OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 34 (2019) 254–258

Archives of CLINICAL NEUROPSYCHOLOGY

Older Adult Normative Data for the Sniffin' Sticks Odor Identification Test

Kayci L. Vickers[1,2], Kathleen Breslin[1,3], David R. Roalf[4], Vidyulata Kamath[5], Sharon X. Xie[6], Paul J. Moberg[1,4], David A. Wolk[1], Dawn Mechanic-Hamilton[1,*]

[1]*Alzheimer's Disease Center, Department of Neurology, Perelman School of Medicine, Philadelphia, PA, USA*
[2]*Department of Psychology, Drexel University, Philadelphia, PA, USA*
[3]*Department of Psychology, Pennsylvania College of Osteopathic Medicine, Philadelphia, PA, USA*
[4]*Brain Behavior Laboratory, Department of Psychiatry, Perelman School of Medicine, Philadelphia, PA, USA*
[5]*Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MA, USA*
[6]*Department of Biostatistics and Epidemiology, Perelman School of Medicine, Philadelphia, PA, USA*

*Corresponding author at: University of Pennsylvania, Department of Neurology, Penn Memory Center, PCAM 2-South, 3400 Civic Center Blvd., Philadelphia, PA 19104, USA. Tel.: +1-215-662-4516; fax: +1-215-615-4695
E-mail address: dawn.mechanic@uphs.upenn.edu

Editorial Decision 14 February 2018; Accepted 16 March 2018

Abstract

Objective: The current study establishes normative Sniffin' Sticks Odor Identification Test (SS-OIT) scores for cognitively intact older adults.
Method: Two hundred and twenty-six cognitively normal older adults were identified as eligible for the current study (Mean Age = 70.49 years; 71.7% female). Important demographic covariates were identified using step-wise regression, and a normative regression equation was developed.
Results: Analyses of the effects of demographic variables (including age, education, and sex) on SS-OIT performance revealed that age was the only significant predictor, $b = -0.07$, $SE_b = .01$, $p < .01$. A final regression equation was determined and normative data are reported in 5-year increments for a number of percentile ranks.
Conclusions: Normative performance on the SS-OIT for adults over the age of 50 was established in a large and demographically diverse sample. These data are needed in order for clinicians to be able to include olfactory testing, a sensitive marker of neurodegeneration, in their assessment armamentarium.

Keywords: Olfactory function; Older adult; Normative data; Odor identification; Neurodegenerative disease; Mild cognitive impairment

Introduction

Olfactory functioning is an important factor in both quality of life and personal safety due to its role in one's ability to experience pleasure when eating and to identify environmental toxins (Pinto, Wroblewski, Kern, Schumm, & McClintock, 2015). Impairments in olfactory functioning are present in a variety of neurodegenerative diseases including Alzheimer's disease (AD), Parkinson's disease, Huntington's disease, vascular dementia, frontotemporal dementia, and multiple system atrophy (Doty, 2017). Importantly, milder olfactory dysfunction is also present during preclinical stages of such diseases (Roalf et al., 2017), and has even been suggested to predict the conversion of mild cognitive impairment (MCI) to AD (Devanand et al., 2010; Roberts et al., 2016). In AD, odor identification is affected in early stages of the disease whereas olfactory threshold, or the lowest perceivable concentration of an odor, is affected in later stages (Eibenstein et al., 2005). This observation has led to the use of odor identification as a more sensitive measure for detecting olfactory dysfunction in the early stages of neurodegenerative diseases. When combined with cognitive screening instruments, one measure of odor identification, the Sniffin' Sticks Odor Identification Test (SS-OIT), has been found to improve diagnostic accuracy in detecting MCI (Quarmley et al., 2016). In order to reliably determine

doi:10.1093/arclin/acy021 Advance Access publication on 28 March 2018

the presence of olfactory abilities and interpret poor performance among those at risk for neurodegeneration, it is important to establish normative data for cognitively intact older adults.

