

Development and validation of odour identification tests for olfactory assessment in Singapore

Xinni Xu^{1,2} MMed, Margaret Ru Xiang Zhang^{1,3} MMed, Terese Huiying Low^{1,3} MMed, Yew Kwang Ong¹ MMed

ABSTRACT

Introduction: Odour recognition is influenced by culture. Odour identification tests need to be adapted to a population to accurately assess olfactory function. This study's objectives were to validate the Singapore version of the Sniffin' Sticks (SS-Sg) and a locally-developed odour recognition test (Scentsor) for Singapore.

Method: This prospective study was performed in 3 otolaryngology outpatient clinics in 3 phases (1 May to 15 November 2024). Phase 1 was a survey evaluation of 93 odour descriptors to identify familiar odour descriptors to be used in the tests (n=414); Phase 2 evaluated and finalised SS-Sg and Scentsor to ensure test odours were recognised by $\geq 75\%$ of healthy controls (n=130); and Phase 3 validated both tests on healthy controls (n=473) to obtain normative data, to determine test-retest reliability (n=50), and to assess the ability to distinguish patients with olfactory loss (n=67).

Results: In Phase 1, the unmodified SS blue and purple sets had 15/32 (46.9%) unfamiliar test odours and 25 unfamiliar distractors combined. In Phase 2, after modification, all odours in SS-Sg and Scentsor were correctly identified by $\geq 75\%$ of controls. In Phase 3, normative data (age 21–83 years) was obtained. Both tests had good test-retest reliability (Pearson's correlation coefficient of 0.88 with $P < 0.001$ for SS-Sg; and at 0.90 with $P < 0.001$ for Scentsor). Both tests differentiated among normosmia, hyposmia and anosmia (SS-Sg scores: 12.6 [± 2.4] versus [vs] 9.8 (± 3.2) vs 6.0 (± 2.3) respectively, $P < 0.001$; Scentsor scores: 14.3 [± 1.8] vs 11.3 [± 2.8] vs 5.8 [± 3.4] respectively, $P < 0.001$).

Conclusion: SS-Sg and Scentsor have been validated to assess olfaction in Singapore.

Keywords: odour identification, olfaction, olfactory assessment, otorhinolaryngology, Sniffin' Sticks

CLINICAL IMPACT

What is New

- This study is the first to develop and validate 2 odour identification tests: Sniffin' Sticks (SS-Sg) and Scentsor, for the Singapore population.
- Normative data were obtained for both tests.

Clinical Implications

- Odour identification tests are sensitive to cultural differences. The unmodified SS is not suitable to be used in Singapore due to the high number of unfamiliar odour descriptors.
- SS-Sg or Scentsor should be used in clinical and scientific assessment of quantitative olfactory function in Singapore.

INTRODUCTION

Olfactory dysfunction affects patients under the care of various medical specialties. Its causes include post-viral upper respiratory tract infections, such as that caused by the SARS-CoV-2 virus, post-traumatic

brain injury and neurodegenerative conditions such as Parkinson's disease. The prevalence of olfactory impairment among community-dwelling, older Singapore residents is estimated to be 34.0%,¹ which is higher than the global average of 26.6% reported in a meta-analysis of olfactory dysfunction in the general population.² In addition, olfactory impairment in this group of Singapore residents was shown to be associated with a 2-fold increased risk of cognitive impairment.¹

Accurate assessment and diagnosis of olfactory dysfunction requires olfactory testing with validated olfactory tests, especially as self-rated olfactory function can correlate poorly with psychophysical olfactory test results.^{3,4} Olfactory testing also allows for tracking and documentation of changes in olfactory function over time. Most clinical olfactory tests, especially screening tests, rely on the ability to identify odours among a few multiple-choice options. However, the recognition of an odour relies not only on the ability to detect the odour, but also on

The Annals is an open access journal, allowing non-commercial use under CC BY-NC-SA 4.0.

¹ Department of Otolaryngology—Head & Neck Surgery, National University Hospital, Singapore

² Department of Otolaryngology—Head & Neck Surgery, Alexandra Hospital, Singapore

³ Department of Otolaryngology—Head & Neck Surgery, Ng Teng Fong General Hospital, Singapore

Correspondence: Dr Xinni Xu, Department of Otolaryngology—Head & Neck Surgery, National University Hospital, 5 Lower Kent Ridge Road, Singapore 119074.

Email: xinni_xu@nuhs.edu.sg

Accepted: 24 March 2025

Published online first: 6 June 2025

familiarity with the test odour and distractor options. Certain odours may not be familiar to people who are not native to the culture or country where the test had been developed. Therefore, adapting odour identification tests and validating them in the test population to obtain normative data is required in order to interpret results accurately.⁵ This can be done by making modifications to odour descriptors in an internationally accepted psychophysical olfactory test, such as the odour identification component of the Sniffin' Sticks⁶ (SS) test⁷⁻¹¹ (Burghart Messtechnik GmbH, Holm, Germany). Alternatively, novel smell identification tests have also been developed using test odours that are familiar to the specific region.¹²⁻¹⁵

The primary objective of this study was to develop and validate smell identification testing for the Singapore population. Two types of smell identification tests were assessed. The first was the 16-item odour identification component of the SS that was adapted to Singapore residents (SS-Sg). The second was an original odour recognition test comprising 16 Singapore-sourced scents named Scentsor, an amalgam of "scents" and "odour recognition".

METHOD

This prospective study was conducted from 1 May to 15 November 2024. It was carried out in 3 phases (Supplemental Fig. S1), in keeping with that of several studies that developed or modified SS for specific populations.^{7-11,16} All participants were recruited from tertiary outpatient otolaryngology clinics from National University Hospital, Ng Teng Fong General Hospital and Alexandra Hospital. This study was approved by the National Healthcare Group Domain Specific Review Board and participants gave informed consent. All experimental procedures were consistent with the Declaration of Helsinki.

SS consists of odour threshold, discrimination and identification testing. For the purposes of this study, only the identification component of SS is used. The SS identification test consists of 16 felt-tip pens, each impregnated with an odour. To present the odour, the cap of the pen is removed and the pen tip is held about 2 cm in front of the participant's nostrils, alternately for about 3 seconds. The participant identifies the correct odour from 4 multiple-choice options in a forced choice manner. The total number of correct answers corresponds to the test score (minimum 0, maximum 16). There are 2 versions of the SS identification, generally referred to as the blue or purple version, based on the colour coding of the Stick. The blue and purple versions have different test odours. These 2 versions were developed so as to have an alternative set of odours for identification testing, especially when patients may be likely to recall the answers.

