Testing shortened versions of smell tests to screen for hyposmia in Parkinson's disease

Stephen D. Auger, MRCP, PhD,[1] Sofia Kanavou, MSc,[2] Michael Lawton, MSc,[2] Yoav Ben-Shlomo, MRCP, PhD, FFPH,[2] Michele T.Hu, FRCP, PhD,[3,4] Anette E. Schrag, FRCP, PhD,[5] Huw R. Morris, FRCP, PhD,[5] Donald G. Grosset, MD,[6] Alastair J. Noyce, MRCP, PhD,[1,7,†]

[1]Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

[2]Population Health Sciences, University of Bristol, Bristol, UK

[3]Oxford Parkinson's Disease Centre (OPDC), University of Oxford, Oxford, UK

[4]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[5]Department of Clinical and Movement Neuroscience, UCL Institute of Neurology, University College London, London, UK.

[6]Department of Neurology, Institute of Neurological Sciences, Queen Elizabeth University Hospital, Glasgow, United Kingdom

[7]Reta Lila Weston Institute and Department of Clinical and Movement Neuroscience, UCL Institute of Neurology, University College London, London, UK.

†Corresponding author:

Dr Alastair Noyce

Preventive Neurology Unit
Wolfson Institute of Preventive Medicine
Charterhouse Square, London, EC1M 6BQ
Telephone: +44 (0) 207 882 5841
Email: a.noyce@qmul.ac.uk

ABSTRACT

Background

Hyposmia is an early feature in neurodegenerative diseases, most notably Parkinson's Disease (PD). Using abbreviated smell tests could provide a cost-effective means for large-scale hyposmia screening. It is unclear whether short smell tests can effectively detect hyposmia in patient populations.

Objectives

To test the ability of short smell combinations to 'pre-screen' for probable hyposmia in people with PD and target administration of more extensive tests, such as the University of Pennsylvania Smell Identification Test (UPSIT).

Methods

We assessed the screening performance of a short 4 smell combination previously derived from use of the 40-item UPSIT in healthy older people and its ability to detect hyposmia in a large cohort of PD patients.

Results

The novel 4 smell combination included Menthol, Clove, Onion and Orange and had a sensitivity of 87.1% (95% confidence interval: 84.9%-89.2%) and specificity of 69.7% (63.3%-75.5%) for detecting hyposmia in patients with PD. A different (also novel) 4-item combination developed using a data-driven approach in PD patients only achieved 81.3% (78.2%-84.4%) sensitivity for equivalent specificity.

Conclusions

A short 4 smell combination derived from a healthy population demonstrated high sensitivity to detect those with hyposmia and PD.

Impaired olfaction (hyposmia) is an early feature of neurodegenerative diseases, most notably Parkinson's disease (PD)[1–6]. The University of Pennsylvania Smell Identification Test (UPSIT), comprising 40 "scratch-and-sniff" microencapsulated odorant strips, is commonly-used worldwide[7]. We previously identified short smell combinations derived from the 40-item UPSIT which could be cost-effective for large-scale hyposmia screening, before targeted administration of the UPSIT[8]. That work was conducted in the healthy general population recruited to the PREDICT-PD cohort study and was compared with only a small sample of patients with PD. Here we tested the screening performance of several short smell combinations in the *Tracking Parkinson's* study that includes data from 1,222 people with PD[9].

METHODS

Validating a short smell test in people with PD

The performance of 5 smells (Menthol, Clove, Onion, Gingerbread and Orange) achieved a balance between brevity and high performance for identifying people with hyposmia against the full 40-item UPSIT (sensitivity 94.1%) in the PREDICT-PD pilot cohort's healthy participants (n=891 for discovery and 191 for validation, as described previously[8]).