Odor identification tests are the most widely used to assess olfactory function, likely as the result of their availability and ease of use. The Sniffin' Sticks test was developed by Kobal and Hummel in Erlangen, Germany (Kobal, Hummel, Sekinger, Barz, Roscher, & Wolf, 1996) and is comprised of three separate subtests measuring chemosensory functioning (including odor threshold, odor discrimination, and odor identification). Normative data were developed for each subtest as well as for an overall score (Hummel, Kobal, Gudziol, & Mackay-Sim, 2007). The SS-OIT subtest includes 16 felt-tipped markers with common odors and is used as a brief screener for olfactory dysfunction because of its ease of administration, low patient burden, and the ability to re-use test stimuli across patients (Hummel et al., 2007). In the validation study, the average SS-OIT score of healthy individuals was between 12.06 and 13.68 across all adult age groups (aged 16 to 55+ years). The validity of the SS-OIT is supported by strong correlations with the University of Pennsylvania's Smell Identification Test (UPSIT), a widely utilized test of olfactory function (Doty, Shaman, & Dann, 1984; Wolfensberger, Schnieper, & Welge-lu, 2000). The SS-OIT has many fewer items than the UPSIT (40 items), making it more efficient and therefore more appropriate for patients who fatigue easily. The SS-OIT is also more economical than the UPSIT for clinics that administer more than 8 olfactory identification tests in a 6-month period. Reliability of the SS-OIT is supported by moderate correlations across test sessions (Albrecht et al., 2008; Hummel et al., 2007; Sorokowska, Albrecht, Haehner, & Hummel, 2015).

While the SS-OIT is a valid and reliable measure (Albrecht et al., 2008; Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997; Wolfensberger et al., 2000), more information is needed to increase its utility in a geriatric US sample. Published normative scores for the SS-OIT combine all individuals over the age of 55, leaving little interpretive value for individuals in the geriatric population. One study using Sniffin' Sticks found that SS-OIT performance significantly decreased with age and that this was more pronounced over 65 years of age (Hummel et al., 1997). Given typical declines in olfactory function with normal aging (Doty et al., 1984; Hummel et al., 1997), the suggestion of more rapid declines among older adults, and the prevalence of olfactory dysfunction in neurodegenerative diseases (Doty, 2017), normative data specific to the geriatric population is vital to the interpretation of SS-OIT results among older adults.

Methods

Participants

Participants were recruited from the Penn Memory Center (PMC) and the Clinical Core of the Alzheimer's Disease Core Center (ADCC) at the University of Pennsylvania. Participants included in the present study were enrolled in National Alzheimer's Coordinating Center (NACC) study through the ADCC as healthy control participants. Individuals eligible for the current study participated in annual evaluations including the following: assessment of personal and family history, neurologic and general physical examinations, neuropsychological assessment, and study partner report of everyday cognitive and functional ability. Participants completed the UDS-2 which includes measures of memory, language, attention, executive function, and processing speed. All data were reviewed by PMC clinicians and consensus opinion determined that participants were cognitively intact. Eligible participants were cognitively normal (MMSE $\geq$ 27) and produced a valid SS-OIT score (n = 292). Participants were excluded if their medical chart revealed evidence of significant neurological comorbidity (including prior history of stroke or transient ischemic attack, parkinsonism, radiation to the head or brain, or seizures; n = 32), uncontrolled hypertension (SBP > 160 or DBP > 95; Qiu, Winblad, & Fratiglioni, 2005; n = 5), history of traumatic brain injury (n = 14), significant psychiatric disorders (bipolar disorder or schizophrenia; n = 3), substance use disorder (n = 5), or if their testing was completed in Spanish (n = 4). Three individuals were excluded for SS-OIT scores which were considered to be extreme values, based on the interquartile range (IQR; i.e., Q1 – [1.5 × IQR] < Acceptable Scores < Q3 + [1.5 × IQR]).