Phase 1: Determining odour recognisability

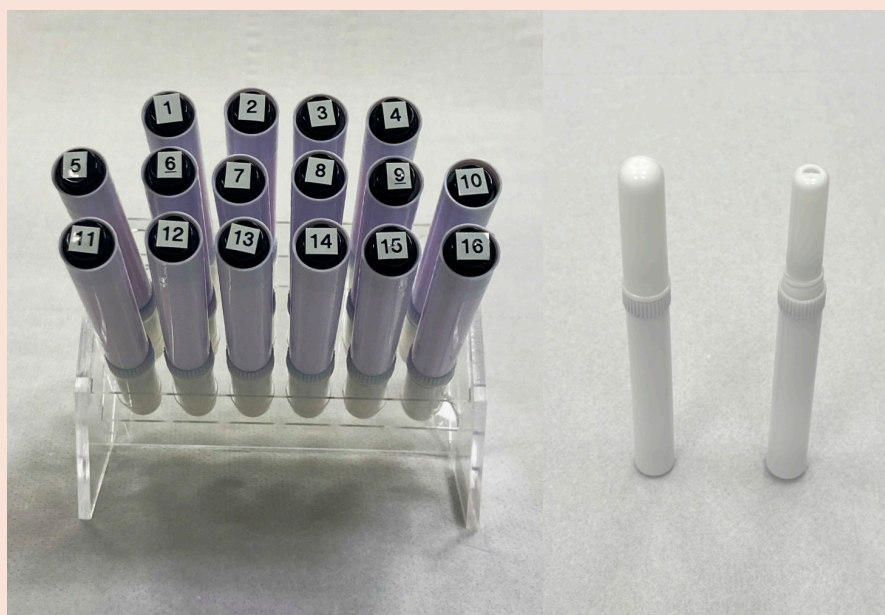
Phase 1 was performed to determine which odours were highly familiar to the Singapore population. A survey was conducted to evaluate a total of 93 odour descriptors. This comprised 69 descriptors used in the SS blue and purple kits as well as 24 new odour descriptors likely to be familiar to Singapore residents. Participants evaluated their familiarity with each odour descriptor using a Likert-type scale ranging from 1 to 5 (i.e. 1=completely unknown odour, 2=unfamiliar, 3=neutral, 4=familiar, 5=highly familiar odour). Scores of 4 and 5 were considered to be familiar to the participant. Odour descriptors identified by $\geq 75\%$ of participants as familiar were defined as recognisable odours. Only recognisable odours were considered candidates for test odours and distractor odour descriptors in the modified olfactory identification tests. The inclusion criteria for survey participants were ≥ 21 years old. The exclusion criterion was cognitive impairment. Surveys with incomplete responses were excluded from analysis.

To form SS-Sg, test odours from the blue and purple sets that were considered to be familiar were selected. Unfamiliar distractor odour descriptors were replaced with familiar descriptors within the same odour family as much as possible (e.g. replacing a non-familiar fruit odour descriptor with a familiar fruit). To form the test odours for Scentsor, 16 odours that were recognisable and locally available were selected as test odours. Only recognisable odour descriptors were used as distractor options. Similar to SS, each test odour from Scentsor had 3 other distractor options. At least 1 of the distractor options belonged to the same odour family as the test odour. To create Scentsor, commercially available unfilled felt-tip pens were filled with liquid odourants dissolved in propylene glycol to a total volume of 4 mL, similar to that of SS.¹⁶ Additionally, the odourised tip was covered such that it would be protected from accidental contact with the subject's nose without affecting odour presentation (Fig. 1). The method of administering and scoring the test was similar to that of SS. Participants were not allowed to go back and change their answers once a new odour had been presented.

Phase 2: Evaluating and finalising the SS-Sg and Scentsor

Phase 2 was performed to test the implemented changes and evaluate if familiar odours were correctly identified in $\geq 75\%$ of participants. Healthy participants were recruited and underwent both SS-Sg and Scentsor at the same sitting. Participants were randomly allocated to undergo either SS-Sg or Scentsor first, using a computer-generated random number table, to reduce bias from smell fatigue. The inclusion criterion was self-rated normal sense of

Fig. 1. Scentsor prototype. A cap is removed before presenting the odour to the subject. The housing over the odourised tip enables delivery of the odour while preventing it from making contact with the subject's nose. If there is any accidental contact, the housing can be cleaned with alcohol wipes without affecting the odour.



smell. This was defined as a score of ≥ 8 on a Visual Analogue Scale-10 (VAS-10) (i.e. 0=not able to smell anything, 5=reduced sense of smell, 10=normal sense of smell). The exclusion criteria were allergic rhinitis, rhinosinusitis, nasal polyposis, sinonasal tumours, previous sinonasal surgery, radiotherapy to the head and neck, smoking, known cognitive impairment, inability to understand or follow instructions for olfactory testing and age < 21 years old. These were determined from history taking. Allergic rhinitis was defined as the presence of ≥ 2 of the following symptoms: rhinorrhoea, nasal obstruction, sneezing and nasal itch,¹⁷ or if the participant had a known positive skin prick test or Immunoglobulin E serum test to environmental allergens. Rhinosinusitis was defined as ≥ 2 symptoms, 1 of which was either nasal obstruction or nasal discharge; and reduction/loss of smell or facial pain/pressure.¹⁸

For each test odour, the percentage of correct answers was obtained. If an odour was correctly identified by $< 75\%$ of participants, the relevant odour descriptors were adjusted. The revised versions of the smell tests were re-evaluated with another group of healthy participants and the percentage of correct answers for each test odour was again obtained.

Phase 3: Validating the SS-Sg and Scentsor

Healthy participants were recruited to generate normative data for both tests. They underwent both tests at the same sitting, the order of which was also

randomly allocated using a computer-generated random number table. The median, mean and 10th percentile scores were obtained and further stratified by age groups and sex. The 10th percentile score was used as the cut-off point between normosmia and hyposmia (i.e. a score below the 10th percentile score was indicative of hyposmia).^{16,19-21} To examine the test-retest reliability of these tests, participants were invited to repeat the tests at 2–8 weeks following the first administration. The time taken to complete each test was measured. The inclusion and exclusion criteria for healthy controls were similar to that of Phase 2.