We assessed screening performance of a combination of 4 of these 5 smells, omitting Gingerbread, in *Tracking Parkinson's* 1222 patients. This was because participants in PREDICT-PD used the 40-item US version of UPSIT while *Tracking Parkinson's* cases completed the 40-item UK version of UPSIT. These two UPSIT versions are broadly similar, but 8 of the 40 smells differ to tailor smells for recognisability in certain populations. We prioritised cross-cultural smells (those appearing in both the US and UK UPSIT versions), which did not include Gingerbread. A positive screen for hyposmia using this 4-item combination was defined when 1 or more smells was identified incorrectly. Positive

and negative screens for hyposmia were compared with hyposmia defined by performance on the full 40-item test, using the same age- and gender-specific thresholds described previously[8].

Deriving a novel set of optimal smells in the PD cohort

We next ran a similar but abridged version of the full analysis we reported previously[8]. We considered all possible combinations of 4, 5 and 6 smells from the full 40 UPSIT smells, testing multiple different score thresholds for defining hyposmia (i.e. at least one, two, three etc. incorrectly identified smells to denote a positive hyposmia screen). These results were compared with participants' scores on the full 40-item test in a 'discovery' cohort, comprising a randomly selected 90% of the participants from *Tracking Parkinson's* (n=1100). The best performing combination of smells (defined by the highest sum of sensitivity and specificity) at each hyposmia threshold was tested in an independent 'validation' cohort, comprising the remaining 10% of participants (n=122). We considered different proportions of data for discovery/validation (including 50% for each, 75%/25%, 80%/20% and 95%/5%), but none achieved notably greater validation screening performance than 90% for discovery and 10% for validation. Screening performance of the best smell combinations is expressed as the values derived from the independent validation set.

Additional analyses

We next compared performance of the smell combinations used in the commercially available 4-item "Pocket Smell Test" to our novel combinations using the same methods. There are two commercially available 4-item "Pocket Smell Tests" (PST-A and PST-B); these are intended for use as a pre-screen for identifying individuals who require a full UPSIT. The PST-A includes chocolate, strawberry, smoke and leather, whereas PST-B includes grape, which does not feature in the UK UPSIT, and so was not further considered.

Finally, we assessed the ability of short smell combinations to distinguish people with PD (in *Tracking Parkinson's*) and those without (in PREDICT-PD). For this analysis, positive and negative screens using the same criteria described above, were compared with whether or not the individual had a diagnosis of PD (rather than hyposmia or not).

RESULTS

Validating a short smell test in people with PD

Based upon total UPSIT scores and age- and sex-specific thresholds, 80.9% of participants (988/1222) were classified as hyposmic in *Tracking Parkinson's*. At the time of olfactory testing, the PD patients

had a mean disease duration of 1.89 years (range 0.36 – 4.50 years). Table 1 shows the screening performance of the Menthol, Clove, Onion and Orange combination, where one or more incorrectly identified smell was used to as a positive screen for hyposmia. A screen was considered negative if participants identified all four smells correctly. The corresponding screening performance values (and 95% confidence intervals) of the 4-item combination were: sensitivity 87.1% (84.9%-89.2%), specificity 69.7% (63.3%-75.5%), PPV 92.3% (90.9%-93.7%), NPV 56.2% (51.7%-60.7%), positive likelihood ratio (LR+) 2.87 (2.36-3.49) and negative likelihood ratio (LR-) 0.18 (0.15-0.22).

Deriving a novel set of optimal smells in the PD cohort

The optimum combinations of 4, 5 or 6 smells, using different cut-offs to define hyposmia (i.e. ≤1, ≤2, ≤3 etc. correctly identified smells) as identified in a 'discovery' cohort of participants, is shown in Figure 1. None of these exceeded the sensitivity of the 4-item combination used in stage 1 (Figure 1), but some had higher PPV.

Additional analyses

For the commercially-available PST-A's combination of smells, the screening performance for detecting hyposmia on the UPSIT was: sensitivity 80.2% (77.5%-82.6%), specificity 72.6% (66.5%-78.3%), PPV 92.5% (90.9%-93.9%), NPV 46.4% (42.8%-50.1%), LR+ 2.93 (2.37-3.62), LR- 0.27 (0.24-0.32). Accordingly, the novel 4 smell combination from stage 1 possessed higher sensitivity for a similar specificity.

