

Unsupervised Clustering of Olfactory Phenotypes

American Journal of Rhinology & Allergy

2022, Vol. 36(6) 796–803

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/19458924221114255

journals.sagepub.com/home/ajr



Rodney J. Schlosser, MD¹ , Judy R. Dubno, PhD¹ ,
Mark A. Eckert, PhD¹ , Andreana M. Benitez, PhD²,
Matthew Gregoski, PhD³, Viswanathan Ramakrishnan, PhD³,
Lois Matthews¹, and Zachary M. Soler, MD¹

Abstract

Background: Current clinical classifications of olfactory function are based primarily upon a percentage of correct answers in olfactory identification testing. This simple classification provides little insight into etiologies of olfactory loss, associated comorbidities, or impact on the quality of life (QOL).

Methods: Community-based subjects underwent olfactory psychophysical testing using Sniffin Sticks to measure threshold (T), discrimination (D), and identification (I). The cognitive screening was performed using Mini-Mental Status Examination (MMSE). Unsupervised clustering was performed based upon T, D, I, and MMSE. Post hoc differences in demographics, comorbidities, and QOL measures were assessed.

Results: Clustering of 219 subjects, mean age 51 years (range 20-93 years) resulted in 4 unique clusters. Cluster 1 was the largest and predominantly younger normosmics. Cluster 2 had the worst olfaction with impairment in nearly all aspects of olfaction and decreased MMSE scores. This cluster had higher rates of smoking, heart disease, and cancer and had the worst olfactory-specific QOL. Cluster 3 had normal MMSE with relative preservation of D and I, but severely impaired T. This cluster had higher rates of smoking and heart disease with moderately impaired QOL. Cluster 4 was notable for the worst MMSE scores, but general preservation of D and I with moderate loss of T. This cluster had higher rates of Black subjects, diabetes, and viral/traumatic olfactory loss.

Conclusion: Unsupervised clustering based upon detailed olfactory testing and cognitive testing results in clinical phenotypes with unique risk factors and QOL impacts. These clusters may provide additional information regarding etiologies and subsequent therapies to treat olfactory loss.

Keywords

olfaction, phenotype, cluster analysis

Introduction

Otorhinolaryngologists are very familiar with classifying patients based upon patterns of sensory loss into relevant phenotypes that provide insight into causative factors that impact subsequent treatment. For example, audiometric patterns separate patients into conductive and sensorineural hearing loss. Patients with conductive loss may benefit from myringotomy and tube if middle ear effusion is present, while patients with sensorineural loss proceed down completely different treatment algorithm pathways. This ability to precisely classify patients based upon sensory testing then allows physicians to recommend treatments specific to each patient. Unfortunately, olfactory categorization remains relatively simplistic.

The most common olfactory tests determine the percentage correct of olfactory identification items which leads to categorization based on the severity of olfactory identification loss.¹

Even more sophisticated testing methods that measure olfactory threshold and discrimination in addition to identification, classify patients based upon overall percentage correct. It is possible that patients may have varying deficits between the olfactory domains, yet have identical overall scores and thus

¹Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina, Charleston, South Carolina

²Department of Neurology, Medical University of South Carolina, Charleston, South Carolina

³Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina

Corresponding Author:

Rodney J. Schlosser, Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina, 135 Rutledge Ave, MSC550, Charleston, SC 29425, USA.

Email: schlossr@musc.edu

are categorized similarly despite heterogeneous olfactory dysfunction (OD). Adults identified as anosmic based upon identification tests are likely to include a wide spectrum of patients. Patients exposed to inhaled irritants may have peripheral sinonasal inflammation, but relative sparing of central olfactory function. Others may have central processing dysfunction due to neurocognitive disorders. Patients who have suffered trauma may have true sensorineural OD and shear injury at the cribriform plate. Finally, many patients may have a multifactorial loss. Given this heterogeneity, it is critical that we begin developing improved phenotypes based upon easily obtained olfactory and other testing that can stratify patients into more homogenous groups that share similar patterns of olfactory loss and enable the development of personalized therapies, both surgical and pharmacological. For example, adults with peripheral OD due to impaired olfactory cleft airflow may respond to intranasal medications or surgery targeted at intranasal structures. Sensorineural OD may have localized fibrosis, ischemic change, shear injury, or metaplasia of the olfactory neuroepithelium. Rather than targeting peripheral factors, these adults may respond to olfactory training to restore functional olfactory neuroepithelium. Finally, patients with central OD due to neurodegenerative changes in central olfactory structures are unlikely to benefit from therapies directed at peripheral or sensorineural causes.

The goal of this study was to use unsupervised clustering to classify OD based upon four common clinical metrics, specifically, the 3 components of Sniffin Sticks testing (threshold, discrimination, and identification) and cognitive assessment using the Mini-Mental Status Examination (MMSE). MMSE was selected based on the known association between OD and cognitive impairment.² The clusters were inspected to confirm that the characterized unique patterns of OD, including clinical symptoms, and if they differed by demographics, comorbidities, or olfactory exposures that may provide insight into the etiology of olfactory loss. Finally, the clusters were also examined to determine if there were also cluster differences in symptoms and quality of life (QOL), based on patient-reported outcome measures (PROMs).