In addition, patients with olfactory loss were recruited to assess whether both tests could discriminate between normosmia, hyposmia and anosmia. Patients who presented to the otolaryngology outpatient clinics with subjective quantitative olfactory loss were recruited. All patients underwent detailed history taking, nasoendoscopy and both olfactory tests (the order of which was randomised) at the same sitting. The inclusion criteria for patients were subjective olfactory loss secondary to rhinosinusitis, nasal polyposis, traumatic brain injury, post-viral upper respiratory tract infection, iatrogenic and idiopathic causes. The exclusion criteria were < 21 years old, VAS-10 score of ≥ 7 , qualitative olfactory dysfunction (parosmia and/or phantosmia) and cognitive impairment. VAS-10 scores of 0–3 were considered to be subjectively anosmic and scores of 4–6 were

considered to be subjectively hyposmic. SS-Sg and Scentsor scores were compared among subjectively, hyposmic, anosmic and healthy controls.

Statistical analysis

Statistical analysis was performed using Stata version 17 (StataCorp LLC, College Station, TX, US). Statistical significance was set at 0.05. Descriptive statistics for continuous variables were presented as mean (standard deviation [SD]) when normality and homogeneity assumptions were satisfied; otherwise median (interquartile range [IQR]) and *n* (%) were presented for categorical variables. The scores were compared by independent sample *t* tests when normality and homogeneity assumptions were satisfied; otherwise Mann Whitney U or Kruskal Wallis test was used. When the Kruskal Wallis test was used, post-hoc testing was done using Dunn's test with Bonferroni correction. The test-retest reliability of the tests was measured using Pearson's correlation coefficient.

RESULTS

Phase 1: Odour familiarity in the Singapore adult population

There were 414 participants (mean age 40.0 [SD 14.9] years, 159 males [38.4%]) who rated their familiarity with 93 odour descriptors (Table 1). There were 288 (69.6%) Chinese, 58 (14.0%) Malays, 40 (9.7%) Indians and 28 (6.8%) belonging to other races. Within the blue SS set, 27 of 52 (51.9%) unique odour descriptors were recognised by <75% of surveyed participants and considered unfamiliar. This included 6 (36.5%) of 16 test odours in the blue set (cinnamon, liquorice, turpentine, clove, rose, anise). Within the purple SS set, 28 of 49 (57.1%) unique odour descriptors were unfamiliar, including 9 (56.3%) test odours (pear, lilac, grapefruit, raspberry, eucalyptus, peach, mushroom, smoked meat, caramel).

All 10 recognisable odours from the SS blue set (orange, leather, peppermint, banana, lemon, garlic, coffee, apple, pineapple, fish) and 6 of 7 odours from the SS purple set with the highest recognition rates (Coke, grass, ginger, coconut, melon, onion) were selected to form SS-Sg. The original distractor odour descriptors were also modified to include only recognisable descriptors. For example, "blackberry" was replaced with "melon", "mustard" with "ketchup", and "fir" with "lemongrass". Scentsor comprised 9 test odours similar to that of SS-Sg (orange, coffee, lemon, garlic, ginger, banana, pineapple, apple, mint) and 7 original test odours (pepper, curry, smoke,

vinegar, pandan, chocolate, durian) with high recognisability rates.

Phase 2: Evaluation of SS-Sg and Scentsor

Seventy participants (mean age 35.5 [10.8] years, 18 males [25.7%]) evaluated SS-Sg and Scentsor. Thirty-five participants (50%) underwent SS-Sg first. In spite of the modifications, 7 test odours (43.8%) in SS-Sg and 3 test odours (18.8%) in Scentsor had identification rates of <75% (Table 2). The lowest performing odour in SS-Sg (banana; 44.9% recognition rate) was replaced with the lavender test odour, the remaining recognisable odour from the SS purple set. The lowest performing odour in Scentsor (pepper; 58.0% recognition rate) was replaced with a coconut test odour. For the remaining underperforming odours, descriptors were further modified using only familiar descriptors identified in Phase 1.

Sixty participants (mean age 39.1 [6.6] years, 28 males [46.7%]) evaluated the second round of modification. All test odours were correctly identified by >75% of participants. This formed the final version of SS-Sg and Scentsor.

Phase 3: Validation of SS-Sg and Scentsor

Normative data

There were 473 healthy controls (mean age 40.8 [14.5] years, age range 21–83 years, 174 (36.8%) males) who validated and provided normative data for both tests (Table 3). Half of them (*n*=236, 49.9%) performed SS-Sg first. There were 391 (82.7%) Chinese, 31 (6.7%) Malays, 33 (7.0%) Indians and 18 (3.9%) belonging to other races. The overall median score was 13.0 (12.0–14.0) for SS-Sg and 15.0 (14.0–15.0) for Scentsor. The 10th percentile scores of controls in the 21–30 years age group were 11 for SS-Sg and 13 for Scentsor. The mean time taken to complete SS-Sg was 5.3 (3.2) minutes, compared to 4.3 (2.6) minutes for Scentsor (*P*<0.001).

Effect of age

There was a significant negative relationship between age and SS-Sg scores (Spearman's rank correlation=−0.156, *P*=0.001) as well as with Scentsor scores (Spearman's rank correlation=−0.213, *P*<0.001) (Fig. 2). Post-hoc testing indicated that in SS-Sg, healthy controls aged >60 years scored significantly lower scores compared to all other age groups (21–50 years [*P*<0.001], 51–60 years [*P*=0.032]) (Table 3). In addition, controls aged 51–60 years scored significantly lower scores compared to those aged 21–30 years (*P*=0.021) and 31–40 years (*P*=0.025).

Table 1. Recognisability of odour descriptors in an adult Singapore population.