For detecting PD cases compared with controls, the screening performance of the Menthol, Clove, Onion, Orange combination was: sensitivity 76.3% (73.8%-78.6%), specificity 69.0% (66.2%-71.8%), PPV 73.6% (71.7%-75.4%), NPV 72.0% (69.8%-74.22%), LR+ 2.46 (2.24-2.71), LR- 0.34 (0.31-0.38). As a comparison, the PST-A smells had sensitivity 70.0% (67.4%-72.6%), specificity 60.1% (57.1%-63.0%), PPV 66.5% (64.6%-68.3%) and NPV 64.0% (61.7%-66.2%), LR+ 1.75 (1.62-1.90), LR- 0.50 (0.45-0.55); and the full 40-item UPSIT had sensitivity 80.9% (78.5%-83.0%), specificity 83.7% (81.4%-85.9%), PPV 84.9% (83.0%-86.6%), NPV 79.5% (77.5%-81.3%), LR+ 4.97 (4.33-5.71), LR- 0.23 (0.20-0.26).

DISCUSSION

An abbreviated combination of 4 smells from the UPSIT (Menthol, Clove, Onion and Orange) retained high sensitivity for identifying individuals with hyposmia in the context of PD. Previous work

demonstrated that these four smells, as well as Gingerbread, have high screening performance for detecting hyposmia in a general population and hinted toward similar screening performance in people with PD[8], an observation which is borne out in the current analyses. This combination of 4 smells also outperformed smells in the commercially available PST-A in differentiating PD cases from controls (on account of high prevalence of hyposmia in PD). Although it performed less well compared with the 40-item UPSIT, these data suggest abbreviated tests used as a screen in the pre-diagnostic phase of PD will not systematically miss a large proportion of those with PD-related olfactory dysfunction.

Our proposed application for an abbreviated smell test would be as part of a two-step approach, whereby individuals complete an abbreviated test as a quick and cost-effective 'pre-screen' to identify individuals who it might be pertinent to consider full smell testing for. Maintaining high sensitivity would be most important for an abbreviated smell test used in this way, to ensure that most people who require full testing are identified, while allowing more cost-effective, targeted administration of the full UPSIT. For example, sending a full UPSIT ($26.95 per test) to identify hyposmia in a cohort of 10,000 healthy older people would cost $269,500. Postage costs, based upon UK pricing, would amount to $4 return for each participant ($40,000 total). Using this 4-item pre-screen ($3.95 per test), 31.0% in PREDICT-PD screen positive and would also be sent a full UPSIT (total cost in tests $122,940). The two-step approach generates different postage costs too. The lighter weight of a 4-item test costs $2 return in the UK ($20,000 for 10,000 participants), with 31.0% incurring a second return postage cost of $4 for the full UPSIT ($12,400 for 10,000 participants). Hence, the total cost for non-discriminative UPSIT administration versus using this 4-item pre-screen amounts to $309,500 vs $155,340 respectively. Given the 4-item combination's 87.1% sensitivity for detecting hyposmia here, this halving of cost would come at the expense of missing approximately 10-15% of people in whom full UPSIT testing would have identified hyposmia.

The attempt to derive a novel "PD-specific" abbreviated smell combination (stage 2) was notable for the much greater variability in smells featuring in winning combinations than was present using the PREDICT-PD cohort's general population. This perhaps reflects greater variability and more erratic trends in olfactory dysfunction in PD cases than controls. Indeed, this is borne out in the fact that the 4 smells identified as performing best in PREDICT-PD's general population outperformed the corresponding 'winning' subset of 4 smells with hyposmia cutoff ≤3 derived from people with PD in every regard (sensitivity 87.1% vs 81.3%, specificity 69.7% vs 61.5%, PPV 92.3% vs 88.6% and NPV

6

56.2% vs 40.7% respectively). This adds further weight to the finding that the high-performing smells identified in PREDICT-PD's general population are a reliable marker for screening hyposmia.

