Supplementary material

Shortened smell tests to predict alpha-synuclein seed amplification assay status. Supplementary Tables S1–S16, numbered in the order the main text cites them.

Abbreviations: Sens, sensitivity; Spec, specificity; Bal Acc, balanced accuracy, the mean of the male and female balanced accuracies (Sens and Spec are computed over all participants unless stated otherwise); PPV and NPV, positive and negative predictive value; AUC, area under the receiver operating characteristic curve; CI, confidence interval; IQR, interquartile range; iRBD, isolated REM-sleep behaviour disorder; PD, Parkinson's disease.

Supplementary Table S1: Complete cases only

Cohort	Test	n (complete)	Short test Sens	Short test Spec	Short test Bal Acc	40 items Sens	40 items Spec	40 items Bal Acc	AUC loss (95% CI)	DeLong p	AUC loss, all participants
Validation	9 odours	434	0.80	0.58	0.67	0.94	0.40	0.66	0.018 (−0.010 to 0.045)	0.20	0.016
iRBD	9 odours	289	0.68	0.84	0.78	0.74	0.84	0.81	0.045 (0.020 to 0.070)	<0.001	0.045
PD	9 odours	311	0.73	0.82	0.78	0.76	0.85	0.80	0.039 (−0.004 to 0.082)	0.073	0.031
Validation	13 odours	434	0.82	0.58	0.68	0.94	0.40	0.66	0.007 (−0.018 to 0.031)	0.59	0.005
iRBD	13 odours	289	0.62	0.88	0.78	0.74	0.84	0.81	0.032 (0.010 to 0.055)	0.005	0.032
PD	13 odours	311	0.74	0.85	0.80	0.76	0.85	0.80	0.008 (−0.021 to 0.037)	0.59	0.011

Participants with any imputed UPSIT item are excluded from Validation (12), iRBD (0) and PD (131); the 17 Discovery participants with an imputed item are kept, so odours and cut-offs are those of the main analysis. AUC loss: 40-item AUC minus the short test's AUC, with DeLong p; the last column gives it for all participants (Supplementary Table S16).

Supplementary Table S2: Strictly baseline and discordant baseline αsyn-SAA results

Scenario	Test	Cohort	n	Cut-offs (male / female)	Sens	Spec	Bal Acc	AUC	AUC loss vs 40 items (95% CI)	DeLong p
main analysis	9 odours	Validation	446	5 / 6	0.80	0.60	0.69	0.75	0.016 (−0.010 to 0.043)	0.22
main analysis	9 odours	iRBD	289	5 / 6	0.68	0.84	0.78	0.83	0.045 (0.020 to 0.070)	<0.001
main analysis	9 odours	PD	442	5 / 6	0.72	0.87	0.79	0.87	0.031 (−0.002 to 0.065)	0.067
main analysis	13 odours	Validation	446	7 / 9	0.82	0.60	0.70	0.77	0.005 (−0.018 to 0.028)	0.66
main analysis	13 odours	iRBD	289	7 / 9	0.62	0.88	0.78	0.84	0.032 (0.010 to 0.055)	0.005
main analysis	13 odours	PD	442	7 / 9	0.71	0.89	0.81	0.89	0.011 (−0.012 to 0.034)	0.35
A: strictly baseline	9 odours	Validation	438	5 / 6	0.80	0.60	0.68	0.75	0.018 (−0.009 to 0.044)	0.19
A: strictly baseline	9 odours	iRBD	287	5 / 6	0.68	0.84	0.78	0.83	0.046 (0.020 to 0.071)	<0.001
A: strictly baseline	9 odours	PD	438	5 / 6	0.71	0.87	0.79	0.87	0.032 (−0.002 to 0.065)	0.064
A: strictly baseline	13 odours	Validation	438	7 / 9	0.82	0.60	0.70	0.77	0.004 (−0.019 to 0.027)	0.72
A: strictly baseline	13 odours	iRBD	287	7 / 9	0.62	0.87	0.78	0.84	0.033 (0.010 to 0.056)	0.005
A: strictly baseline	13 odours	PD	438	7 / 9	0.71	0.89	0.80	0.89	0.012 (−0.011 to 0.035)	0.32
B: discordant reversed	9 odours	Validation	446	5 / 6	0.80	0.60	0.69	0.75	0.016 (−0.010 to 0.043)	0.22
B: discordant reversed	9 odours	iRBD	289	5 / 6	0.68	0.85	0.79	0.84	0.044 (0.018 to 0.069)	<0.001
B: discordant reversed	9 odours	PD	442	5 / 6	0.72	0.87	0.79	0.87	0.031 (−0.002 to 0.065)	0.067
B: discordant reversed	13 odours	Validation	446	7 / 9	0.82	0.60	0.70	0.77	0.005 (−0.018 to 0.028)	0.66
B: discordant reversed	13 odours	iRBD	289	7 / 9	0.62	0.89	0.80	0.85	0.033 (0.011 to 0.056)	0.004
B: discordant reversed	13 odours	PD	442	7 / 9	0.71	0.89	0.81	0.89	0.011 (−0.012 to 0.034)	0.35
A+B	9 odours	Validation	438	5 / 6	0.80	0.60	0.68	0.75	0.018 (−0.009 to 0.044)	0.19
A+B	9 odours	iRBD	287	5 / 6	0.68	0.85	0.79	0.84	0.044 (0.018 to 0.070)	<0.001
A+B	9 odours	PD	438	5 / 6	0.71	0.87	0.79	0.87	0.032 (−0.002 to 0.065)	0.064
A+B	13 odours	Validation	438	7 / 9	0.82	0.60	0.70	0.77	0.004 (−0.019 to 0.027)	0.72
A+B	13 odours	iRBD	287	7 / 9	0.62	0.88	0.79	0.85	0.034 (0.011 to 0.057)	0.004
A+B	13 odours	PD	438	7 / 9	0.71	0.89	0.80	0.89	0.012 (−0.011 to 0.035)	0.32

Main analysis: αsyn-SAA status as in the rest of the paper. A: the 21 participants without a definitive baseline result, whose first later result (year 1 or 2) was used, are excluded (Discovery 7, of whom 6 in the balanced subset; Validation 8; iRBD 2; PD 4). B: the 3 participants with both a positive and a negative baseline result (Discovery 2, iRBD 1) are given the other result. A+B: both. In each scenario the cut-offs are re-fitted on the modified balanced Discovery subset; the odours are unchanged.