Rank	Odour descriptor	Recognition rate	Rank	Odour descriptor	Recognition rate	Rank	Odour descriptor	Recognition rate	Rank	Odour descriptor	Recognition rate
1	Durian ^a	97.8%	25	Bread	85.3%	49	Peanut	76.3%	73	Grapefruit	54.6%
2	Orange	96.9%	26	Coke	84.5%	50	Dried anchovy ^a	75.9%	74	Spearmint	53.4%
3	Coffee	95.7%	27	Mint	84.5%	51	Chlorine ^a	75.1%	75	Chive	51.5%
4	Lemon	95.4%	28	Cheese	84.1%	52	Bleach ^a	75.0%	76	Plum	50.7%
5	Curry ^a	95.1%	29	Lemongrass ^a	83.6%	53	Nail polish remover ^a	74.9%	77	Eucalyptus	47.8%
6	Smoke	93.7%	30	Peanut butter ^a	82.4%	54	Wood	74.2%	78	Chamomile	47.6%
7	Garbage ^a	93.5%	31	Peppermint	82.4%	55	Honey	73.2%	79	Clove	45.9%
8	Onion	92.3%	32	Ketchup ^a	82.1%	56	Grape	72.7%	80	Anise	45.9%
9	Fish	91.8%	33	Petrol ^a	82.0%	57	Rose	72.7%	81	Walnut	45.2%
10	Wall paint ^a	91.8%	34	Grass	81.9%	58	Cinnamon	72.2%	82	Nutmeg	42.8%
11	Cigarette	91.1%	35	Vanilla	81.4%	59	Peach	71.0%	83	Raspberry	39.9%
12	Garlic	90.3%	36	Cookie	81.2%	60	Menthol	71.0%	84	Currant	40.1%
13	Vinegar ^a	90.1%	37	Sweat ^a	81.2%	61	Pear	70.3%	85	Turpentine	39.1%
14	Ginger	89.6%	38	Jackfruit ^a	79.2%	62	Wine	69.8%	86	Rum	38.4%
15	Pandana	89.4%	39	Melon	79.0%	63	Parsley	69.1%	87	Salami	36.5%
16	Banana	88.7%	40	Honeydew ^a	78.9%	64	Ham	68.6%	88	Blackberry	31.6%
17	Chocolate	88.7%	41	Shoe leather ^a	78.8%	65	Rubber	67.9%	89	Winegum	29.5%
18	Pineapple	87.9%	42	Pomelo ^a	78.6%	66	Caramel	67.2%	90	Liquorice	23.2%
19	Mango ^a	87.8%	43	Strawberry	78.5%	67	Mushroom	65.5%	91	Lilac	19.1%
20	Limea	87.8%	44	Leather	78.3%	68	Smoked meat	64.5%	92	Sauerkraut	15.9%
21	Pepper	87.4%	45	Glue	77.3%	69	Carrot	63.3%	93	Fir	6.8%
22	French fries ^a	86.7%	46	Coconut	77.3%	70	Cherry	58.9%			
23	Apple	86.5%	47	Lavender	76.8%	71	Mustard	58.2%			
24	Toast ^a	85.6%	48	Papaya ^a	76.4%	72	Paprika	55.6%			

^a Odour descriptor not in the original Sniffin' Sticks (blue and purple)

Table 2. Identification rates of test odours in SS-Sg and Scentsor in Phase 2.

SS-Sg			Scentsor	
Number	Odour	Identification rate	Odour	Identification rate
1	Orange	97.1%	Pepper ^c	58.0%
2	Leather ^a	75.9%	Orange	97.1%
3	Peppermint	98.6%	Coffee	97.1%
4	Banana ^b	44.9%	Lemon	96.6%
5	Lemon	94.2%	Curry	72.4%
6	Garlic	69.6%	Smoke	76.8%
7	Coffee	94.2%	Garlic	92.8%
8	Apple	49.3%	Vinegar	88.4%
9	Pineapple	84.1%	Ginger	75.1%
10	Fish	82.3%	Pandan	95.7%
11	Coke	89.9%	Banana	88.4%
12	Grass	58.0%	Chocolate	72.4%
13	Ginger	71.0%	Pineapple	88.4%
14	Coconut	82.6%	Apple	88.6%
15	Melon	78.3%	Mint	88.4%
16	Onion	65.2%	Durian	92.3%

^a Replaced with the descriptor “shoe leather”

^b Replaced with “lavender” test odour in Phase 3

^c Replaced with “coconut” test odour in Phase 3

With Scentsor, healthy controls aged >60 years scored significantly lower scores compared to those aged 21–50 years ($P<0.001$). Healthy controls aged 51–60 years also scored significantly lower scores compared to those aged 21–40 years ($P<0.001$).

Effect of sex

With SS-Sg, there was no significant difference in overall median score between males and females (13 [12–14] for both sexes, $P=0.596$) (Table 3). There was also no significant difference in scores between sexes in all age groups.

With Scentsor, females scored higher than males overall (15 [14–16] versus [vs] 14 [13–15], $P<0.001$). Females scored higher than males in age groups 21–30 years (15 [15–16] vs 14 [14–15], $P<0.001$), 31–40 years (15 [14–16] vs 14 [13–15], $P<0.001$) and 41–50 years (15 [14–16] vs 14 [13–15], $P<0.001$).

Test-retest reliability

Fifty healthy controls (mean age 39.2 [5.2] years, 20 [40.0%] males) repeated both tests after a mean

duration of 2.6 (2.0) weeks. Pearson’s correlation coefficient was 0.88 ($P<0.001$) for SS-Sg and 0.90 ($P<0.001$) for Scentsor.

Patients with subjective olfactory loss

Sixty-seven patients (mean age 50.5 [15.9] years, 34 [50.8%] males) with subjective olfactory loss were recruited. The causes for olfactory loss were chronic rhinosinusitis ($n=40$, 59.7%), post-viral ($n=19$, 28.4%), idiopathic ($n=6$, 9.0%) and post-traumatic brain injury ($n=2$, 3.0%). Forty-two patients had self-rated hyposmia and 25 had self-rated anosmia.

With SS-Sg, there was a significant difference in scores across controls, self-rated hyposmias and self-rated anosmia (χ^2 [2]=85.8, $P<0.001$) (Fig. 3). The mean scores were 12.6 (2.4) for controls, 9.8 (3.2) for self-rated hyposmias and 6.0 (2.3) for self-rated anosmias. Post-hoc comparisons indicated that the mean score of controls was significantly higher than that of self-rated hyposmias ($P<0.001$), which was in turn significantly higher than the scores of self-rated anosmias ($P<0.001$).

Table 3. Healthy controls, SS-Sg and Scentsor scores by age and sex.