Methods

Enrollment

Subjects 18 years of age and older were recruited from the general population by word of mouth, by posting notices, and through referrals of family and friends from clinics associated with the Medical University of South Carolina (MUSC). The study was approved by the MUSC Institutional Review Board (HR E-607R). Recruited subjects were not seeking medical care and enrollment was not determined by perceived olfactory ability. Subjects taking immunosuppressive medications including systemic steroids and chemotherapy were excluded as were those diagnosed with autoimmune disease or dementia of any form including Parkinson's disease and

Alzheimer's disease which may render olfactory testing less reliable. Subjects with active upper respiratory tract infections were also excluded, as were those unable to complete written forms in English. Written informed consent was obtained and the subject data was de-identified. The study was conducted prior to the COVID pandemic.

Demographics and Comorbidities

Olfactory-specific information pertaining to demographics, medical comorbidities, potential olfactory exposures, and risk factors was collected from the subject at the time of the visit using a written questionnaire. Demographic information included subject age at enrollment, sex, race, and ethnicity. Self-reported medical comorbidities, including allergic rhinitis, asthma, diabetes, and depression were collected from each subject, and subjects were further instructed to include all existing, physician-diagnosed conditions. Nasal endoscopy was performed to evaluate for polyps or other abnormalities. Olfactory-specific information was collected including prior sinus surgery and repeated or prolonged exposures to materials such as smoking, chemical odors, exhaust fumes, ionizing radiation, and silica. Subjects were asked if they had suffered a perceived olfactory loss due to head trauma or viral infection. During the initial visit, a trained research coordinator performed the MMSE.

Quantitative Olfactory Testing

Psychophysical olfactory testing was performed using "Sniffin Sticks" testing (Burghardt, Wedel, Germany), consisting of felt pens with odorant substances.³ The testing consisted of threshold (T), discrimination (D), and identification (I) subtests, which form a composite TDI summed score. Subjects were asked to close their eyes during the threshold and discrimination portions of testing, while the identification portion required their eyes remain open. Using dilutions of n-butanol, the threshold portion of the test was performed using a single-staircase triple-forced choice procedure whereby the subject selected the pen with the odorant among the set of 3 pens presented. Correct identification of the odorant pen in two consecutive trials resulted in a decrease in the concentration (increase in dilution) of the odorant pen for the next trial, whereas 1 incorrect identification of the odorant pen resulted in an increase in concentration for the next trial. This procedure was conducted until 7 reversals were completed, and the dilution level at the final four reversals was averaged to produce the resultant threshold score (a higher threshold score indicates better olfactory function). The discrimination subtest also uses triplet sets of odorant pens, of which 2 contain the same odorant and the third, a different odorant. Sixteen sets of 3 are presented in random order, and the subject earned one point for correctly identifying the pen containing the odorant that differs from the other 2 of the triplet. The discrimination score was reported as the total number of correct responses

(out of 16). The identification subtest presents 16 suprathreshold odorants individually, along with 4 answer choices for each pen. The total number of correct responses (out of 16) was reported as the identification score. The scores for threshold (T), discrimination (D), and identification (I) were then summed to produce a composite TDI score. Olfactory categories of anosmic, hyposmic, and normosmic were established using previously published normative values.

Patient-reported outcome measures

Subjects completed validated PROMs examining numerous aspects of QOL potentially associated with olfactory loss,⁴⁻⁶ including: (1) olfactory-specific QOL using Questionnaire for Olfactory Disorders-Negative Statements (QOD); (2) depression-specific QOL using Patient Health Questionnaire 9 (PHQ9); (3) social isolation using University of California, Los Angeles (UCLA) social isolation scale; (4) visual analog scales (VAS) assessing the impact of OD upon mood, food enjoyment, social interactions, safety, hygiene, sex, cooking, appetite, and weight changes; and (5) VAS of cardinal symptoms of cytokine release syndrome including nasal blockage, nasal drainage, sinus pain/pressure, and overall olfaction.

Statistical Analysis

Analyses were conducted using version 27 of SPSS (IBM Corporation). A nonhierarchical (K-means, normal mixture) was performed, using the 4 defining variables, namely T, D, I, and cognitive function scores. In an initial exploratory step, the possible number of clusters was identified using Ward's linkage method with squared Euclidean distances and a bottom-up agglomerative dendrogram. We examined between 2 and 8 cluster solutions. Not surprisingly, a 2-cluster solution differentiated between positive z-scores and negative z-scores on outcome measures likely indicative of younger individuals with closer to normal olfaction versus others with poorer olfaction. As shown in Figure 1 (i.e. dendrogram) a 4-cluster solution emerged at a rescaled distance of ~3.0 (range 0-25) and was favored over a 3-cluster solution based on cluster distinction.