Supplementary Table S3: Test length vs performance across 30 re-splits (Figure 2 data)

Odours (n)	Runs	Discovery Bal Acc	Validation Bal Acc, median (IQR)	Validation Bal Acc, 95% range	iRBD Bal Acc	PD Bal Acc	AUC loss: Validation	AUC loss: iRBD	AUC loss: PD
1	60	0.66	0.64 (0.63–0.66)	0.58–0.69	0.69	0.75	0.111	0.165	0.160
2	60	0.68	0.64 (0.62–0.66)	0.59–0.68	0.71	0.76	0.081	0.122	0.111
3	60	0.71	0.66 (0.63–0.67)	0.61–0.71	0.73	0.76	0.060	0.085	0.073
4	60	0.72	0.67 (0.66–0.68)	0.63–0.69	0.73	0.77	0.051	0.064	0.061
5	60	0.74	0.67 (0.66–0.68)	0.63–0.71	0.76	0.77	0.040	0.049	0.048
6	49	0.75	0.68 (0.66–0.70)	0.63–0.71	0.76	0.79	0.036	0.046	0.037
7	52	0.75	0.68 (0.67–0.69)	0.63–0.72	0.76	0.78	0.029	0.037	0.031
8	49	0.76	0.69 (0.67–0.70)	0.65–0.73	0.77	0.78	0.023	0.032	0.025
9	**48**	**0.76**	**0.69 (0.68–0.70)**	**0.65–0.72**	**0.78**	**0.79**	**0.017**	**0.027**	**0.025**
10	40	0.76	0.69 (0.68–0.71)	0.66–0.73	0.78	0.80	0.016	0.026	0.018
11	26	0.77	0.69 (0.69–0.71)	0.66–0.73	0.77	0.80	0.013	0.024	0.011
12	36	0.77	0.70 (0.69–0.71)	0.66–0.73	0.78	0.80	0.004	0.021	0.014
13	30	0.77	0.71 (0.69–0.73)	0.68–0.74	0.78	0.80	−0.001	0.018	0.008
14	22	0.78	0.71 (0.70–0.73)	0.67–0.74	0.77	0.80	−0.001	0.017	0.007

The control cohort was re-split 30 times into Discovery and Validation (60/40, same stratified rule), and the genetic algorithm was run twice on each Discovery set (60 runs, main-analysis settings). Each run's best test per length was scored on its own Validation set and on the full iRBD and PD cohorts. Runs: number of runs that returned a test of that length; lengths returned by fewer than 20 runs are omitted. Values are medians across runs; for Validation also the IQR and the 95% range (2.5th to 97.5th percentile), as drawn in Figure 2A. AUC loss: 40-item AUC minus the test's AUC. Imputed item values are those of the main analysis.

Supplementary Table S3 (continued): odours in the 137 best tests of 8–10 odours (Figure 2B)

UPSIT item	Odour	Tests containing it (of 137)	%	In the 9-odour test
13	Licorice/Anise	113	82	yes
35	Grape Candy	106	77	yes
29	Apple	82	60	—
09	Leather	73	53	yes
10	Coconut	73	53	—
08	Clove	58	42	—
19	Chocolate	58	42	yes
14	Baby Powder	55	40	—
37	Soap	53	39	yes
05	Motor Oil	52	38	yes
40	Peanut	51	37	—
01	Pizza	48	35	—
07	Banana	45	33	yes
04	Cherry	41	30	—

UPSIT item	Odour	Tests containing it (of 137)	%	In the 9-odour test
02	Bubble Gum	40	29	yes
34	Pine	37	27	—
39	Rose	30	22	—
17	Strawberry	27	20	—
16	Gasoline	26	19	—
06	Mint	24	18	—
20	Wintergreen Gum	21	15	—
32	Grass	21	15	—
38	Natural Gas	20	15	—
26	Pineapple	19	14	yes
12	Raspberry	16	12	—
30	Watermelon	10	7	—
25	Dill Pickle	8	6	—
03	Menthol/Eucalyptus	6	4	—
22	Cotton Candy	3	2	—
28	Orange	2	1	—
24	Garlic	1	1	—
23	Peach	1	1	—
27	Rubber Tire	1	1	—
21	Flower	1	1	—
31	Sandalwood/Incense	1	1	—
15	Cinnamon	1	1	—

Each run's best test at 8, 9 and 10 odours (137 tests across the 60 runs), as drawn in Figure 2B. Never selected: Onion (11), Cedar (18), Smoke (33), Lemon (36).

Supplementary Table S4: Other comparators, including a 40-item test with a cut-off fitted on Discovery

Test	Cut-off ≤ male / female	Discovery Bal Acc	Validation Bal Acc	Validation Sens	Validation Spec	iRBD Bal Acc	iRBD Sens	iRBD Spec	PD Bal Acc	PD Sens	PD Spec
9 odours	5 / 6	0.78	0.69	0.80	0.60	0.78	0.68	0.84	0.79	0.72	0.87
13 odours	7 / 9	0.78	0.70	0.82	0.60	0.78	0.62	0.88	0.81	0.71	0.89
40 items, ≤15th percentile	normative	0.67	0.68	0.94	0.43	0.81	0.74	0.84	0.82	0.76	0.87
40 items, cut-off fitted on Discovery	22 / 28	0.70	0.71	0.83	0.60	0.79	0.63	0.90	0.78	0.67	0.89
Age and sex only (logistic regression)‡	—	0.53	0.59	0.54	0.64	0.62	0.59	0.65	0.47	0.39	0.54

Fitted cut-offs are sex-specific and maximise balanced accuracy on the balanced Discovery subset, as for the short tests (ties go to the lower cut-off). The AUC comparison with the 40-item test does not depend on the cut-off and is in Supplementary Table S16.

‡Logistic regression on age and sex, fitted on the balanced Discovery subset and positive at a predicted probability of 0.5 or more; its Sens and Spec are means of the male and female values.

Supplementary Table S5: Discrimination among normosmic and hyposmic participants

Cohort	Test	Participants	n	αsyn-SAA+ / −	Short test AUC (95% CI)	40 items AUC (95% CI)	AUC loss	DeLong p
Validation	9 odours	normosmic	88	18 / 70	0.81 (0.68–0.93)	0.81 (0.70–0.93)	0.007	0.86
Validation	9 odours	hyposmic	358	264 / 94	0.63 (0.56–0.70)	0.65 (0.58–0.72)	0.018	0.39
Validation	9 odours	all	446	282 / 164	0.75 (0.71–0.80)	0.77 (0.72–0.82)	0.016	0.22
iRBD	9 odours	normosmic	127	53 / 74	0.72 (0.63–0.81)	0.82 (0.75–0.90)	0.103	0.005
iRBD	9 odours	hyposmic	162	148 / 14	0.52 (0.39–0.64)	0.60 (0.42–0.77)	0.080	0.17
iRBD	9 odours	all	289	201 / 88	0.83 (0.78–0.88)	0.87 (0.83–0.92)	0.045	<0.001
PD	9 odours	normosmic	135	96 / 39	0.81 (0.73–0.90)	0.83 (0.75–0.90)	0.011	0.72
PD	9 odours	hyposmic	307	301 / 6	0.62 (0.35–0.88)	0.73 (0.50–0.96)	0.116	0.15
PD	9 odours	all	442	397 / 45	0.87 (0.80–0.94)	0.90 (0.85–0.95)	0.031	0.067
Validation	13 odours	normosmic	88	18 / 70	0.84 (0.73–0.95)	0.81 (0.70–0.93)	−0.029	0.38
Validation	13 odours	hyposmic	358	264 / 94	0.64 (0.57–0.71)	0.65 (0.58–0.72)	0.011	0.57
Validation	13 odours	all	446	282 / 164	0.77 (0.72–0.82)	0.77 (0.72–0.82)	0.005	0.66
iRBD	13 odours	normosmic	127	53 / 74	0.75 (0.66–0.83)	0.82 (0.75–0.90)	0.076	0.012
iRBD	13 odours	hyposmic	162	148 / 14	0.51 (0.33–0.69)	0.60 (0.42–0.77)	0.086	0.17
iRBD	13 odours	all	289	201 / 88	0.84 (0.79–0.90)	0.87 (0.83–0.92)	0.032	0.005
PD	13 odours	normosmic	135	96 / 39	0.81 (0.72–0.89)	0.83 (0.75–0.90)	0.018	0.47
PD	13 odours	hyposmic	307	301 / 6	0.75 (0.54–0.96)	0.73 (0.50–0.96)	−0.016	0.77
PD	13 odours	all	442	397 / 45	0.89 (0.84–0.94)	0.90 (0.85–0.95)	0.011	0.35