Age	21–30 years		31–40 years		41–50 years		51–60 years		>60 years	
n	133		137		82		66		55	
SS-Sg										
Median (IQR)	13 (12-14)		13 (11-14)		13 (12-14)		13 (12-14)		12 (9-13)	
Mean (SD)	13.1 (1.8)		12.8 (2.4)		12.9 (2.1)		12.4 (2.7)		10.7 (3.0)	
10th percentile	11		10		11		9		7	
Sex	Male (n=48)	Female (n=85)	Male (n=47)	Female (n=90)	Male (n=31)	Female (n=51)	Male (n=23)	Female (n=43)	Male (n=25)	Female (n=30)
Mean (SD)	13.3 (1.6)	13.0 (1.9)	12.3 (2.8)	13.0 (2.1)	13.2 (1.9)	12.7 (2.2)	12.2 (2.9)	12.5 (2.7)	9.8 (3.3)	11.4 (2.6)
Median (IQR)	13 (12-15)	13 (12-14)	13 (11-14)	13 (12-14)	14 (13-14)	13 (12-14)	12 (11-14)	13 (12-14)	10 (7-12)	12 (10-13)
10th percentile	11	10	9	10.5	11	10	9	11	5	7
Scentsor										
Median (IQR)	15 (14-16)		15 (14-16)		15 (14-15)		14 (14-15)		14 (12-15)	
Mean (SD)	14.8 (1.3)		14.3 (1.7)		14.4 (1.5)		14.0 (1.8)		13.0 (2.5)	
10th percentile	13		12		13		12		10	
Sex	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Mean (SD)	14.1 (1.0)	15.1 (1.2)	13.3 (1.8)	14.8 (1.4)	13.8 (1.2)	14.8 (1.6)	13.6 (1.9)	14.3 (1.7)	12.3 (2.7)	13.5 (2.3)
Median (IQR)	14 (14-15)	15 (15-16)	14 (13-15)	15 (14-16)	14 (13-15)	15 (14-16)	14 (14-15)	15 (14-15)	12 (12-14)	14 (12-15)
10th percentile	13	14	11	13	12	13	12	13	9	10

IQR: interquartile range; SD: standard deviation

Fig. 2. Scatterplot of SS-Sg and Scentsor scores in male and female subjects over age.

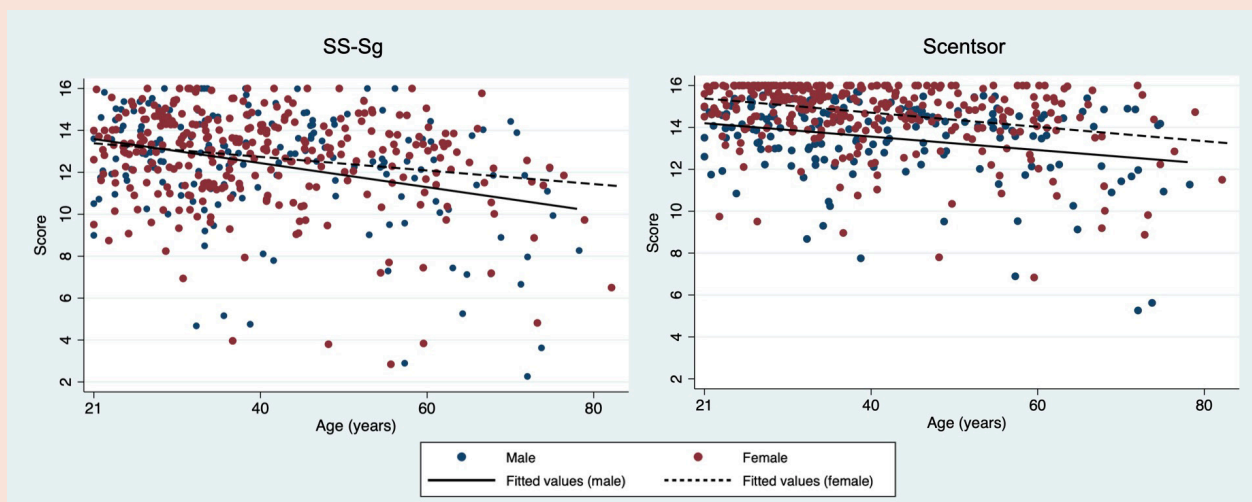
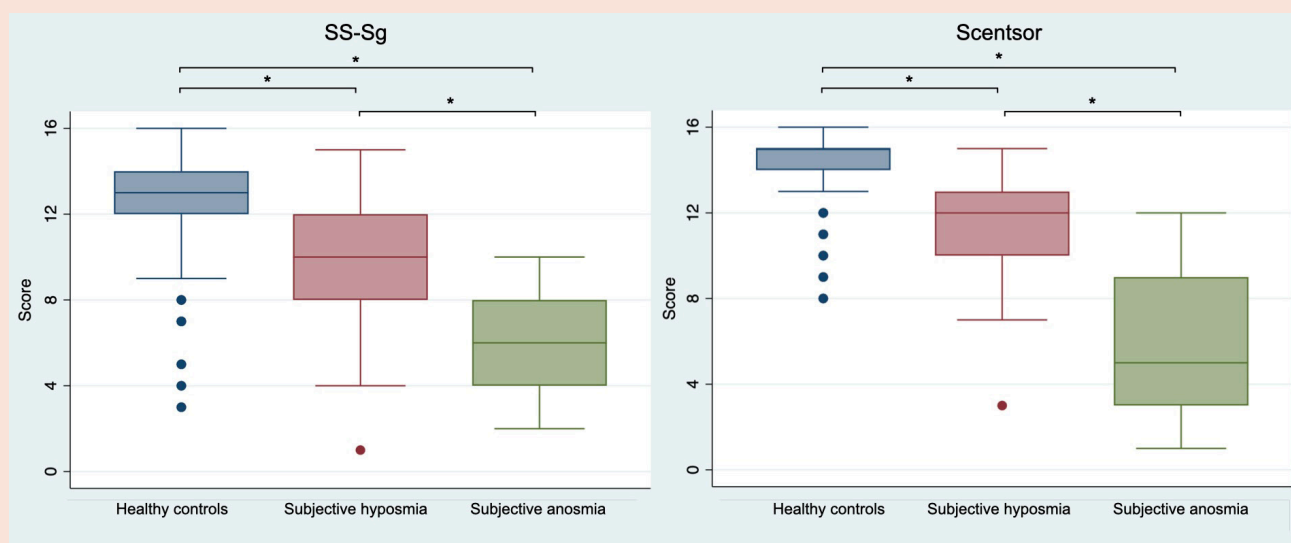


Fig. 3. Boxplot of SS-Sg and Scentsor scores in healthy controls, subjectively hyposmic and subjectively anosmic patients. Asterisk indicates a statistically significant difference.