Based on the initial exploratory cluster analyses, a nonhierarchical (K-means, normal mixture) was performed, using the same 4 defining variables (ie, T, D, I, and cognitive function). Silhouette scores revealed Cluster 1 (.387) as having the most defined cluster followed by Clusters 2, 3, and 4 (silhouette scores of .288, .281, and .188, respectively). Further,

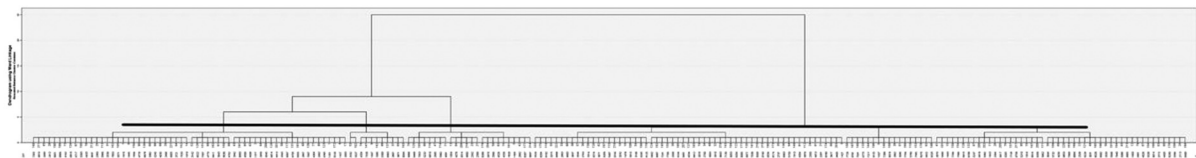


Figure 1. Depiction of bottom-up dendrogram demonstrating that a 4 cluster solution provides clusters that are more alike than other cluster solutions ranging from 2 to 8.

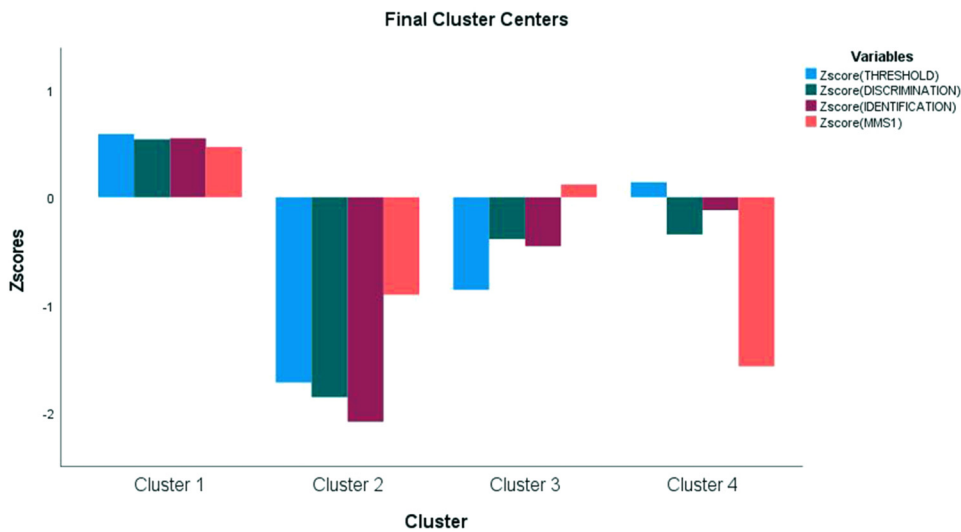


Figure 2. Z scores to standardize variables of T, D, I, and MMSE demonstrate the contribution of each variable to cluster composition. For example, cluster 1 is driven by higher scores in all variables, while cluster 2 is driven by lower scores in all variables. Cluster 3 is driven by lower T scores with increased MMSE, while cluster 4 is driven by lower MMSE.

Abbreviations: T, threshold; D, discrimination; I, identification; MMSE, Mini-Mental Status Examination.

variable Z-scores were well-defined to cluster centers as shown in Figure 2.

Differences in demographics, comorbidities, exposures, other risk factors, and QOL measures were assessed across clusters using a series of analyses of variance and subsequent Bonferroni corrected post hoc analyses for continuous outcomes and chi-square analyses with Bonferroni corrected z -tests for examination of frequency data. Alternative Welch Robust Tests of Means and Games–Howell post hoc analyses for continuous outcomes and Fisher’s exact test with subsequent Bonferroni corrected z -tests for examinations of frequency data were used if data assumptions were violated. If alternative Welch Robust Test of Means were significant and subsequent Games–Howell post hoc analyses failed to yield significant results, variables were reanalyzed using nonparametric Kruskal–Wallis analyses with all pairwise comparisons.

Results

Cohort Characteristics

A total of 219 subjects were enrolled, with cohort characteristics detailed in Table 1. The average age of the cohort was 51.1 years (range = 20–93 years), and females made up the majority (61%). Consistent with the population of our state, 21% of subjects were Black with the majority White (73%). Notable comorbidities potentially related to olfaction included allergic rhinitis (30%), depression (16%), diabetes (10%), hypertension (27%), head injury (10%), and ever had smoke exposure (37%). The olfactory function of the entire cohort is detailed in Table 2. The overall mean TDI score was 29.24 (SD = 7.0), and 46.6% of the cohort was considered normosmic. The overall MMSE score was 28.9 (SD = 1.5).