Normosmic: PPMI UPSIT percentile above 15; hyposmic: at or below 15. Scores (odours correctly identified; 40-item total) are compared by AUC, which does not depend on the cut-offs. AUC loss: 40-item AUC minus the short test's AUC. Confidence intervals and p values by DeLong's method.

Supplementary Table S6: Clinic vs remote administration of the UPSIT

Cohort	Test	Administration	n	αsyn-SAA+ / −	Short test Sens	Short test Spec	Short test AUC	40 items Sens	40 items Spec	40 items AUC	AUC loss	DeLong p
Validation	9 odours	all	446	282 / 164	0.80	0.60	0.75	0.94	0.43	0.77	0.016	0.22
Validation	9 odours	clinic	38	8 / 30	0.75	0.87	0.84	0.75	0.90	0.88	0.033	0.32
Validation	9 odours	remote	408	274 / 134	0.80	0.54	0.73	0.94	0.32	0.74	0.013	0.41
Validation	9 odours	enrolled as healthy controls (all clinic)	32	5 / 27	0.80	0.85	0.82	0.80	0.89	0.85	0.030	0.39
PD	9 odours	all (99.5% clinic)	442	397 / 45	0.72	0.87	0.87	0.76	0.87	0.90	0.031	0.067
iRBD	9 odours	all (96.9% clinic)	289	201 / 88	0.68	0.84	0.83	0.74	0.84	0.87	0.045	<0.001
Validation	13 odours	all	446	282 / 164	0.82	0.60	0.77	0.94	0.43	0.77	0.005	0.66
Validation	13 odours	clinic	38	8 / 30	0.63	0.90	0.86	0.75	0.90	0.88	0.017	0.46
Validation	13 odours	remote	408	274 / 134	0.82	0.53	0.74	0.94	0.32	0.74	0.004	0.76
Validation	13 odours	enrolled as healthy controls (all clinic)	32	5 / 27	0.60	0.89	0.84	0.80	0.89	0.85	0.007	0.76
PD	13 odours	all (99.5% clinic)	442	397 / 45	0.71	0.89	0.89	0.76	0.87	0.90	0.011	0.35
iRBD	13 odours	all (96.9% clinic)	289	201 / 88	0.62	0.88	0.84	0.74	0.84	0.87	0.032	0.005

Administration mode as recorded by PPMI. Hyposmia-enriched controls were mostly tested remotely and PD and iRBD participants almost all in clinic, so mode is confounded with cohort and, within the controls, with recruitment arm (all participants enrolled as healthy controls were tested in clinic; the other 6 clinic-tested Validation participants were enrolled for hyposmia). AUC loss: 40-item AUC minus the short test's AUC; p values by DeLong's method.

Supplementary Table S7: Cut-offs fitted on the full rather than the class-balanced Discovery cohort

Test	Cut-offs, balanced Discovery (male / female)	Cut-offs, full Discovery (male / female)	Validation Bal Acc, balanced fit	Validation Bal Acc, full fit
1 odour	0 / 0	0 / 0	0.66	0.66
2 odours	1 / 1	1 / 1	0.65	0.65
3 odours	1 / 2	1 / 2	0.69	0.69
4 odours	2 / 2	2 / 2	0.66	0.66
5 odours	2 / 2	2 / 2	0.65	0.65
6 odours	3 / 3	3 / 3	0.66	0.66
7 odours	3 / 4	3 / 4	0.70	0.70
8 odours	4 / 5	4 / 5	0.69	0.69
9 odours	5 / 6	5 / 6	0.69	0.69
10 odours	6 / 7	6 / 7	0.69	0.69
13 odours	7 / 9	7 / 9	0.70	0.70
14 odours	8 / 10	8 / 10	0.70	0.70
15 odours	8 / 10	8 / 10	0.71	0.71
16 odours	9 / 11	9 / 11	0.69	0.69
17 odours	10 / 12	10 / 12	0.69	0.69
40 items, cut-off fitted on Discovery	22 / 28	22 / 28	0.71	0.71

Odours unchanged; cut-offs re-fitted on all 671 Discovery participants (426 αsyn-SAA+ / 245 αsyn-SAA−) instead of the class-balanced subset of 438. iRBD and PD are not shown: with the odours fixed, their metrics depend only on the cut-offs.

Supplementary Table S8: Alternative cut-off rules, including age-adjusted cut-offs

Test	Cut-off rule	Cut-off (male / female)	Cohort	Sens	Spec	Bal Acc
9 odours	Balanced-accuracy optimum (main analysis)	5 / 6	Discovery†	0.84	0.68	0.78
9 odours	Balanced-accuracy optimum (main analysis)	5 / 6	Validation	0.80	0.60	0.69
9 odours	Balanced-accuracy optimum (main analysis)	5 / 6	iRBD	0.68	0.84	0.78
9 odours	Balanced-accuracy optimum (main analysis)	5 / 6	PD	0.72	0.87	0.79
9 odours	Target Discovery sensitivity ≥0.85	6 / 7	Discovery†	0.93	0.43	0.68
9 odours	Target Discovery sensitivity ≥0.85	6 / 7	Validation	0.92	0.46	0.68
9 odours	Target Discovery sensitivity ≥0.85	6 / 7	iRBD	0.79	0.75	0.78
9 odours	Target Discovery sensitivity ≥0.85	6 / 7	PD	0.87	0.80	0.85
9 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.70 / −0.14	Discovery†	0.82	0.70	0.78
9 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.70 / −0.14	Validation	0.72	0.63	0.67
9 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.70 / −0.14	iRBD	0.62	0.85	0.75