Similarly, with Scentsor, there was a significant difference in scores across controls, self-rated hyposmias and self-rated anosmias ($\chi^2 [2]=99.9$, $P<0.001$). The median scores were 14.3 (1.8) for controls, 11.3 (2.8) for self-rated hyposmias and 5.8 (3.4) for self-rated anosmias. Post-hoc comparisons indicated that the mean score of controls was significantly higher than that of self-rated hyposmias ($P<0.001$), which was in turn, significantly higher than that of self-rated anosmias ($P<0.001$).

DISCUSSION

Odour identification tests are sensitive to cultural differences. Our study has shown that the unmodified SS was not applicable to the Singapore population, with over half of the odour descriptors in the blue and purple sets unfamiliar to most Singaporeans. This highlights the need for odour identification tests to be adapted and validated in the population of interest before it can be applied.

In this study, SS and Scentsor were adapted and developed respectively for Singapore. Both tests can be used to reliably measure olfactory function in Singapore. They have been validated for the Singapore population and normative data from a large group of healthy controls were obtained. The median scores of 13.0 and 15.0 with SS-Sg and Scentsor, respectively, for normosmic individuals is in keeping with the average scores of 13.7–14.8 in Malaysia,¹¹ China²² and Taiwan,²³ where 16-item odour identification scores have been validated. Both tests have good test-retest reliability. They can distinguish between normosmic, hyposmic and anosmic patients. Both kits are portable, re-usable and arguably cost-

effective. Both tests can be administered in about 5 minutes or less on average, which makes it suitable to be integrated into a busy clinical practice. This will allow accurate assessment and follow-up of olfaction, rather than relying only on self-reported olfactory function. They can also be combined with odour threshold and discrimination tests to form the comprehensive threshold, discrimination and identification olfactory test, which is useful in tracking olfactory dysfunction.

The scores for functional anosmia, defined as quantitatively reduced olfaction to the extent that the sense of smell is not useful in daily life,⁵ is determined by the empiric distribution of scores obtained by anosmic subjects.¹⁹ Obtaining a score of more than 7 by chance is extremely unlikely; a score of 7 and less is therefore considered to be consistent with functional anosmia. The explanation for this is that identification tests are carried out using the forced multiple-choice procedure. Even with anosmia, some chance correct answers are expected. The probability of obtaining a certain number of correct answers by chance can be calculated using the binomial distribution equation. From this, the probability of obtaining more than 7 correct answers is less than 5%.²⁰ A score of 7 or less cannot be distinguished from a purely random result and is therefore interpreted as being consistent with anosmia.²⁰ This was congruous with the findings in our study, in which anosmic patients had mean scores of 6.0 and 5.8 on SS-Sg and Scentsor, respectively.

On the other hand, hyposmia is differentiated from normosmia using the 10th percentile of the distribution of scores in normosmic subjects.^{16,19-21} The

reference group is that of young, healthy adults as they represent the best olfactory performing cohort.^{19,24} In our study, the 10th percentile score of subjects aged 21–30 years was 11 for SS-Sg and 13 for Scentsor. This means that subjects who achieve a score of 8–10 in SS-Sg and a score of 8–12 in Scentsor would be considered hyposmic. Our patients with hyposmia had mean scores of 9.8 and 11.3 on SS-Sg and Scentsor, respectively, which was in keeping with these hyposmia score ranges defined by the normative data. A subject's test score should be defined primarily by this range, even though it can also be interpreted with respect to age group and sex. For example, a 75-year-old male who scores 11 on Scentsor can be considered "normosmic" relative to his age group and sex. However, in terms of his overall ability to smell, he is still considered to be hyposmic. This is similar in principle to the assessment of other human senses. Using hearing assessment as an analogy, if a 75-year-old has an audiogram pure tone average of 35 decibels, his diagnosis is mild hearing loss, even if this level of hearing is reasonably acceptable for his age.

An age-related decline in scores was observed in age groups above 50 years with both SS-Sg and Scentsor, which is consistent with the findings of other studies carried out in other populations.^{8,24–26} This was largely to be expected, as olfactory function generally declines with age.^{27,28} With regard to sex, females scored significantly better than males in Scentsor; however, no significant sex-related differences in scores were observed with the SS-Sg. In general, females typically tend to perform better than males in odour detection and recognition tasks,²⁹ although the difference in absolute test scores is probably small.

Ideally, we would have adapted only 1 SS set, rather than have a combination of blue and purple Sticks. The blue set was the preferred candidate for adaptation as it had fewer unfamiliar test odours than the purple set. However, it was difficult to replace all 6 unfamiliar test odours (cinnamon, liquorice, turpentine, clove, rose, anise) with similar-smelling alternative descriptors that were also familiar to Singaporeans. In the Malaysian version of the SS, "liquorice" was replaced with "fennel seed"¹¹ and in the Taiwanese version of the SS, "clove" was replaced with "wood",²³ both of which are unfamiliar odours in Singapore. This highlights that even among close neighbours in Asia, there is considerable cultural diversity which influences familiarity with odours. Hence, SS-Sg was formed by combining familiar odours from the blue and purple SS. This creates a slight drawback in that Sticks will have to be procured individually to form the SS-Sg. However, having an adapted version of the SS gives an option to assess

olfaction using a set of odours that is standardised internationally.

Scentsor was created as a pragmatic alternative odour identification test using locally-sourced odours. The general feedback from healthy participants was that Scentsor's odours were more recognisable compared to SS-Sg, even when test odours descriptors were the same (e.g. ginger). This might account for the shorter mean test duration with Scentsor, as well as the slightly lower overall SS-Sg scores in spite of adapting all odour descriptors to familiar ones. Familiarity with odours may become more crucial when testing older subjects or screening for neurodegenerative disorders. This is because identifying odours relies on the ability to detect the stimulus as well as to match it to the correct identity, of which the latter requires intact executive function and semantic memory.³⁰ We partially adopted the SS method of odour storage and presentation as it is practical, inexpensive and the odours are minimally exposed to the environment, which prolongs odour longevity and integrity. In addition, for hygiene purposes, we modified the design to protect the odourised tip such that it is not able to touch the participant's nose. In the event that there is accidental contact, the device can be easily disinfected with an alcohol wipe without affecting the odour. As Scentsor is produced locally, it may be a more cost-effective option. For these reasons, Scentsor is our preferred odour identification test.