The resulting sample (n = 226) was 71.7% female (n = 162) and ranged in age from 50 to 93 years old (M = 70.49, SD = 8.57). The sample was 63.7% Caucasian (n = 144), 32.7% African American (n = 74), 3.1% Hispanic (n = 7), and 0.4% other (n = 1). Participants obtained between 9 and 20 years of education (M = 15.74, SD = 2.76). The mean MMSE score was 29.14 (SD = 0.92) and SS-OIT ranged from 7 to 16 (M = 12.87, SD = 1.91).

Materials

During the SS-OIT, 16 odorants were presented in felt-tip pens, which were approximately 14 cm long and contained an inner diameter of 1.3 cm (Hummel et al., 2007). These pens contain liquid odorants or odorants dissolved in propylene glycol. The experimenter removes the cap of the pen for approximately 3–5 s and the tip of the pen is placed below the participant's nostrils.

Following presentation of the pen, participants were presented with four multiple-choice options and asked to choose one. Total scores range from 0 to 16. No other tests of olfaction were administered, including other components of the Sniffin' Sticks test.

Statistical Analyses

During exploratory data analyses, mean and standard deviations in SS-OIT were calculated for age-stratified groups (see Table 1). The effects of demographic variables on SS-OIT were analyzed with step-wise linear regression analyses. The full regression model included age, sex, years of education, and all two-way interactions between these variables. These variables were selected as they were identified in the past to affect SS-OIT performance (e.g., Pinto et al., 2015). Although a quadratic age term was considered, visual inspection of a scatter plot revealed no evidence for non-linear relationship between age and SS-OIT. Age at time of testing and years of education were continuous variables. Sex was dummy coded, such that 0 = female and 1 = male. The full regression model was reduced by using a series of step-wise hierarchical regression models, which excluded non-significant predictors in subsequent iterations.

The assumptions of regression analysis were tested for each model: normal distribution of residuals (Kolmogorov–Smirnov tests on the residual values), homoscedasticity (applying the Levene test to residuals), and multicollinearity (Belsley, Kuh, & Welsch, 2004). The skew of the distribution is within acceptable limits for the sample size (Skew = –0.81; Tabachnick & Fidell, 2013, p. 80–81; Gravetter & Wallnau, 2013, p. 50–51). Although non-normality was present (as represented by significant Kolmogorov–Smirnov test, $p < .01$) on SS-OIT, normality tests of this nature can be overly sensitive with larger sample sizes given the presence of enough power to detect non-normality even when not clinically meaningful (Maxwell & Delaney, 2004, p. 115). In line with this, analyses of Q-Q plots revealed that age, education, and SS-OIT scores were sufficiently normal to continue with regression-based analyses and transformations were not performed (Maxwell & Delaney, 2004, p. 114; Pallant, 2013, p. 66).

Next, regression-based normative data were established through a four-step procedure, including the following steps: (1) expected SS-OIT test scores were computed using the final multiple regression model, (2) residual scores were calculated (i.e., observed score – expected score), (3) residuals were standardized, and (4) standardized residuals were converted into percentile values (see Oosterhuis, van der Ark, & Sijtsma, 2016 for more complete explanation of procedures).

A Bonferroni correction was applied to avoid Type I errors. Given that two models were tested, a conservative alpha level of .025 was applied as the statistical threshold of significance ($p = .05/2 = .025$). All statistical tests are two sided. All analyses were conducted with IBM's Statistical Package for Social Sciences, Version 24 (SPSS).

Results

Preliminary analysis revealed that age at the time of testing was a significant predictor of SS-OIT scores ($b = –.07$, $SE_b = .01$, $p < .01$). Neither sex ($b = –.33$, $SE_b = .27$, $p = .22$) nor education ($b = .07$, $SE_b = .04$, $p = .10$) were significant predictors and all two-way interactions between age, sex, and education were non-significant. Therefore, the normative regression equation included only age as a predictor ($b = –.07$, $SE_b = .01$, $p < .01$, Adjusted $R^2 = .11$).