Some limitations need to be noted. The high hyposmia prevalence (using UPSIT) in this cohort of people with PD (80.9%) is expected given the strong association between hyposmia and PD. In the PREDICT-PD general population, hyposmia prevalence using the same criteria was 16.2%. This warrants attention as high prevalence drives higher PPV and lower NPV. Indeed, as is clear in Figure 1, PPV was consistently higher than NPV for every smell combination in *Tracking Parkinson's* participants. Great care is necessary if attempting to extrapolate PPV/NPV to other populations, given their high context dependency. There are likely to be other sociodemographic differences between PREDICT-PD and *Tracking Parkinson's* participants additional to disease status. These potential confounders could have influenced test performance, but the observation that the original 4 smells out-performed the data driven approach suggests this was limited. The analysis also extrapolated performance of 4-item smells tests from 4 smell combinations taken within a larger 40-item smell test. There is a chance that true performance with dedicated 4-item tests may differ from the 4 smell combinations identified in this work.

Our four-smell combination retains a degree of cross-cultural relevance, given that all four smells feature in both the UK and US versions of the UPSIT. However, testing here was only in individuals based in the UK including a fairly restricted set of ethnicities. Further validation would be required including other ethnic minorities and in other countries to assess broader external validity. There also remain some unavoidable differences between the UK and US versions, most notably the use of different distractor options for some smells. In the analysis of screening performance for identifying PD cases vs controls, the cases (in *Tracking Parkinson's*) all completed the UK version of the UPSIT, whereas the controls (in PREDICT-PD) completed the US version. This could have impaired performance of controls versus cases given that it was a UK population completing a US test version. However, this would likely have led to underestimation rather than overestimation of screening performance owing to a reduced difference between cases and controls.

Prospective follow-up of these and similar cohorts will be able to provide additional information regarding possible clinical implications of having hyposmia in people already diagnosed with PD or whether disease duration is related to incidence of hyposmia.

7

In conclusion, the abbreviated 4 smell combination of Menthol, Clove, Onion and Orange retains high sensitivity to detect those with hyposmia in the context of PD. This subset of smells has cross-cultural relevance and out-performed attempts to derive a separate PD-specific combination of smells for hyposmia screening.

AUTHOR CONTRIBUTIONS

1. Research project: A. Conception, B. Organization, C. Execution; 2. Statistical Analysis: A. Design, B. Execution, C. Review and Critique; 3. Manuscript Preparation: A. Writing of the first draft, B. Review and Critique.

S.D.A.: 1A, 1C, 2A, 2B, 2C, 3A, 3B

S.K.: 1C, 3B

M.L.: 1B, 3B

Y.B-S.: 1B, 3B

M.T.H.: 1B, 3B

A.E.S.: 1B, 3B

H.R.M.: 1B, 3B

D.G.G.: 1A, 1B, 3B

A.J.N.: 1A, 1B, 1C, 2A, 2C, 3B

8

ETHICAL COMPLIANCE STATEMENT

Ethical approval was obtained from multicentre ethics committees and the relevant local research and development departments. Everyone participating in the studies provided written informed consent to have their data used for research upon enrolment. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this work is consistent with those guidelines.

REFERENCES

1.   Haehner A, Hummel T, Hummel C, Sommer U, Junghanns S, Reichmann H. Olfactory loss may be a first sign of idiopathic Parkinson's disease. *Mov Disord*. 2007;22(6):839-842. doi:10.1002/mds.21413

2.   Haehner A, Boesveldt S, Berendse HW, et al. Prevalence of smell loss in Parkinson's disease - a multicenter study. *Parkinsonism Relat Disord*. 2009;15(7):490-494. doi:10.1016/j.parkreldis.2008.12.005

3.   Nalls MA, McLean CY, Rick J, et al. Diagnosis of Parkinson's disease on the basis of clinical and

9

genetic classification: a population-based modelling study. *Lancet Neurol*. 2015;14(10):1002-1009. doi:10.1016/S1474-4422(15)00178-7