Test	Cut-off rule	Cut-off (male / female)	Cohort	Sens	Spec	Bal Acc
9 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.70 / −0.14	PD	0.65	0.89	0.78
9 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.29 / 0.16	Discovery†	0.85	0.59	0.75
9 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.29 / 0.16	Validation	0.80	0.55	0.67
9 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.29 / 0.16	iRBD	0.66	0.83	0.75
9 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.29 / 0.16	PD	0.72	0.87	0.80
13 odours	Balanced-accuracy optimum (main analysis)	7 / 9	Discovery†	0.86	0.67	0.78
13 odours	Balanced-accuracy optimum (main analysis)	7 / 9	Validation	0.82	0.60	0.70
13 odours	Balanced-accuracy optimum (main analysis)	7 / 9	iRBD	0.62	0.88	0.78
13 odours	Balanced-accuracy optimum (main analysis)	7 / 9	PD	0.71	0.89	0.81
13 odours	Target Discovery sensitivity ≥0.85	8 / 9	Discovery†	0.87	0.60	0.73
13 odours	Target Discovery sensitivity ≥0.85	8 / 9	Validation	0.88	0.58	0.71
13 odours	Target Discovery sensitivity ≥0.85	8 / 9	iRBD	0.73	0.84	0.81
13 odours	Target Discovery sensitivity ≥0.85	8 / 9	PD	0.80	0.80	0.81
13 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.61 / 0.05	Discovery†	0.86	0.67	0.78
13 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.61 / 0.05	Validation	0.80	0.59	0.69
13 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.61 / 0.05	iRBD	0.59	0.86	0.75
13 odours	Age-adjusted, balanced-accuracy optimum	z ≤ −0.61 / 0.05	PD	0.68	0.89	0.79
13 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.26 / 0.04	Discovery†	0.85	0.62	0.74
13 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.26 / 0.04	Validation	0.85	0.57	0.69
13 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.26 / 0.04	iRBD	0.69	0.84	0.77
13 odours	Age-adjusted, target sensitivity ≥0.85	z ≤ −0.26 / 0.04	PD	0.73	0.84	0.79

Odours fixed; each rule is fitted on the balanced Discovery subset and applied unchanged to the other cohorts. Target sensitivity: in each sex, the smallest cut-off that reaches the target Discovery sensitivity. Age-adjusted: scores converted to z-scores against αsyn-SAA-negative Discovery participants (linear in age, per sex), with a sex-specific z threshold. †Discovery figures are optimistic: the odours and cut-offs were chosen there.

Supplementary Table S9: Local search at each length from 1 to 20 odours

Odours (n)	Local search vs genetic algorithm (Discovery)	Genetic algorithm: Discovery Bal Acc	Local search: Discovery Bal Acc	Cut-off ≤ male / female	Validation Bal Acc	iRBD Bal Acc	PD Bal Acc	Odours (local search)
1	same test as the genetic algorithm	0.668	0.668	0 / 0	0.66	0.69	0.75	Licorice/Anise
2	same test as the genetic algorithm	0.684	0.684	1 / 1	0.65	0.71	0.78	Licorice/Anise, Grape Candy
3	same test as the genetic algorithm	0.713	0.713	1 / 2	0.69	0.71	0.77	Leather, Licorice/Anise, Grape Candy
4	same test as the genetic algorithm	0.733	0.733	2 / 2	0.66	0.73	0.77	Banana, Leather, Licorice/Anise, Grape Candy
5	same test as the genetic algorithm	0.747	0.747	2 / 2	0.65	0.72	0.73	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise
6	same test as the genetic algorithm	0.758	0.758	3 / 3	0.66	0.74	0.76	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Grape Candy
7	same test as the genetic algorithm	0.767	0.767	3 / 4	0.70	0.75	0.78	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Watermelon, Grape Candy
8	same test as the genetic algorithm	0.775	0.775	4 / 5	0.69	0.76	0.78	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Pineapple, Grape Candy, Soap
9	same test as the genetic algorithm	0.778	0.778	5 / 6	0.69	0.78	0.79	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Chocolate, Pineapple, Grape Candy, Soap
10	local search higher	0.770	0.775	6 / 7	0.70	0.79	0.81	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Chocolate, Pineapple, Pine, Grape Candy, Soap
11	no genetic-algorithm test of this length	—	0.780	6 / 7	0.67	0.78	0.80	Pizza, Motor Oil, Banana, Leather, Licorice/Anise, Cinnamon, Rubber Tire, Apple, Grape Candy, Soap, Rose
12	no genetic-algorithm test of this length	—	0.777	6 / 8	0.68	0.79	0.77	Bubble Gum, Motor Oil, Mint, Banana, Leather, Licorice/Anise, Gasoline, Chocolate, Pineapple, Grape Candy, Soap, Natural Gas
13	genetic algorithm higher	0.781	0.779	7 / 9	0.67	0.77	0.79	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Gasoline, Chocolate, Pineapple, Watermelon, Pine, Grape Candy, Soap, Peanut
14	equal, different test	0.780	0.780	8 / 10	0.69	0.80	0.79	Bubble Gum, Banana, Clove, Leather, Licorice/Anise, Gasoline, Chocolate, Dill Pickle, Pineapple, Watermelon, Sandalwood/Incense, Pine, Grape Candy, Soap
15	local search higher	0.779	0.780	8 / 11	0.70	0.78	0.81	Bubble Gum, Cherry, Motor Oil, Banana, Leather, Licorice/Anise, Baby Powder, Chocolate, Wintergreen Gum, Dill Pickle, Pineapple, Pine, Soap, Natural Gas, Peanut
16	genetic algorithm higher	0.781	0.779	9 / 11	0.71	0.81	0.77	Bubble Gum, Motor Oil, Mint, Banana, Clove, Leather, Licorice/Anise, Cinnamon, Gasoline, Chocolate, Wintergreen Gum, Garlic, Pineapple, Grape Candy, Soap, Peanut
17	local search higher	0.782	0.784	9 / 11	0.71	0.80	0.77	Bubble Gum, Motor Oil, Mint, Banana, Clove, Leather, Licorice/Anise, Baby Powder, Cinnamon, Gasoline, Chocolate, Pineapple, Grass, Grape Candy, Soap, Natural Gas, Peanut
18	no genetic-algorithm test of this length	—	0.781	10 / 13	0.69	0.80	0.80	Bubble Gum, Cherry, Motor Oil, Banana, Clove, Leather, Coconut, Licorice/Anise, Baby Powder, Gasoline, Wintergreen Gum, Garlic, Dill Pickle, Pine, Grape Candy, Soap, Natural Gas, Peanut

Odours (n)	Local search vs genetic algorithm (Discovery)	Genetic algorithm: Discovery Bal Acc	Local search: Discovery Bal Acc	Cut-off ≤ male / female	Validation Bal Acc	iRBD Bal Acc	PD Bal Acc	Odours (local search)
19	no genetic-algorithm test of this length	—	0.792	11 / 13	0.71	0.79	0.79	Pizza, Motor Oil, Banana, Clove, Leather, Coconut, Licorice/Anise, Baby Powder, Gasoline, Chocolate, Peach, Garlic, Pineapple, Rubber Tire, Grape Candy, Soap, Natural Gas, Rose, Peanut
20	no genetic-algorithm test of this length	—	0.778	11 / 14	0.68	0.82	0.80	Pizza, Motor Oil, Banana, Clove, Leather, Coconut, Raspberry, Licorice/Anise, Baby Powder, Gasoline, Garlic, Dill Pickle, Pineapple, Rubber Tire, Sandalwood/Incense, Grape Candy, Soap, Natural Gas, Rose, Peanut

For each length from 1 to 20 odours, a greedy start and 30 random starts were each improved by single-odour substitutions on the balanced Discovery subset, with the genetic algorithm's objective and cut-off rule; the best test per length is shown. The cut-offs, Validation, iRBD and PD Bal Acc and odours are those of the local-search test; nothing was selected on Validation, iRBD or PD.

