Age, mean (SD)	Female, %	Tongue pressure, mean (SD)	Hip strength, mean (SD)	Calf circumference, mean (SD)	Positive WST, %	Survey
HDS-R missing	447	75.4 (7.0)	56.4	22.5 (9.4)	22.8 (8.7)	33.3 (3.1)	12.1	2024
HDS-R recorded	2955	72.6 (5.5)	52.5	26.3 (8.2)	25.5 (7.9)	34.0 (2.9)	4.4	2024
HDS-R missing	490	75.2 (6.4)	55.9	21.0 (6.4)	24.0 (8.0)	33.4 (3.3)	14.9	2025
HDS-R recorded	3008	73.7 (5.5)	52.1	22.6 (6.6)	26.1 (7.7)	33.8 (3.3)	8.6	2025

Supplementary Table S9. Positive and negative percent agreement for the objective model at candidate thresholds

Threshold source	Threshold	Participants screened, %	Positive percent agreement, %	Negative percent agreement, %
Selected in derivation cohort, applied unchanged	0.026	67.0	81.2	34.3
Selected within validation cohort	0.062	21.4	49.1	81.1
Fixed	0.040	43.1	62.0	58.6
Fixed	0.060	23.0	49.1	79.4
Fixed	0.080	13.5	35.0	88.4
Fixed	0.100	8.1	25.2	93.5
Fixed	0.150	3.1	12.4	97.7
Fixed	0.200	1.4	8.1	99.3

Supplementary Appendix

Model equation (linear predictor of the five-predictor objective model)

logit(p) = -4.970571 - 0.036224 x maximum tongue pressure (kPa)

- 0.051393 x maximum grip strength (kg)

- 0.041185 x calf circumference (cm)

+ 0.076424 x age (years)

- 0.015550 x Hasegawa Dementia Scale-Revised score

Predicted probability p = 1 / (1 + exp(-logit(p))).

Model performance in the temporal validation cohort (n = 2,765; 230 events)

Area under the curve	0.661 (95% CI 0.620 to 0.702)
Calibration-in-the-large	0.064 (95% CI -0.54 to 0.66)
Calibration slope	0.796 (95% CI 0.61 to 1.00)
Brier score	0.074 before, 0.073 after logistic recalibration

Software and reproducibility

R version 4.6.1 (2026-06-24 ucrt).

Random seed: 20261012.

Packages: readxl, dplyr, pmsampsize, rms, mice, dcurves, pROC.

Complete analysis scripts are archived at [REPOSITORY, DOI].