Defining Clusters

Cluster analysis resulted in 4 clusters as shown in Table 2. As expected, there were significant differences in essentially all cluster-defining variables (T, D, I, MMSE). Cluster 1 was the largest and demonstrated the best TDI and MMSE scores, mean scores of 33.82 and 29.60, respectively. Cluster 2 had the poorest olfactory function with TDI in the anosmic range, a mean score of 13.56, and impairment of all aspects of olfaction. Cluster 3 had relative preservation of D and I with severely impaired T, mean of 3.91, and normal MMSE scores. The final cluster (4) was notable for the lowest MMSE scores, mean of 26.59, but general preservation of D and I function with moderately impaired T, mean score of 6.64.

Post hoc Differences in Demographics, Comorbidities, and Risk Factors

Cluster 1 was notable for being the youngest cluster (mean age of 43.73 years) which was not surprising given their normal olfactory function (Tables 2 and 3). Cluster 2 demonstrated higher rates of smoking, heart disease, and cancer.

Table 1. Overall Characteristics of Study Population (N = 219).

Variable	Count (percentage) or Mean \pm SD
Sex	
Male	86 (39.3%)
Female	133 (60.7%)
Age (years)	51.1 \pm 18.0
Race	
White	159 (72.6%)
Black	46 (21.0%)
Asian	8 (3.7%)
Other	6 (2.7%)
Body mass index	27.4 + 6.0
Tobacco	
Ever used	82 (37.4%)
Currently use	21 (9.6%)
Ever exposed to second-hand smoke	113 (51.6%)
Currently exposed to second-hand smoke	12 (5.5%)
Alcohol	
Current use	137 (62.6%)
Former use	17 (7.8%)
Comorbidities, risk factors	
Vertigo/dizziness	19 (8.7%)
Tinnitus	24 (11%)
Head injury	22 (10%)
Ear ache	16 (7.3%)
Thyroid issues	25 (11.4%)
Heart disease	21 (9.6%)
Stroke	3 (1.3%)
Kidney issues	12 (5.5%)
Arthritis	45 (20.5%)
Diabetes	21 (9.6%)
Hypertension	60 (27.4%)
Memory deficits	5 (2.3%)
Cancer	24 (11.0%)
Depression	36 (16.4%)
Apnea	21 (9.6%)
Allergic rhinitis	65 (29.7%)
Asthma	16 (7.3%)
Cholesterol	44 (20.1%)
Acid reflux	28 (12.8%)
Depression	36 (16.4%)
Anxiety	28 (12.8%)
Nasal polyps	4 (1.8%)

This cluster trended toward being the oldest, with a mean age of 67.53 years. Cluster 3 also had higher rates of smoking and heart disease with the addition of arthritis. Cluster 4 was unique in that it has higher rates of Black subjects (52%). It also had higher rates of diabetes, hypertension, and a history of upper respiratory infection (URI) or trauma associated with OD.

Patient-Reported Outcomes

Table 4 demonstrates that there were global differences in olfactory-specific symptoms and QOL between clusters. In

Table 2. Cluster-defining Variables Among Clusters.

	Overall	Cluster 1	Cluster 2	Cluster 3	Cluster 4	P-value
Cluster distribution	219 (100%)	124 (56.6%)	21 (9.6%)	45 (20.5%)	29 (13.2%)	
Threshold	6.28 ± 2.72	7.86 ± 1.62 ²³⁴	1.56 ± 1.07 ¹³⁴	3.91 ± 1.79 ¹²⁴	6.64 ± 1.75 ¹²³	<.001
Discrimination	10.98 ± 2.63	12.40 ± 1.73 ²³⁴	6.10 ± 2.19 ¹³⁴	9.96 ± 1.61 ¹²	10.07 ± 1.98 ¹²	<.001
Identification	11.97 ± 2.91	13.56 ± 1.48 ²³⁴	5.90 ± 2.00 ¹³⁴	10.64 ± 2.17 ¹²	11.62 ± 2.08 ¹²	<.001
Total TDI	29.24 ± 7.04	33.82 ± 2.86 ²³⁴	13.56 ± 3.35 ¹³⁴	24.51 ± 3.00 ¹²⁴	28.33 ± 3.62 ¹²³	<.001
MMSE	28.91 ± 1.48	29.60 ± 0.64 ²³⁴	27.57 ± 2.16 ¹³	29.09 ± 0.63 ¹²⁴	26.59 ± 1.52 ¹³	<.001

TDI: Threshold, Discrimination, Identification, MMSE: Mini-Mental Status Exam.

Note: (1) Welch Robust Test and Games–Howell Post hoc test used when unequal homogeneity is present. (2) Superscript numbers indicate clusters with significant post hoc differences. (3) Total TDI was not used as a cluster-defining variable, but is included for completeness.

Table 3. Post hoc Analysis of Demographics, Comorbidities and Risk Factors Across Clusters.