There are limitations with this study. Healthy controls did not undergo nasoendoscopy to examine the nasal cavity and olfactory cleft. However, we applied stringent inclusion and exclusion criteria to ensure participants were likely to have normal olfactory function. There were more female participants who provided normative data, which could have raised the overall scores, even if the difference in scores between sex may be small. There were fewer participants aged above 50 years old who provided normative data. Future studies may be needed to obtain normative olfactory data in older age groups.

Within Asia, countries that have developed or adapted olfactory identification tests to their own populations include China,²² Malaysia,¹¹ Japan,^{31,32} South Korea,^{33,34} Taiwan,^{14,23} Thailand³⁵ and Vietnam.¹⁵ The authors hope that the validation of SS-Sg and Scentsor will standardise olfactory testing carried out in Singapore. The potential applications for a validated olfactory test are vast. Olfactory impairment is a key symptom of chronic rhinosinusitis. In severe type 2 inflammatory chronic rhinosinusitis with nasal polyps, treatment with biologics requires documentation of anosmia on smell testing,^{18,36} and biologics themselves have the potential to improve anosmia.^{37,38} During the COVID-19 pandemic, olfactory loss was a

prominent symptom,³⁹ and anosmia was reported to have high specificity for the SARS-CoV-2 infection.⁴⁰ Olfactory tests have been used as screening tools for SARS-CoV-2 in the community,⁴¹ emergency department,⁴² among healthcare workers⁴³ and inpatients.⁴⁴ Olfactory loss is an early symptom of neurodegenerative conditions, in particular Parkinson's disease⁴⁵⁻⁴⁷ and Alzheimer's disease.⁴⁸ In older adults, severe olfactory dysfunction is also an early indicator of frailty, a syndrome of accelerated physiologic decline and increased vulnerability to external stressors beyond normal ageing.⁴⁹ These are relevant issues in our ageing society. An odour identification test may be a practical screening tool for early detection of neurodegeneration or frailty in vulnerable individuals.

CONCLUSION

SS-Sg and Scentsor are 2 odour recognition tests that have been validated for the Singapore population. Either of these tests, referenced against the normative data obtained in this study, should be used when assessing olfaction in Singapore.

Supplementary Fig. S1. STROBE flow diagram of participant recruitment.

Acknowledgment

The authors would like to express their sincere gratitude to: Mr HS Kwan for being our consultant perfumer; senior staff nurses Barbara Oh, Ong Susan, Tan Elyn and Yee Weiling for carrying out olfactory testing; Ms Serene Ng for coordinating the study; and Prof Thomas Hummel for his advice on the process of adapting olfactory identification tests.

Ethics statement

This study was approved by the National Healthcare Group Domain Specific Review Board and participants gave informed consent. All experimental procedures were consistent with the Declaration of Helsinki.

Declaration

All the authors have no affiliations or financial involvement with any commercial organisation with a direct financial interest in the subject or materials discussed in the manuscript.

REFERENCES

- Man REK, Wong CMJ, Gupta P, et al. Understanding the Prevalence and Risk Factor Profile of Olfactory Impairment and Its Impact on Patient Health Indicators and Economic Outcomes in Community-Dwelling Older Asian Adults. *Innov Aging* 2024;8:igae088.
- Desiato VM, Levy DA, Byun YJ, et al. The Prevalence of Olfactory Dysfunction in the General Population: A Systematic Review and Meta-analysis. *Am J Rhinol Allergy* 2021;35:195-205.
- Philpott CM, Wolstenholme CR, Goodenough PC, et al. Comparison of subjective perception with objective measurement of olfaction. *Otolaryngol Head Neck Surg* 2006;134:488-90.
- Wehling E, Lundervold AJ, Espeset T, et al. Even cognitively well-functioning adults are unaware of their olfactory dysfunction: Implications for ENT clinicians and researchers. *Rhinology* 2015;53:89-94.
- Whitcroft KL, Altundag A, Balungwe P, et al. Position paper on olfactory dysfunction: 2023. *Rhinology* 2023;61:1-108.
- Kobal G, Hummel T, Sekinger B, et al. "Sniffin' Sticks": screening of olfactory performance. *Rhinology* 1996;34:222-6.
- Oleszkiewicz A, Taut M, Sorokowska A, et al. Development of the Arabic version of the "Sniffin' Sticks" odor identification test. *Eur Arch Otorhinolaryngol* 2016;273:1179-84.
- Ribeiro JC, Simoes J, Silva F, et al. Cultural Adaptation of the Portuguese Version of the "Sniffin' Sticks" Smell Test: Reliability, Validity, and Normative Data. *PLoS One* 2016;11:e0148937.
- Niklassen AS, Ovesen T, Fernandes H, et al. Danish validation of sniffin' sticks olfactory test for threshold, discrimination, and identification. *Laryngoscope* 2018;128:1759-66.
- Balungwe P, Huat C, Matanda R, et al. Adaptation of the Sniffin' Sticks Test in South-Kivu. *Eur Ann Otorhinolaryngol Head Neck Dis* 2020;137:467-71.
- Sai-Guan L, Husain S, Zahedi FD, et al. Cultural Adaptation of Sniffin' Sticks Smell Identification Test: The Malaysian Version. *Iran J Otorhinolaryngol* 2020;32:213-22.
- Nordin S, Bramerson A, Liden E, et al. The Scandinavian Odor-Identification Test: development, reliability, validity and normative data. *Acta Otolaryngol* 1998;118:226-34.
- Kobayashi M, Saito S, Kobayakawa T, et al. Cross-cultural comparison of data using the odor stick identification test for Japanese (OSIT-J). *Chem Senses* 2006;31:335-42.
- Hsu NI, Lai JT, Shen PH. Development of Taiwan Smell Identification Test: a quick office-based smell screening test for Taiwanese. *Am J Rhinol Allergy* 2015;29:e50-4.
- Tran TN, Thi Dang TH, Thai TT, et al. Development and validation of the Vietnamese smell identification test. *Parkinsonism Relat Disord* 2023;113:105494.
- Hummel T, Sekinger B, Wolf SR, et al. "Sniffin' sticks": olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses* 1997;22:39-52.
- Wise SK, Damask C, Roland LT, et al. International consensus statement on allergy and rhinology: Allergic rhinitis - 2023. *Int Forum Allergy Rhinol* 2023;13:293-859.
- Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. *Rhinology* 2020;58:1-464.
- Kobal G, Klimek L, Wolfensberger M, et al. Multicenter investigation of 1,036 subjects using a standardized method for the assessment of olfactory function combining tests of odor identification, odor discrimination, and olfactory thresholds. *Eur Arch Otorhinolaryngol* 2000;257:205-11.
- Wolfensberger M, Schnieper I, Welge-Lüssen A. Sniffin' Sticks: a new olfactory test battery. *Acta Otolaryngol* 2000;120:303-6.
- Gellrich J, Sparing-Paschke LM, Thieme T, et al. Normative data for olfactory threshold and odor identification in children and adolescents. *Int J Pediatr Otorhinolaryngol* 2019;123:5-9.
- Su B, Wu D, Wei Y. Development of Chinese odor identification test. *Ann Transl Med* 2021;9:499.
- Shu CH, Yuan BC, Lin SH, et al. Cross-cultural application of the "Sniffin' Sticks" odor identification test. *Am J Rhinol* 2007;21:570-3.