4.   Ross GW, Petrovitch H, Abbott RD, et al. Association of olfactory dysfunction with risk for future Parkinson's disease. *Ann Neurol*. 2008;63(2):167-173. doi:10.1002/ana.21291

5.   Siderowf A, Stern MB. Premotor Parkinson's disease: Clinical features, detection, and prospects for treatment. *Ann Neurol*. 2009;64(S2):S139-S147. doi:10.1002/ana.21462

6.   Siderowf A, Jennings D, Eberly S, et al. Impaired olfaction and other prodromal features in the Parkinson At-Risk Syndrome study. *Mov Disord*. 2012;27(3):406-412. doi:10.1002/mds.24892

7.   Doty RL, Shaman P, Dann M. Development of the university of pennsylvania smell identification test: A standardized microencapsulated test of olfactory function. *Physiol Behav*. 1984;32(3):489-502. doi:10.1016/0031-9384(84)90269-5

8.   Joseph T, Auger SD, Peress L, et al. Screening performance of abbreviated versions of the UPSIT smell test. *J Neurol*. 2019;266(8):1897-1906. doi:10.1007/s00415-019-09340-x

9.   Malek N, Swallow DMA, Grosset KA, et al. Tracking Parkinson's: Study Design and Baseline Patient Data. *J Parkinsons Dis*. 2015;5(4):947-959. doi:10.3233/JPD-150662

TABLE AND FIGURE LEGENDS

Table 1 – Screening performance of Menthol, Clove, Onion and Orange, using one or more incorrectly identified smell to define a positive screen, in 1222 people with recent onset Parkinson's disease.

Figure 1 – **Screening performance in the validation cohort for each 'winning' smell combination, at the different cut-offs used for defining hyposmia for subsets of 4 to 6 smells.** The 4 smell combination described in Stage 1 is included at the top for comparison. Cell shading corresponds to the value of the number they contain. Values below 50 contain no shading, darker shading corresponds to higher values (in brackets of 5). The highest value in each category is underlined. PPV = positive predictive value, NPV = negative predictive value.

Smells		Cut-off for hyposmia	Sensitivity	Specificity	PPV	NPV
4 smells	Clove, Menthol, Onion, Orange	≤3	87.1	69.7	92.3	56.6
4 smells	Cherry, Grapefruit, Liquorice, Pizza	≤1	72.9	76.9	92.1	43.5
	Coconut, Liquorice, Menthol, Peanut	≤2	77.1	84.6	94.9	50.0
	Clove, Leather, Peanut, Raspberry	≤3	81.3	61.5	88.6	47.1
5 smells	Cherry, Chutney, Grapefruit, Liquorice, Pizza	≤1	66.7	84.6	94.1	40.7
	Cherry, Coconut, Grass, Liquorice, Orange	≤2	72.9	92.3	97.2	48.0
	Clove, Leather, Liquorice, Orange, Peanut	≤3	76.0	76.9	92.4	46.5
	Baby powder, Clove, Leather, Peanut, Raspberry	≤4	84.4	57.7	88.0	50.0
6 smells	Cherry, Chutney, Grapefruit, Liquorice, Pizza, Turpentine	≤1	59.4	92.3	96.6	38.1
	Cherry, Clove, Coconut, Grapefruit, Liquorice, Turpentine	≤2	76.0	88.5	96.1	50.0
	Cherry, Coconut, Grass, Liquorice, Menthol, Orange	≤3	80.2	88.5	96.3	54.8
	Baby powder, Clove, Leather, Liquorice, Menthol, Orange	≤4	78.1	80.8	93.8	50.0
	Baby powder, Leather, Menthol, Peanut, Raspberry, Smoke	≤5	85.4	61.5	89.1	53.3

Table 1 – **Screening performance of Menthol, Clove, Onion and Orange, using one or more incorrectly identified smell to define a positive screen, in 1222 people with recent onset Parkinson's**.

	Positive screen	Negative screen	Total
Hyposmia	861 (87.1%)	127 (12.9%)	988
Normosmia	71 (30.3%)	163 (69.7%)	234
Total	932	290	1222