	Overall (N = 219)	Cluster 1 (N = 124)	Cluster 2 (N = 21)	Cluster 3 (N = 45)	Cluster 4 (N = 29)	P-value
Demographics						
Race						
White	159 (72.6%)	101 (81.5%) ⁴	15 (71.4%)	30 (66.7%)	13 (44.8%) ¹	<.001
Black	46 (21%)	14 (11.3%) ⁴	6 (28.6%)	11 (24.4%)	15 (51.7%) ¹	
Asian/other	14 (6.4%)	9 (7.3%)	0 (0.0%)	4 (8.9%)	1 (3.4%)	
Gender						
Male	86 (39.3%)	45 (36.3%)	6 (28.6%)	23 (51.1%)	12 (41.4%)	.244
Female	133 (60.7%)	79 (63.7%)	15 (71.4%)	22 (48.9%)	17 (58.6%)	
Age	51.07 ± 17.97	43.73 ± 16.73 ¹²³	67.53 ± 14.46 ¹	59.76 ± 15.42 ¹	57.06 ± 12.66 ¹	<.001
Comorbidities						
Heart disease	29 (13.2%)	4 (3.2%) ²³	5 (23.8%) ¹	11 (24.4%) ¹	1 (3.4%)	<.001
Diabetes	21 (9.6%)	6 (4.8%) ⁴	2 (9.5%)	7 (15.6%)	6 (20.7%) ¹	.017
Hypertension	60 (27.4%)	21 (16.9%) ⁴	8 (38.1%)	16 (35.6%)	15 (51.7%) ¹	<.001
Arthritis	45 (20.5%)	17 (13.7%) ³	7 (33.3%)	16 (35.6%) ¹	5 (17.2%)	.007
Cancer	24 (11.0%)	9 (7.3%) ²	7 (33.3%) ¹⁴	7 (15.6%)	1 (3.4%) ²	.004
Asthma	16 (7.3%)	6 (4.8%)	3 (14.3%)	6 (13.3%)	1 (3.4%)	.110
Cholesterol	44 (20.1%)	19 (15.3%)	4 (19.0%)	13 (28.9%)	8 (27.6%)	.367
GERD	28 (12.8%)	19 (15.3%)	1 (4.8%)	6 (13.3%)	2 (6.9%)	.516
Depression	36 (16.4%)	21 (16.9%)	2 (9.5%)	8 (17.8%)	5 (17.2%)	.888
Anxiety	28 (12.8%)	20 (16.1%)	1 (4.8%)	3 (6.7%)	4 (13.8%)	.306
Allergic rhinitis	65 (29.7%)	35 (28.2)	6 (28.6%)	16 (35.6%)	8 (27.6%)	.897
Risk factors						
Smoker now/ever	82 (37.4%)	35 (40.6%) ²³	13 (61.9%) ¹	26 (57.8%) ¹	8 (27.6%)	<.001
Current EtOH	137 (62.6%)	84 (67.7%) ⁴	14 (66.7%)	28 (62.2%)	11 (37.9%) ¹	.028
Former EtOH	17 (7.8%)	3 (2.4%) ²⁴	4 (19.0%) ¹	4 (8.9%)	6 (20.7%) ¹	<.001
URI OD	16 (7.3%)	3 (2.4%) ⁴	3 (14.3%)	5 (11.1%)	5 (17.2%) ¹	<.001
Trauma OD	3 (1.4%)	0 (0.0%) ⁴	1 (4.8%)	0 (0.0%)	2 (6.9%) ¹	<.001
Other cause OD	9 (4.1%)	3 (2.4%) ²	4 (19.0%) ¹	1 (2.2%)	1 (3.4%)	.037

Abbreviations: GERD, gastroesophageal reflux disease; EtOH, alcohol; URI, upper respiratory infection; OD, olfactory dysfunction.

Note: (1) Fisher exact test utilized when expected cell counts <5. (2) Superscript numbers indicate significant mean differences at $P < .05$.

general, clusters 1 and 4 showed very little impact of any OD upon QOL. Cluster 2, not surprisingly given the severity of psychophysical loss, had the greatest impact on olfaction-specific symptoms and QOL. Cluster 3 reported moderate impacts upon olfactory-specific PROMs. Interestingly, clusters 2 and 3 also had significant nasal symptoms of blockage and drainage, potentially indicating associations with peripheral etiologies, such as smoke

exposure or other conditions. Other PROMs to assess depression and social isolation showed no differences between clusters.

Discussion

Machine learning in olfactory research has included pattern-based odor detection, development of olfactory

Table 4. Post hoc Analysis of Patient-Reported Outcome Measures Across Clusters.