Normative data were developed using the four-step procedure described previously. Following norming procedures, the final regression equation to determine predicted SS-OIT scores based on age was: $(–.07 * \text{Age}) + 18.05 = \text{SS-OIT}$, $SD(\text{residual}) = .998$. As an example, for an individual who is 78 years old and has an obtained SS-OIT score of 12, the equation to determine predicted SS-OIT score would be: $(-.07 * 78) + 18.05 = 12.59$. Once predicted and obtained SS-OIT scores are calculated, the difference between them becomes that individual's residual score (i.e., obtained SS-OIT – predicted SS-OIT = residual; $12 - 12.59 = -.59$). Then, the residual score is divided by the *SD*(residual) value ($-.59 \div .998 = -.59$) to provide a standardized (Z) score. A Z-score of –.59 on a normal distribution corresponds to performance at the 28th percentile, meaning the obtained score of 12 for the individual in the example corresponds to performance at the 28th percentile for their age. While this equation can be used as described for

Table 1. Sniffin' Sticks Odor Identification Test (SS-OIT) scores by age group

Age group	*n*	Mean age (*SD*)	Mean years of education (*SD*)	Mean SS-OIT score (*SD*)
50–60 years	23	58.94 (2.55)	15.78 (2.76)	13.83 (1.64)
61–65 years	49	63.14 (1.43)	15.37 (2.77)	13.37 (1.45)
66–70 years	54	67.81 (1.49)	15.56 (2.52)	12.80 (1.86)
71–75 years	38	72.61 (1.20)	16.05 (2.84)	13.13 (1.93)
76–80 years	27	77.74 (1.46)	16.07 (2.63)	12.59 (1.53)
81+ years	35	85.20 (3.00)	15.91 (3.19)	11.57 (2.29)

Table 2. Conversion of obtained SS-OIT score to percentile rank by age

		Age (in years)[a]								
		50	55	60	65	70	75	80	85	90
SS-OIT	16	93	96	98	99	99	>99	>99	>99	—
	15	67	79	88	93	97	99	99	99	>99
	14	29	42	56	69	80	89	94	97	99
	13	6	11	20	31	44	58	71	82	89
	12	1	1	3	7	12	21	33	46	60
	11	<1	<1	<1	1	2	4	7	14	23
	10	<.01	<.01	<1	<1	<1	<1	1	2	4
	9	—	—	<.01	<.01	<.01	<1	<1	<1	<1
	8	—	—	—	—	—	<.01	<.01	<.01	<1
	7	—	—	—	—	—	—	—	—	<.01

[a]Exact percentiles based on age and obtained SS-OIT scores may be calculated by using the equation −.07 Age + 18.05 = Predicted SS-OIT and following instructions in-text.

a specific age, the authors have also included a user-friendly normative table (Table 2), which calculates corresponding percentile ranks based on obtained SS-OIT performance across many ages.

Discussion

The primary aim of the current study was to establish the normal range of performance on the SS-OIT among cognitively intact older adults. To determine the influence of various demographic variables, including age, sex, and years of education, preliminary regression analysis was used to identify important covariates. Age was determined to be the only variable that significantly influenced SS-OIT performance. A linear age effect was included, such that an expected decrease of .07 SS-OIT points occurred with each year of increased age. Sex did not significantly impact SS-OIT scores and education was determined to be highly collinear with the effects of age.

A normative regression equation was determined, which can be used with individuals over the age of 50, based on established normative procedures (Oosterhuis et al., 2016). For ease of clinical use, a normative table displaying percentile ranks was made available as part of this publication. While this extends previous work which also identified age as a predictor of SS-OIT performance (e.g., Pinto et al., 2015), it adds to the literature by providing normative data for older adults, including both Caucasian and African-American individuals. This is particularly important within a US sample, as SS-OIT was validated in Europe and norms have not been established in a US population (Rumeau, Nguyen, & Jankowski, 2016).