Supplementary Table S9 (continued): all performance metrics for the genetic algorithm's and the local search's test at each length

Odours (n)	Test	Cohort	n	Cut-off ≤ male / female	Bal Acc	Sens	Spec	PPV	NPV	AUC	AUC loss vs 40 items
1	Both (same test)	Discovery†	438	0 / 0	0.67	0.74	0.59	0.65	0.70	0.67	0.053
1	Both (same test)	Validation	446	0 / 0	0.66	0.72	0.65	0.78	0.57	0.68	0.087
1	Both (same test)	iRBD	289	0 / 0	0.69	0.60	0.82	0.88	0.47	0.71	0.165
1	Both (same test)	PD	442	0 / 0	0.75	0.68	0.80	0.97	0.22	0.74	0.160
2	Both (same test)	Discovery†	438	1 / 1	0.68	0.82	0.56	0.65	0.75	0.69	0.034
2	Both (same test)	Validation	446	1 / 1	0.65	0.83	0.55	0.76	0.65	0.72	0.055
2	Both (same test)	iRBD	289	1 / 1	0.71	0.70	0.76	0.87	0.53	0.74	0.135
2	Both (same test)	PD	442	1 / 1	0.78	0.80	0.76	0.97	0.30	0.79	0.111
3	Both (same test)	Discovery†	438	1 / 2	0.71	0.84	0.58	0.66	0.78	0.73	−0.014
3	Both (same test)	Validation	446	1 / 2	0.69	0.79	0.56	0.76	0.61	0.72	0.053
3	Both (same test)	iRBD	289	1 / 2	0.71	0.55	0.86	0.90	0.46	0.77	0.108
3	Both (same test)	PD	442	1 / 2	0.77	0.66	0.84	0.97	0.22	0.83	0.073
4	Both (same test)	Discovery†	438	2 / 2	0.73	0.81	0.64	0.69	0.77	0.75	−0.029
4	Both (same test)	Validation	446	2 / 2	0.66	0.78	0.60	0.77	0.62	0.73	0.036
4	Both (same test)	iRBD	289	2 / 2	0.73	0.63	0.81	0.88	0.49	0.80	0.076
4	Both (same test)	PD	442	2 / 2	0.77	0.73	0.80	0.97	0.25	0.84	0.067
5	Both (same test)	Discovery†	438	2 / 2	0.75	0.67	0.78	0.75	0.70	0.76	−0.042
5	Both (same test)	Validation	446	2 / 2	0.65	0.61	0.72	0.79	0.52	0.73	0.043
5	Both (same test)	iRBD	289	2 / 2	0.72	0.52	0.89	0.91	0.45	0.82	0.051
5	Both (same test)	PD	442	2 / 2	0.73	0.57	0.89	0.98	0.19	0.86	0.044
6	Both (same test)	Discovery†	438	3 / 3	0.76	0.72	0.74	0.74	0.73	0.76	−0.045
6	Both (same test)	Validation	446	3 / 3	0.66	0.72	0.66	0.78	0.57	0.74	0.033
6	Both (same test)	iRBD	289	3 / 3	0.74	0.59	0.86	0.91	0.48	0.82	0.052
6	Both (same test)	PD	442	3 / 3	0.76	0.66	0.87	0.98	0.22	0.86	0.037

Odours (n)	Test	Cohort	n	Cut-off ≤ male / female	Bal Acc	Sens	Spec	PPV	NPV	AUC	AUC loss vs 40 items
7	Both (same test)	Discovery†	438	3 / 4	0.77	0.77	0.71	0.73	0.76	0.76	−0.043
7	Both (same test)	Validation	446	3 / 4	0.70	0.76	0.65	0.79	0.61	0.75	0.026
7	Both (same test)	iRBD	289	3 / 4	0.75	0.55	0.90	0.92	0.46	0.82	0.054
7	Both (same test)	PD	442	3 / 4	0.78	0.65	0.89	0.98	0.22	0.87	0.031
8	Both (same test)	Discovery†	438	4 / 5	0.78	0.81	0.69	0.73	0.79	0.76	−0.043
8	Both (same test)	Validation	446	4 / 5	0.69	0.77	0.63	0.78	0.61	0.75	0.024
8	Both (same test)	iRBD	289	4 / 5	0.76	0.60	0.86	0.91	0.49	0.83	0.049
8	Both (same test)	PD	442	4 / 5	0.78	0.68	0.87	0.98	0.23	0.87	0.035
9	Both (same test)	Discovery†	438	5 / 6	0.78	0.84	0.68	0.72	0.81	0.77	−0.046
9	Both (same test)	Validation	446	5 / 6	0.69	0.80	0.60	0.77	0.63	0.75	0.016
9	Both (same test)	iRBD	289	5 / 6	0.78	0.68	0.84	0.91	0.53	0.83	0.045
9	Both (same test)	PD	442	5 / 6	0.79	0.72	0.87	0.98	0.26	0.87	0.031
10	Genetic algorithm	Discovery†	438	6 / 7	0.77	0.83	0.69	0.73	0.80	0.75	−0.035
10	Genetic algorithm	Validation	446	6 / 7	0.69	0.78	0.61	0.78	0.62	0.75	0.017
10	Genetic algorithm	iRBD	289	6 / 7	0.77	0.61	0.88	0.92	0.50	0.83	0.045
10	Genetic algorithm	PD	442	6 / 7	0.76	0.66	0.87	0.98	0.22	0.87	0.029
10	Local search	Discovery†	438	6 / 7	0.78	0.86	0.65	0.71	0.82	0.76	−0.039
10	Local search	Validation	446	6 / 7	0.70	0.84	0.58	0.77	0.67	0.76	0.007
10	Local search	iRBD	289	6 / 7	0.79	0.72	0.83	0.91	0.57	0.83	0.047
10	Local search	PD	442	6 / 7	0.81	0.76	0.87	0.98	0.29	0.88	0.025
11	Local search	Discovery†	438	6 / 7	0.78	0.88	0.65	0.71	0.85	0.75	−0.028
11	Local search	Validation	446	6 / 7	0.67	0.87	0.52	0.76	0.70	0.76	0.015
11	Local search	iRBD	289	6 / 7	0.78	0.71	0.83	0.90	0.55	0.85	0.021
11	Local search	PD	442	6 / 7	0.80	0.73	0.84	0.98	0.26	0.88	0.018
12	Local search	Discovery†	438	6 / 8	0.78	0.83	0.69	0.73	0.80	0.76	−0.041
12	Local search	Validation	446	6 / 8	0.68	0.77	0.60	0.77	0.60	0.76	0.015
12	Local search	iRBD	289	6 / 8	0.79	0.60	0.90	0.93	0.49	0.85	0.023
12	Local search	PD	442	6 / 8	0.77	0.65	0.87	0.98	0.22	0.88	0.021
13	Genetic algorithm	Discovery†	438	7 / 9	0.78	0.86	0.67	0.72	0.82	0.75	−0.035
13	Genetic algorithm	Validation	446	7 / 9	0.70	0.82	0.60	0.78	0.65	0.77	0.005
13	Genetic algorithm	iRBD	289	7 / 9	0.78	0.62	0.88	0.92	0.50	0.84	0.032
13	Genetic algorithm	PD	442	7 / 9	0.81	0.71	0.89	0.98	0.26	0.89	0.011
13	Local search	Discovery†	438	7 / 9	0.78	0.83	0.68	0.72	0.80	0.75	−0.035
13	Local search	Validation	446	7 / 9	0.67	0.76	0.59	0.76	0.59	0.76	0.009
13	Local search	iRBD	289	7 / 9	0.77	0.58	0.89	0.92	0.48	0.84	0.038
13	Local search	PD	442	7 / 9	0.79	0.68	0.89	0.98	0.24	0.88	0.024
14	Genetic algorithm	Discovery†	438	8 / 10	0.78	0.87	0.65	0.71	0.84	0.76	−0.036
14	Genetic algorithm	Validation	446	8 / 10	0.70	0.84	0.58	0.77	0.68	0.77	0.001