	Overall	Cluster 1	Cluster 2	Cluster 3	Cluster 4	P-value
Olfaction-specific measures						
QODNS	4.73 ± 7.42	2.99 ± 4.38 ²	12.55 ± 13.30 ¹	5.73 ± 7.99	5.21 ± 7.99	.004
Olfaction VAS symptom scores						
Mood	0.67 ± 1.69	0.34 ± 0.92 ²³⁴	1.64 ± 2.89 ¹	0.97 ± 1.93 ¹	0.91 ± 2.28 ¹	<.001*
Enjoy food	1.39 ± 2.64	0.70 ± 1.61 ²	3.71 ± 3.91 ¹	1.94 ± 3.09	1.78 ± 3.20	<.001
Social interactions	0.69 ± 1.73	0.38 ± 1.09 ²³	2.06 ± 3.11 ¹	1.56 ± 0.23 ¹	0.97 ± 2.33	.002*
Safety	1.14 ± 2.42	0.59 ± 1.63 ²	4.66 ± 3.75 ¹³⁴	1.10 ± 2.26 ²	1.02 ± 2.29 ²	<.001
Hygiene	0.71 ± 1.85	0.43 ± 1.14	2.51 ± 3.37	0.59 ± 1.18	.77 ± 2.17	.055
sex life	0.36 ± 1.23	0.52 ± 1.52	1.33 ± 2.62	0.34 ± 0.62	0.87 ± 2.25	.249
Difficult cooking	0.96 ± 2.18	0.47 ± 1.41 ²	3.02 ± 3.45 ¹	1.16 ± 2.19	1.27 ± 2.80	.004
Appetite change	1.06 ± 2.36	0.63 ± 1.77 ¹²³	2.04 ± 3.19 ¹	1.52 ± 2.64 ¹	1.49 ± 3.08 ¹	<.001*
Weight change	0.71 ± 1.84	0.39 ± 1.12	1.19 ± 2.58	1.04 ± 2.19	1.22 ± 2.72	.081
Other measures						
PHQ9	2.22 ± 3.96	2.40 ± 4.46	2.81 ± 5.15	1.60 ± 2.13	2.00 ± 2.73	.591
DJG	1.40 ± 1.55	1.21 ± 1.56	1.95 ± 1.66	1.36 ± 1.35	1.90 ± 1.59	.055
UCLA	3.73 ± 1.26	3.75 ± 1.25	4.00 ± 1.55	3.58 ± 1.27	3.66 ± 1.08	.630
OCES	.51 ± 1.13	.38 ± .91	1.00 ± 2.38	.73 ± 1.13	.53 ± .90	.169
CRS VAS symptom scores						
Nasal blockage	2.20 ± 2.65	1.68 ± 2.25 ³	3.71 ± 3.29	3.10 ± 3.06 ¹	1.81 ± 2.41	.006
Nasal discharge	2.50 ± 2.89	2.05 ± 2.57 ⁴	3.42 ± 3.58	3.50 ± 3.15	2.17 ± 2.86 ¹	.032
Reduction smell	1.59 ± 2.86	0.82 ± 1.88 ²³	5.47 ± 4.29 ¹³⁴	2.46 ± 3.17 ¹²⁴	0.75 ± 1.88 ²³	<.001
Sinus pain pressure	1.04 ± 1.94	0.94 ± 1.87	1.49 ± 2.39	1.06 ± 1.95	1.11 ± 1.92	.677

Abbreviations: QODNS, Questionnaire for Olfactory Disorders–Negative Statements; VAS, visual analog scale; PHQ9, Patient Health Questionnaire 9; UCLA, University of California, Los Angeles Social Isolation Scale; OCES, Olfactory Cleft Endoscopy Score; CRS, cytokine release syndrome; DJG, De Jong Gierveld Scale.

Note: (1) Welch Robust Test and Games–Howell post hoc test used when unequal homogeneity is present. *Indicates subsequent Kruskal–Wallis performed with all pairwise comparisons when Games–Howell Post hoc was not significant. (2) Superscript numbers indicate significant post hoc differences.

biomarkers, prediction of chemical properties of odorant molecules, and pattern recognition of olfactory phenotypes.⁷ In this study, we focused upon unsupervised clustering of olfactory phenotypes and demonstrated 1 normosmic group and 3 discrete groups of dysosmic subjects across the spectrum of OD. The goal of unsupervised clustering is to identify novel groups of subjects and avoid preconceived investigator bias. We could have elected to include numerous variables, such as age, sex, or race as cluster-defining variables. However, this approach likely would result in clustering similar to current classifications based largely upon variables deemed relevant by investigators and potentially subject to bias. Post hoc analysis did demonstrate differences in age and race between clusters and ultimately some combination of patient-specific variables with olfactory and cognitive testing may provide the most insightful and useful for clinical management.

Additionally, current classifications of OD rely largely upon patient history regarding past trauma, viral URI, or other potential causes. As shown in Table 3, relying solely upon the history of such risk factors may not always accurately group patients. One of the most common patient-reported clinical factors, viral URI, actually had representation among all 4 clusters. So while

cluster 4 had the highest percentage of URI-associated OD, there were still significant numbers of subjects who reported URI-associated OD in the other 3 clusters. This may be due to differences in viral-induced deficits along the olfactory pathway, with some viral-OD subjects suffering only peripheral loss, while others may lose sensorineural or even central function. Thus unsupervised clustering allows us to consider additional groupings in an unbiased fashion that may lend additional insight into patient history or demographics alone. This method of classifying patients is not unlike current assessments of patients who report hearing loss that includes both clinical factors and sensory testing. Patients who report traumatic hearing loss can suffer conductive loss if injury occurs to the tympanic membrane or ossicles. On the other hand, they may suffer injury to the cochlea or auditory nerve and have a sensorineural loss. Thus clinical history alone does not provide a complete picture of the pathology. Relying solely upon hearing tests in patients with hearing loss is useful to classify conductive versus sensorineural loss, but does not take into account multiple etiologies that cause each type of hearing loss. Ultimately, otolaryngologists use both clinical history and sophisticated sensory tests for assessment. It is expected that future investigations into better classifications of olfactory loss will include both clinical factors as

well as a refined understanding of olfactory and cognitive testing.