24. Oleszkiewicz A, Schriever VA, Croy I, et al. Updated Sniffin' Sticks normative data based on an extended sample of 9139 subjects. *Eur Arch Otorhinolaryngol* 2019;276:719-28.
25. Doty RL, Shaman P, Applebaum SL, et al. Smell identification ability: changes with age. *Science* 1984;226:1441-3.
26. 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;32:489-502.
27. Murphy C, Schubert CR, Cruickshanks KJ, et al. Prevalence of olfactory impairment in older adults. *JAMA* 2002;288:2307-12.
28. Wilson RS, Arnold SE, Tang Y, et al. Odor identification and decline in different cognitive domains in old age. *Neuroepidemiology* 2006;26:61-7.
29. Brand G, Millot JL. Sex differences in human olfaction: between evidence and enigma. *Q J Exp Psychol B* 2001;54:259-70.
30. Hedner M, Larsson M, Arnold N, et al. Cognitive factors in odor detection, odor discrimination, and odor identification tasks. *J Clin Exp Neuropsychol* 2010;32:1062-7.
31. Takagi SF. A standardized olfactometer in Japan. A review over ten years. *Ann N Y Acad Sci* 1987;510:113-8.
32. Saito S, Ayabe-Kanamura S, Takashima Y, et al. Development of a smell identification test using a novel stick-type odor presentation kit. *Chem Senses* 2006;31:379-91.
33. Hong SC, Yoo YS, Kim ES, et al. Development of KVSS Test (Korean Version of Sniffin' Sticks Test). *Korean J Otolaryngol-Head Neck Surg* 1999;42:855-60.
34. Ha JG, Kim J, Nam JS, et al. Development of a Korean Culture-Friendly Olfactory Function Test and Optimization of a Diagnostic Cutoff Value. *Clin Exp Otorhinolaryngol* 2020;13:274-84.
35. Kasemsuk N, Thanaviratnanich S, Piroomchai P. A study of 30 odors panel smell identification test, smell detection threshold and University of Pennsylvania Smell Identification Test (UPSIT) in Thailand. *Auris Nasus Larynx* 2020;47:1003-8.
36. Hellings PW, Alobid I, Anselmo-Lima WT, et al. EUFOREA/EPOS2020 statement on the clinical considerations for chronic rhinosinusitis with nasal polyps care. *Allergy* 2024;79:1123-33.
37. Bachert C, Han JK, Desrosiers M, et al. Efficacy and safety of dupilumab in patients with severe chronic rhinosinusitis with nasal polyps (LIBERTY NP SINUS-24 and LIBERTY NP SINUS-52): results from two multicentre, randomised, double-blind, placebo-controlled, parallel-group phase 3 trials. *Lancet* 2019;394:1638-50.
38. Mullol J, Lund VJ, Wagenmann M, et al. Mepolizumab improves sense of smell in severe chronic rhinosinusitis with nasal polyps: SYNAPSE. *Rhinology* 2024;62:320-9.
39. Borsetto D, Hopkins C, Philips V, et al. Self-reported alteration of sense of smell or taste in patients with COVID-19: a systematic review and meta-analysis on 3563 patients. *Rhinology* 2020;58:430-6.
40. Struyf T, Deeks JJ, Dinnes J, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. *Cochrane Database Syst Rev* 2022;5:CD013665.
41. Hunter SR, Zola A, Ho E, et al. Using SCENTinel(R) to predict SARS-CoV-2 infection: insights from a community sample during dominance of Delta and Omicron variants. *Front Public Health* 2024;12:1322797.
42. Keck JW, Bush M, Razick R, et al. Performance of formal smell testing and symptom screening for identifying SARS-CoV-2 infection. *PLoS One* 2022;17:e0266912.
43. Cao AC, Nimmo ZM, Mirza N, et al. Objective screening for olfactory and gustatory dysfunction during the COVID-19 pandemic: a prospective study in healthcare workers using self-administered testing. *World J Otorhinolaryngol Head Neck Surg* 2021;8:249-56.
44. Ziuzia-Januszewska L, Dobrzynski P, Slaczka K, et al. Simple Disposable Odor Identification Tests for Predicting SARS-CoV-2 Positivity. *Int J Environ Res Public Health* 2021;18:10185.
45. Haehner A, Hummel T, Reichmann H. Olfactory dysfunction as a diagnostic marker for Parkinson's disease. *Expert Rev Neurother* 2009;9:1773-9.
46. Chen H, Shrestha S, Huang X, et al. Olfaction and incident Parkinson disease in US white and black older adults. *Neurology* 2017;89:1441-7.
47. Marrero-Gonzalez P, Iranzo A, Bedoya D, et al. Prodromal Parkinson disease in patients with idiopathic hyposmia. *J Neurol* 2020;267:3673-82.
48. Bouhaben J, Delgado-Lima AH, Delgado-Losada ML. The role of olfactory dysfunction in mild cognitive impairment and Alzheimer's disease: A meta-analysis. *Arch Gerontol Geriatr* 2024;123:105425.
49. Wang E, Wroblewski KE, McClintock MK, et al. Examining the Longitudinal Relationship Between Olfactory Dysfunction and Frailty in Community-Dwelling, older US Adults. *Otolaryngol Head Neck Surg* 2024;171:261-8.