Odours (n)	Test	Cohort	n	Cut-off ≤ male / female	Bal Acc	Sens	Spec	PPV	NPV	AUC	AUC loss vs 40 items
14	Genetic algorithm	iRBD	289	8 / 10	0.83	0.71	0.89	0.93	0.57	0.85	0.021
14	Genetic algorithm	PD	442	8 / 10	0.81	0.73	0.87	0.98	0.27	0.89	0.014
14	Local search	Discovery†	438	8 / 10	0.78	0.86	0.66	0.72	0.82	0.74	−0.024
14	Local search	Validation	446	8 / 10	0.69	0.81	0.58	0.77	0.64	0.76	0.009
14	Local search	iRBD	289	8 / 10	0.80	0.64	0.89	0.93	0.52	0.85	0.023
14	Local search	PD	442	8 / 10	0.79	0.72	0.87	0.98	0.26	0.88	0.018
15	Genetic algorithm	Discovery†	438	8 / 10	0.78	0.85	0.68	0.73	0.82	0.76	−0.037
15	Genetic algorithm	Validation	446	8 / 10	0.71	0.84	0.59	0.78	0.68	0.76	0.010
15	Genetic algorithm	iRBD	289	8 / 10	0.80	0.66	0.89	0.93	0.53	0.87	0.009
15	Genetic algorithm	PD	442	8 / 10	0.80	0.72	0.87	0.98	0.26	0.89	0.015
15	Local search	Discovery†	438	8 / 11	0.78	0.88	0.65	0.71	0.85	0.75	−0.033
15	Local search	Validation	446	8 / 11	0.70	0.84	0.57	0.77	0.67	0.77	−0.003
15	Local search	iRBD	289	8 / 11	0.78	0.62	0.85	0.91	0.49	0.85	0.020
15	Local search	PD	442	8 / 11	0.81	0.71	0.89	0.98	0.26	0.89	0.008
16	Genetic algorithm	Discovery†	438	9 / 11	0.78	0.84	0.69	0.73	0.81	0.75	−0.034
16	Genetic algorithm	Validation	446	9 / 11	0.69	0.78	0.60	0.77	0.62	0.77	0.002
16	Genetic algorithm	iRBD	289	9 / 11	0.80	0.65	0.88	0.92	0.52	0.85	0.022
16	Genetic algorithm	PD	442	9 / 11	0.79	0.69	0.89	0.98	0.25	0.90	0.006
16	Local search	Discovery†	438	9 / 11	0.78	0.84	0.69	0.73	0.81	0.75	−0.029
16	Local search	Validation	446	9 / 11	0.71	0.83	0.62	0.79	0.67	0.77	−0.003
16	Local search	iRBD	289	9 / 11	0.81	0.65	0.91	0.94	0.53	0.86	0.012
16	Local search	PD	442	9 / 11	0.77	0.67	0.87	0.98	0.23	0.88	0.020
17	Genetic algorithm	Discovery†	438	10 / 12	0.78	0.86	0.68	0.73	0.83	0.76	−0.036
17	Genetic algorithm	Validation	446	10 / 12	0.69	0.82	0.58	0.77	0.66	0.77	0.000
17	Genetic algorithm	iRBD	289	10 / 12	0.82	0.69	0.88	0.93	0.55	0.86	0.016
17	Genetic algorithm	PD	442	10 / 12	0.80	0.72	0.89	0.98	0.26	0.90	−0.002
17	Local search	Discovery†	438	9 / 11	0.78	0.84	0.69	0.73	0.81	0.76	−0.037
17	Local search	Validation	446	9 / 11	0.71	0.80	0.63	0.79	0.65	0.77	0.005
17	Local search	iRBD	289	9 / 11	0.80	0.63	0.90	0.93	0.51	0.87	0.009
17	Local search	PD	442	9 / 11	0.77	0.64	0.89	0.98	0.22	0.90	0.004
18	Local search	Discovery†	438	10 / 13	0.78	0.86	0.68	0.73	0.83	0.75	−0.027
18	Local search	Validation	446	10 / 13	0.69	0.84	0.57	0.77	0.67	0.78	−0.005
18	Local search	iRBD	289	10 / 13	0.80	0.65	0.88	0.92	0.52	0.86	0.011
18	Local search	PD	442	10 / 13	0.80	0.71	0.89	0.98	0.25	0.89	0.010
19	Local search	Discovery†	438	11 / 13	0.79	0.89	0.66	0.72	0.85	0.75	−0.034
19	Local search	Validation	446	11 / 13	0.71	0.86	0.58	0.78	0.70	0.76	0.008
19	Local search	iRBD	289	11 / 13	0.79	0.68	0.86	0.92	0.54	0.86	0.012
19	Local search	PD	442	11 / 13	0.79	0.70	0.89	0.98	0.25	0.89	0.017

Odours (n)	Test	Cohort	n	Cut-off ≤ male / female	Bal Acc	Sens	Spec	PPV	NPV	AUC	AUC loss vs 40 items
20	Local search	Discovery†	438	11 / 14	0.78	0.89	0.65	0.72	0.85	0.74	−0.024
20	Local search	Validation	446	11 / 14	0.68	0.82	0.55	0.76	0.64	0.77	0.005
20	Local search	iRBD	289	11 / 14	0.82	0.66	0.90	0.94	0.54	0.87	0.005
20	Local search	PD	442	11 / 14	0.80	0.69	0.89	0.98	0.25	0.89	0.015

Where both searches found the same test it is shown once; no genetic-algorithm run returned 11, 12 or 18–20 odours. AUC loss: 40-item AUC minus the test's AUC. †Discovery figures are optimistic: the odours and cut-offs were chosen there. PPV and NPV apply only at each cohort's own αsyn-SAA prevalence.