Machine learning-based solely upon olfactory function (T, D, and I) has been performed previously by Lotsch et al.⁸ There are several notable differences with our study that provide interesting additional perspectives. The first is the population studied. Lotsch et al included OD patients who were presenting to a Smell and Taste Clinic, while we specifically chose to study a community-based sample who were not seeking medical care in the hopes our results would be broadly applicable for future screening efforts as well as understanding unrecognized OD. In addition, we chose to include an instrument for cognitive dysfunction, the MMSE, to define the phenotypes, in addition to olfactory function. We also reviewed differences in comorbidities that may contribute to OD such as diabetes, cardiovascular disease, and hypertension between olfactory phenotypes. We felt these were important aspects to consider given the known associations between olfaction and these factors. While we did include these additional variables, future studies must consider that the number of these variables may track together, such as age, race, cognitive function, diabetes, and other risk factors, thus large population-based studies are needed. Additionally, clusters 2 and 4 were relatively smaller than other clusters. Ultimately, cluster size is dependent upon the population studied and variables used to define clusters. Thus, studying an older population with higher rates of cognitive dysfunction would likely result in increased numbers of cluster 2 and 4 subjects. Similarly, our community-based cohort had relatively few numbers of patients (only 13 of 219 subjects) with any history of chronic rhinosinusitis or prior sinonasal surgery, so we were unable to assess this variable.

Considering both clustering studies yield several important similarities, as well as differences. The first cluster observed by Lotsch et al⁸ had some impairment of D with relative preservation of T and I and was dominated by healthy controls. Our cluster 1 had the best olfactory function with an average TDI score of 33.8, which is in the normal range. Unlike Lotsch et al,⁸ we did not see the impairment of D, perhaps due to variations in the population studied as mentioned above.

The second cluster noted by Lotsch et al⁸ had the most severe impairment of all aspects of T, D, and I. This cluster had an overrepresentation of congenital anosmics and possibly traumatic OD. This is similar to our cluster 2, as we found dramatic impairment across all aspects of olfactory function with near-complete loss of T. We did not have any subjects with known congenital OD and we did not find an increased prevalence of trauma in this group. This cluster, however, had higher rates of smoking, heart disease, and cancer of any type. It is unknown if patients in this cluster have suffered a widespread insult that is impacting the entire olfactory pathway or if this is simply the result of focal peripheral etiologies that have

progressed in severity over time. Our cluster 2 had the poorest TDI scores of any cluster and reported nasal symptoms of both blockage and drainage, as well as impaired olfactory-specific symptoms and poorer QOL. Others have shown that olfactory-specific QOL is associated with etiology, psychophysical function, age, disease duration, and sex.⁹ Interestingly, prior studies have shown that flavor identification is a better predictor of QOL than orthonasal olfactory testing.¹⁰ Thus there is a complex relationship between psychophysical chemosensory testing, demographic factors, and QOL.

Lotsch et al⁸ had a third cluster with moderately impaired T, relative preservation of D and I, and overrepresentation of post-infectious OD. This would be similar in olfactory function to our cluster 3 which also had an increased prevalence of URI-associated OD. Again, it is unclear if this cluster has peripheral causes of OD with preservation of most central function or not. Cluster 3 had higher rates of smoking, heart disease, and arthritis compared to normal subjects. It could be hypothesized that these factors contribute to peripheral inflammation in the nasal cavity and lead to the impairment of T, with relative preservation of suprathreshold function. Cluster 3 had moderate increases in nasal symptoms of blockage and drainage, consistent with possible peripheral factors such as smoking that contributed to the olfactory loss and impaired T.

Given our inclusion of the MMSE as a cluster-defining variable, we found a unique cluster 4. This group, which was 13% of our overall cohort, has the lowest MMSE. While the MMSE is simply a screening tool, this group may have central/cognitive factors associated with their olfactory loss. When patients are cognitively impaired, it is often difficult to predict exactly which aspects of an olfactory psychophysical test will be impacted. Prior reports have suggested that suprathreshold tests, I, in particular, are impacted to a greater degree than T.¹¹ In our study, we found mild impairment of D and I, with some preservation of T compared to normosmics. Whether this indicates normal nasal cavity and olfactory neuroepithelium function with impaired central processing remains an area for further study. This group also had higher rates of the Black race, diabetes, hypertension, and history of OD associated with URI or trauma. Given the exploratory nature of this preliminary study, it is uncertain which of these factors are independent. Black individuals are known to have higher rates of diabetes mellitus and hypertension, but given the size of our study, we are unable to control for these multiple variables. Prior studies have shown that MMSE scores are lower in Black and Asian individuals.¹² Thus whether this association is due to true variation in rates of dementia or the potential impact of race, education or socioeconomic status upon test scores requires further study. Cluster 4 subjects had essentially no nasal symptoms and normal scores in olfactory VAS symptoms and QOD. Assuming this was not due to cognitive influences, this indicates a lack of