This study adds to the discussion on sex effects; no significant effect of sex was identified in our sample. Some studies have found sex differences in SS-OIT performance (e.g., Pinto et al., 2015). Other studies have argued that with larger sample sizes, sex-based discrepancies are negligible (Hummel et al., 2007). The present study supports that with greater sample sizes, negligible differences exist between older men and women on SS-OIT performance.

The current study is limited by the restricted age range; the regression equation cannot be applied to individuals under the age of 50 (i.e., 45-year olds). While this was not the specific goal of the study, it is important to note that regression-based norms lack generalizability to ages outside of the sample, as the relationship between predictors may change with age (i.e., may be better classified as a quadratic relationship). The relatively high educational attainment within the present sample may also limit the generalizability of findings, given the potential for increased cognitive reserve among individuals who have more formal education (Meng & D'Arcy, 2012). Increased cognitive reserve may theoretically, in turn, increase resistance to decline in olfactory functioning over time, though this relationship has not yet been shown. Future studies on the SS-OIT should include educational attainment in order to provide greater information about the impact it may have on odor identification, particularly among older adults. Due to these limitations, the normative data should not be applied to individuals with less than 12 years of education. Additionally, the sample used for the current study was largely composed of Caucasian and African-American individuals living in the US, which limits generalizability to other racial and cultural groups. Given that past studies have identified both age and cultural factors as potentially impacting performance, (Pinto et al., 2015; Yaffe et al., 2017) future studies should aim to collect normative data for the SS-OIT for a larger age range within the US population and include groups underrepresented in the current study.

It is important to consider that it is likely that the present sample includes individuals with preclinical AD (Devanand et al., 2010). Although detection of AD was not the explicit purpose of this study, the presence of AD pathology in individuals who present with normal cognition on psychometric measures may reduce the average scores expected by age. For this reason, future normative studies should aim to account for molecular biomarkers of AD and longitudinal data on cognitive stability over time, in order to provide data on a robust sample that is likely to remain cognitively normal. In the future, it will be useful to compare normative data stratified by biomarker status with "mixed" samples to understand the impact of preclinical AD in a population identified as cognitively normal.

The current study provides valuable normative data on odor identification performance on the SS-OIT in a large sample of cognitively normal adults over age 50. This data provides researchers and clinicians with a tool to reliably and accurately establish the integrity of olfactory function. Use of this normative data, along with history, physical and neuropsychological assessment, will aid in the early detection of neurodegenerative disorders.

Funding

This work was supported by National Institutes of Health, National Institute on Aging P30 AG10124; NIMH K01 MH102609 (Dr. Roalf).

References

Albrecht, J., Anzinger, A., Kopietz, R., Schopf, V., Kleemann, A. M., Pollatos, O., et al. (2008). Test-retest reliability of the olfactory detection threshold test of the Sniffin' Sticks. *Chemical Senses*, *33*, 461–467. https://doi.org/10.1093/chemse/bjn013.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Devanand, D. P., Tabert, M. H., Cuasay, K., Manly, J. J., Schupf, N., Brickman, A. M., et al. (2010). Olfactory identification deficits and MCI in a multi-ethnic elderly community sample. *Neurobiology of Aging*, *31*, 1593–1600. https://doi.org/10.1016/j.neurobiolaging.2008.09.008.

Doty, R. L. (2017). Olfactory dysfunction in neurodegenerative diseases: Is there a common pathological substrate? *The Lancet Neurology*, *16*, 478–488.

Doty, R. L., Shaman, P., Applebaum, S. L., Giberson, R., Siksorski, L., & Rosenberg, L. (1984a). Smell identification ability: Changes with age. *Science*, *226*, 1441–1443.

Doty, R. L., Shaman, P., & Dann, M. (1984b). Development of the University of Pennsylvania Smell Identification Test: A standardized microencapsulated test of olfactory function. *Physiology & Behavior*, *32*, 489–502.

Eibenstein, A., Fioretti, A. B., Simaskou, M. N., Sucapane, P., Mearelli, S., Mina, C., et al. (2005). Olfactory screening test in mild cognitive impairment. *Neurological Sciences*. https://doi.org/10.1007/s10072-005-0453-2.

Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the behavioral sciences* (9th ed.). Belmont, CA: Cengage Learning.

Hummel, T., Kobal, G., Gudziol, H., & Mackay-Sim, A. (2007). Normative data for the "Sniffin' Sticks" including tests of odor identification, odor discrimination, and olfactory thresholds: An upgrade based on a group of more than 3,000 subjects. *European Archives of Oto-Rhino-Laryngology*. https://doi.org/10.1007/s00405-006-0173-0.

Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., & Kobal, G. (1997). "Sniffin" Sticks': Olfactory performance assessed by the combined testing of odour identification, odor discrimination and olfactory threshold. *Chemical Senses*, *22*, 39–52. https://doi.org/10.1093/chemse/22.1.39.

Kobal, G., Hummel, T., Sekinger, B., Barz, S., Roscher, S., & Wolf, S. (1996). "Sniffin' Sticks": Screening of olfactory performance. *Rhinology*, *34*, 222–226.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*, *Vol. 1*). New York: Psychology Press.

Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the cognitive reserve hypothesis: A systematic review with meta-analyses and qualitative analyses. *PLoS One*, *7*, e38268.

Oosterhuis, H. E. M., van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*, 191–202. https://doi.org/10.1177/1073191115580638.

Pallant, J. (2013). *SPSS survival guide*. New York: Open University Press.

Pinto, J. M., Wroblewski, K. E., Kern, D. W., Schumm, L. P., & McClintock, M. K. (2015). The rate of age-related olfactory decline among the general population of older U.S. adults. *Journals of Gerontology – Series A Biological Sciences and Medical Sciences*, *70*, 1435–1441. https://doi.org/10.1093/gerona/glv072.

Qiu, C., Winblad, B., & Fratiglioni, L. (2005). The age-dependent relation of blood pressure to cognitive function and dementia. *Lancet Neurology*, *4*, 487–499. https://doi.org/10.1016/S1474-4422(05)70141-1.

Quarmley, M., Moberg, P. J., Mechanic-Hamilton, D., Kabadi, S., Arnold, S. E., Wolk, D. A., et al. (2016). Odor identification screening improves diagnostic classification in incipient Alzheimer's disease. *Journal of Alzheimer's Disease*, *55*, 1497–1507. https://doi.org/10.3233/JAD-160842.

Roalf, D. R., Moberg, M. J., Turetsky, B. I., Brennan, L., Kabadi, S., Wolk, D. A., et al. (2017). A quantitative meta-analysis of olfactory dysfunction in mild cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, *88*, 226–232. https://doi.org/10.1136/jnnp-2016-314638.

Roberts, R. O., Christianson, T. J., Kremers, W. K., Mielke, M. M., Machulda, M. M., Vassilaki, M., et al. (2016). Association between olfactory dysfunction and amnestic mild cognitive impairment and Alzheimer disease dementia. *JAMA Neurology*, *73*, 93–101.

Rumeau, C., Nguyen, D. T., & Jankowski, R. (2016). How to assess olfactory performance with the Sniffin' Sticks test. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *133*, 203–206.

Sorokowska, A., Albrecht, E., Haehner, A., & Hummel, T. (2015). Extended version of the "Sniffin' sticks" identification test: Test-retest reliability and validity. *Journal of Neuroscience Methods*, *243*, 111–114. https://doi.org/10.1016/j.jneumeth.2015.01.034.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.

Wolfensberger, M., Schnieper, I., & Welge-lu, A. (2000). Sniffin'Sticks: A New Olfactory Test Battery. *Acta Otolaryngology*, *120*, 303–306. https://doi.org/10.1080/000164800750001134.

Yaffe, K., Freimer, D., Chen, H., Asao, K., Rosso, A., Rubin, S., et al. (2017). Olfaction and risk of dementia in a biracial cohort of older adults. *Neurology*, *88*, 456–462.