Supplementary Table S10: Odours excluded from the internationally portable test

UPSIT item	Odour	Reason for exclusion	Note
02	Bubble Gum	North-American confectionery flavour	In the 9-odour test
13	Licorice/Anise	North-American confectionery flavour; regional familiarity varies widely	In the 9-odour test
20	Wintergreen Gum	Wintergreen is largely a North-American flavouring	—
22	Cotton Candy	North-American confectionery flavour	—
25	Dill Pickle	North-American food preparation	—
35	Grape Candy	Artificial grape is a North-American confectionery flavour	In the 9-odour test

The list was fixed before the constrained search was run; the remaining 34 odours formed the candidate pool.

Supplementary Table S11: Best internationally portable test at each length

Odours (n)	Cut-off ≤ male / female	Discovery Bal Acc	Genetic-algorithm seeds that found it (of 10)	Odours
1	0 / 0	0.63	10	Banana
2	1 / 1	0.66	10	Cherry, Leather
3	1 / 1	0.69	10	Pizza, Banana, Leather
4	2 / 2	0.71	3	Leather, Garlic, Apple, Watermelon
5	2 / 3	0.73	6	Banana, Leather, Gasoline, Apple, Watermelon
6	3 / 4	0.74	1	Clove, Leather, Gasoline, Apple, Watermelon, Peanut
7	4 / 4	0.74	7	Pizza, Motor Oil, Banana, Leather, Pineapple, Apple, Soap
8	4 / 4	0.75	1	Pizza, Motor Oil, Banana, Leather, Baby Powder, Chocolate, Watermelon, Grass
9	5 / 6	0.75	2	Cherry, Motor Oil, Mint, Banana, Clove, Leather, Apple, Soap, Rose
10	5 / 6	0.75	2	Cherry, Motor Oil, Banana, Leather, Chocolate, Apple, Watermelon, Grass, Soap, Peanut
11	6 / 7	0.76	2	Cherry, Motor Oil, Banana, Leather, Chocolate, Apple, Watermelon, Grass, Pine, Soap, Peanut
12	6 / 7	0.76	1	Pizza, Motor Oil, Banana, Clove, Leather, Baby Powder, Gasoline, Pineapple, Watermelon, Grass, Rose, Peanut
13	7 / 8	0.76	1	Pizza, Motor Oil, Banana, Leather, Baby Powder, Gasoline, Chocolate, Pineapple, Watermelon, Grass, Soap, Rose, Peanut
14	8 / 9	0.76	1	Pizza, Motor Oil, Banana, Leather, Baby Powder, Gasoline, Chocolate, Garlic, Pineapple, Apple, Watermelon, Grass, Soap, Peanut
15	8 / 9	0.76	2	Pizza, Cherry, Motor Oil, Banana, Leather, Baby Powder, Cinnamon, Chocolate, Garlic, Pineapple, Apple, Soap, Natural Gas, Rose, Peanut
16	9 / 10	0.76	1	Pizza, Motor Oil, Banana, Clove, Leather, Baby Powder, Gasoline, Chocolate, Garlic, Pineapple, Rubber Tire, Watermelon, Soap, Natural Gas, Rose, Peanut

Genetic algorithm, settings and selection rule as in the main analysis (balanced Discovery subset, 219 αsyn-SAA+ / 219 αsyn-SAA−), with the candidate pool restricted to the 34 odours not excluded in Supplementary Table S10.

Supplementary Table S12: Internationally portable 9-odour test (Cherry, Motor Oil, Mint, Banana, Clove, Leather, Apple, Soap, Rose)

Estimates (95% CI); cut-offs ≤5 men / ≤6 women

Cohort	n	Bal Acc	Sens	Spec	PPV	NPV	AUC	9-odour test Bal Acc	9-odour test AUC
Discovery†	438	0.75 (0.71–0.79)	0.83 (0.77–0.87)	0.65 (0.58–0.71)	0.70 (0.66–0.74)	0.79 (0.74–0.84)	0.74	0.78	0.77
Validation	446	0.69 (0.64–0.74)	0.83 (0.78–0.87)	0.59 (0.51–0.66)	0.77 (0.74–0.81)	0.66 (0.60–0.73)	0.77	0.69	0.75
iRBD	289	0.84 (0.79–0.88)	0.75 (0.68–0.80)	0.85 (0.76–0.91)	0.92 (0.88–0.96)	0.60 (0.54–0.66)	0.87	0.78	0.83
PD	442	0.81 (0.75–0.85)	0.71 (0.66–0.75)	0.89 (0.77–0.95)	0.98 (0.97–1.00)	0.26 (0.22–0.29)	0.88	0.79	0.87

†Discovery figures are optimistic: the odours and cut-offs were chosen there. Intervals: Wilson score for Sens and Spec; stratified bootstrap (2,000 replicates, by αsyn-SAA status × sex) for Bal Acc, PPV and NPV. The last two columns give the 9-odour test for comparison. PPV and NPV apply only at each cohort's own αsyn-SAA prevalence.

Supplementary Table S13: Portable vs 9-odour test, same participants (main analysis)

Cohort	AUC difference, portable − 9-odour test (95% CI)	DeLong p
Discovery†	−0.027 (−0.057 to 0.002)	0.072
Validation	0.016 (−0.010 to 0.043)	0.23
iRBD	0.037 (0.009 to 0.066)	0.010
PD	0.012 (−0.022 to 0.046)	0.50

†Discovery figures are optimistic: the odours were chosen there. Positive differences favour the portable test; p values by DeLong's method. Results across 30 re-splits are in Supplementary Table S14.

Supplementary Table S14: Portable 9-odour test vs 9-odour test from all 40 odours, across 30 re-splits of the control cohort

Measure	All 40 odours, median (IQR)	Portable (34 odours), median (IQR)	Paired difference, portable − all 40 odours, median (IQR)	Re-splits favouring the portable test
Discovery Bal Acc	0.762 (0.754 to 0.773)	0.753 (0.742 to 0.758)	−0.0106 (−0.0168 to −0.0061)	1 / 30
Validation Bal Acc	0.692 (0.679 to 0.701)	0.688 (0.668 to 0.698)	−0.0069 (−0.0219 to 0.0112)	13 / 30
Validation AUC	0.756 (0.746 to 0.762)	0.759 (0.744 to 0.767)	−0.0004 (−0.0120 to 0.0181)	13 / 30
iRBD Bal Acc	0.783 (0.768 to 0.790)	0.779 (0.763 to 0.792)	−0.0001 (−0.0195 to 0.0164)	14 / 30
iRBD AUC	0.846 (0.841 to 0.856)	0.845 (0.838 to 0.854)	−0.0006 (−0.0112 to 0.0067)	13 / 30
PD Bal Acc	0.784 (0.777 to 0.801)	0.800 (0.781 to 0.818)	0.0137 (−0.0066 to 0.0320)	20 / 30
PD AUC	0.875 (0.867 to 0.884)	0.883 (0.877 to 0.892)	0.0070 (0.0001 to 0.0159)	22 / 30

The control cohort was re-split 30 times under the main stratified rule. In each re-split, the best 9-odour test from all 40 odours and the best portable 9-odour test (34 odours) were selected on the balanced Discovery subset by local search (a greedy start and 12 random starts, improved by single-odour substitutions; on the main-analysis split it selects the same portable test as the genetic algorithm) and scored on that re-split's Validation set and on the full iRBD and PD cohorts. Re-splits favouring the portable test: paired difference above zero. No re-split reached the iRBD AUC difference of the main analysis (Supplementary Table S13).