awareness/impact of their olfactory loss upon daily activities and suggests that this group would be unlikely to seek medical attention. Thus improved screening strategies are needed to identify these individuals in the general population. In addition, it should be noted that MMSE is typically used to detect frank dementia. Thus future studies may consider more sensitive instruments to detect mild cognitive impairment.

The goal of clustering is to group individuals who have similar characteristics to determine unique etiologies and targeted therapies. This preliminary study showed 4 distinct groups of subjects from the community who varied by cognitive function, as well as by the severity and specific aspect of OD. These novel clusters do appear to have some utility in identifying potential causative factors and expand upon prior studies using machine learning to identify unique clinical phenotypes. Further studies are needed using biomarkers, novel imaging, and other methods to determine precise mechanisms of OD that occur in each of these groups and will enable clinicians to develop targeted therapies. Others have used cluster analysis based upon gustatory function, T, D, I, age of OD onset, and family history of Parkinson's disease,¹³ thus machine-based learning based upon chemosensory function holds promise and future clinical utility well beyond simple classifications that currently exist.

Clustering may be useful in directing subsequent treatment strategies, such as nasal surgery, nasal steroids, or olfactory training. This further supports our findings of the varying impact of unique clusters or groups of OD upon QOL and should be investigated in future studies.

Declaration of Conflicting Interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institute on Deafness and Other Communication Disorders (NIDCD), one of the National Institutes of Health, Bethesda MD (5R01DC019078-01A1; Principal Investigator: RJS, Co-Investigators: JRD, MAE, AMB, MG, VR, and ZMS).

ORCID iDs

Rodney J. Schlosser  <https://orcid.org/0000-0001-6480-0275>

Judy R. Dubno  <https://orcid.org/0000-0003-2340-4721>

Mark A. Eckert  <https://orcid.org/0000-0001-9286-7717>

References

- Schubert CR, Cruickshanks KJ, Murphy C, et al. Olfactory impairment in adults. *Ann N Y Acad Sci*. 2009;1170(1):531–536.
- Hackett K, Krikorian R, Giovannetti T, et al. Utility of the NIH Toolbox for assessment of prodromal Alzheimer's disease and dementia. *Alzheimers Dement (Amst)*. 2018;10(1):764–772.
- Oleszkiewicz A, Schriever VA, Croy I, Hähner A, Hummel T. Updated Sniffin' Sticks normative data based on an extended sample of 9139 subjects. *Eur Arch Otorhinolaryngol*. 2019;276(3):719–728.
- Hummel T, Whitcroft KL, Andrews P, et al. Position paper on olfactory dysfunction. *Rhinol Suppl*. 2017;54(26):1–30.
- Grossberg GT, Beck D, Zaidi SNY. Rapid depression assessment in geriatric patients. *Clin Geriatr Med*. 2017;33(3):383–391.
- Ge L, Yap CW, Ong R, Heng HB. Social isolation, loneliness and their relationships with depressive symptoms: a population-based study. *PLoS One*. 2017 Aug 23;12(8):e0182145.
- Lotsch J, Kringel D, Hummel T. Machine learning in human olfactory research. *Chem Senses*. 2019;44(1):11–22.
- Lotsch J, Hummel T, Ultsch A. Machine-learned pattern identification in olfactory subtest results. *Sci Rep*. 2016;6(1):35688.
- Zou LQ, et al. Association between olfactory function and quality of life in patients with olfactory disorders: a multicenter study in over 760 participants. *Rhinology*. 2021;59(2):164–172.
- Oleszkiewicz A, Park D, Resler K, et al. Quality of life in patients with olfactory loss is better predicted by flavor identification than by orthonasal olfactory function. *Chem Senses*. 2019;44(6):371–377.
- Wang Q, Chen B, Zhong X, et al. Olfactory dysfunction is already present with subjective cognitive decline and deepens with disease severity in the Alzheimer's disease Spectrum. *J Alzheimers Dis*. 2021;79(2):585–595.
- Mukadam N, Lewis G, Mueller C, Werbeloff N, Stewart R, Livingston G. Ethnic differences in cognition and age in people diagnosed with dementia: a study of electronic health records in two large mental healthcare providers. *Int J Geriatr Psychiatry*. 2019;34(3):504–510.
- Lotsch J, Haehner A, Hummel T. Machine-learning-derived rules set excludes risk of Parkinson's disease in patients with olfactory or gustatory symptoms with high accuracy. *J Neurol*. 2020;267(2):469–478.