Supplementary Table S15: Odours in each test

Odours (n)	Odours	Genetic-algorithm seeds that found it (of 10)
1	Licorice/Anise	10
2	Licorice/Anise, Grape Candy	10
3	Leather, Licorice/Anise, Grape Candy	10
4	Banana, Leather, Licorice/Anise, Grape Candy	10
5	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise	10
6	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Grape Candy	10
7	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Watermelon, Grape Candy	6
8	Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Pineapple, Grape Candy, Soap	8
9	**Bubble Gum, Motor Oil, Banana, Leather, Licorice/Anise, Chocolate, Pineapple, Grape Candy, Soap**	**4**
10	Bubble Gum, Motor Oil, Leather, Licorice/Anise, Chocolate, Garlic, Grape Candy, Soap, Natural Gas, Peanut	1
13	**Bubble Gum, Cherry, Banana, Leather, Licorice/Anise, Baby Powder, Dill Pickle, Pineapple, Pine, Grape Candy, Soap, Natural Gas, Peanut**	**1**
14	Bubble Gum, Motor Oil, Mint, Banana, Clove, Leather, Licorice/Anise, Gasoline, Chocolate, Pineapple, Watermelon, Pine, Grape Candy, Soap	1
15	Bubble Gum, Banana, Clove, Leather, Coconut, Licorice/Anise, Baby Powder, Gasoline, Dill Pickle, Pineapple, Grape Candy, Soap, Natural Gas, Rose, Peanut	1
16	Bubble Gum, Motor Oil, Banana, Clove, Leather, Raspberry, Licorice/Anise, Baby Powder, Gasoline, Chocolate, Pineapple, Pine, Grape Candy, Soap, Natural Gas, Peanut	1
17	Bubble Gum, Motor Oil, Banana, Clove, Leather, Licorice/Anise, Baby Powder, Gasoline, Strawberry, Chocolate, Dill Pickle, Pineapple, Pine, Grape Candy, Soap, Natural Gas, Peanut	1

Odours are listed in UPSIT booklet order. Seeds = how many of the 10 genetic-algorithm runs reached this exact test. Tests found by local search are in Supplementary Table S9.

Supplementary Table S16: Comparison with 95% CIs and paired tests

Estimates (95% CI)

Cohort	Test	Bal Acc	Sens	Spec	PPV	NPV	AUC
Discovery†	9 odours	0.78 (0.74–0.82)	0.84 (0.78–0.88)	0.68 (0.62–0.74)	0.72 (0.69–0.76)	0.81 (0.76–0.85)	0.77
Discovery†	13 odours	0.78 (0.74–0.82)	0.86 (0.81–0.90)	0.67 (0.60–0.73)	0.72 (0.68–0.76)	0.82 (0.78–0.87)	0.75
Discovery†	40 items	0.67 (0.62–0.71)	0.93 (0.89–0.96)	0.39 (0.33–0.46)	0.61 (0.58–0.63)	0.85 (0.79–0.91)	0.72
Validation	9 odours	0.69 (0.63–0.73)	0.80 (0.75–0.84)	0.60 (0.52–0.67)	0.77 (0.74–0.81)	0.63 (0.57–0.69)	0.75
Validation	13 odours	0.70 (0.65–0.75)	0.82 (0.77–0.86)	0.60 (0.52–0.67)	0.78 (0.74–0.81)	0.65 (0.59–0.72)	0.77
Validation	40 items	0.68 (0.63–0.72)	0.94 (0.90–0.96)	0.43 (0.35–0.50)	0.74 (0.71–0.76)	0.80 (0.71–0.87)	0.77
iRBD	9 odours	0.78 (0.72–0.83)	0.68 (0.61–0.74)	0.84 (0.75–0.90)	0.91 (0.87–0.95)	0.53 (0.48–0.59)	0.83
iRBD	13 odours	0.78 (0.73–0.83)	0.62 (0.55–0.69)	0.88 (0.79–0.93)	0.92 (0.88–0.96)	0.50 (0.46–0.55)	0.84
iRBD	40 items	0.81 (0.76–0.86)	0.74 (0.67–0.79)	0.84 (0.75–0.90)	0.91 (0.88–0.95)	0.58 (0.53–0.65)	0.87
PD	9 odours	0.79 (0.73–0.85)	0.72 (0.67–0.76)	0.87 (0.74–0.94)	0.98 (0.96–0.99)	0.26 (0.22–0.30)	0.87
PD	13 odours	0.81 (0.75–0.86)	0.71 (0.66–0.75)	0.89 (0.77–0.95)	0.98 (0.97–1.00)	0.26 (0.22–0.30)	0.89
PD	40 items	0.82 (0.75–0.87)	0.76 (0.71–0.80)	0.87 (0.74–0.94)	0.98 (0.97–0.99)	0.29 (0.25–0.34)	0.90

Paired differences vs the 40-item UPSIT (Sens, Spec: short test minus 40 items; AUC loss: 40 items minus short test)

Cohort	Test	AUC loss (95% CI)	DeLong p	Sens difference (95% CI)	McNemar p (Sens)	Spec difference (95% CI)	McNemar p (Spec)
Validation	9 odours	0.016 (−0.010 to 0.043)	0.22	−0.138 (−0.184 to −0.092)	<0.001	0.171 (0.110 to 0.238)	<0.001
Validation	13 odours	0.005 (−0.018 to 0.028)	0.66	−0.121 (−0.170 to −0.074)	<0.001	0.171 (0.110 to 0.232)	<0.001
iRBD	9 odours	0.045 (0.020 to 0.070)	<0.001	−0.060 (−0.124 to 0.000)	0.073	0.000 (0.000 to 0.000)	1.00
iRBD	13 odours	0.032 (0.010 to 0.055)	0.005	−0.114 (−0.174 to −0.050)	<0.001	0.034 (0.000 to 0.080)	0.25
PD	9 odours	0.031 (−0.002 to 0.065)	0.067	−0.043 (−0.086 to −0.003)	0.060	0.000 (−0.089 to 0.089)	1.00
PD	13 odours	0.011 (−0.012 to 0.034)	0.35	−0.048 (−0.091 to −0.008)	0.032	0.022 (−0.067 to 0.111)	1.00

†Discovery figures are optimistic: the short tests' odours and cut-offs were chosen there. Intervals: Wilson score for Sens and Spec; stratified bootstrap (2,000 replicates, by αsyn-SAA status × sex) for Bal Acc, PPV, NPV and the Sens and Spec differences; DeLong for the AUC difference. PPV and NPV hold only at each cohort's own αsyn-SAA prevalence. In iRBD the 9-odour and 40-item tests classify all 88 αsyn-SAA-negative participants identically, hence the zero-width specificity interval. McNemar: exact two-sided test on discordant classifications.