Unsupervised Machine Learning Identified Distinct Population Clusters Based on Symptoms of Oral Pain, Psychological Distress, and Sleep Problems

Nontawat Chuinsiri

Institute of Dentistry, Suranaree University of Technology, Nakhon Ratchasima, Thailand

ABSTRACT

Objectives: The aims of this study were to explore the use of unsupervised machine learning in clustering the population based on reports of oral pain, psychological distress, and sleep problems and to compare demographic and socio-economic characteristics as well as levels of functional domains (work, social, and leisure) between clusters. **Materials and Methods:** In this cross-sectional study, a total of 1613 participants from the National Health and Nutrition Examination Survey in 2017–2018 were analyzed. Five variables, including oral pain, depression, anxiety, sleep apnea, and excessive daytime sleepiness, were selected for cluster analysis using the k-medoids clustering algorithm. The distribution of categorical variables between clusters was assessed using χ^2 test. One-way analysis of variance and Kruskal–Wallis H test were used to compare numerical variables as appropriate. **Results:** Five distinct clusters were identified: healthy, norm, anxiety, apnea-comorbid, and pain-comorbid. The apnea-comorbid cluster had mean age of 59 years and higher proportion of men. The pain-comorbid cluster had mean age of 56 years and higher proportion of women. Whites constituted a majority of both comorbid clusters. The pain-comorbid cluster demonstrated the least percentage of individuals with college degree, the lowest income, and significant impairment in all functional domains. **Conclusion:** Through the use of unsupervised machine learning, the clusters with comorbidity of oral pain, psychological distress, and sleep problems have emerged. Major characteristics of the comorbid clusters included mean age below 60 years, White, and low levels of education and income. Functional domains were significantly impaired. The comorbid clusters thus call for public health intervention.

KEYWORDS: *Anxiety, depression, intrinsic sleep disorders, pain, unsupervised machine learning*

INTRODUCTION

Pain is one of the most common complaints prompting patients to seek dental care.[1] The prevalence of orofacial pain (OFP) based on community surveys varied between countries, ranging from 4% to 57%.[2] The prevalence increased over the years and 22% of those with first onset pain progressed toward persistent pain.[3] Significant impact of OFP on quality of life was also previously described.[4]

OFP is often accompanied with high levels of psychological distress, depression, and anxiety.[4,5] Individuals with depressive and anxiety symptoms are at increased risk of developing OFP and vice

Address for correspondence: Dr Nontawat Chuinsiri, Institute of Dentistry, Suranaree University of Technology, 111 Mahawittayalai road, Tumbon Suranaree, District Mueang, Nakhon Ratchasima 30000, Thailand. E-mail: cnontawat@sut.ac.th

Access this article online	
Quick Response Code:	**Website:** www.jispcd.org
	DOI: 10.4103/jispcd.JISPCD_131_21

How to cite this article: Chuinsiri N. Unsupervised machine learning identified distinct population clusters based on symptoms of oral pain, psychological distress, and sleep problems. J Int Soc Prevent Communit Dent 2021;11:531-8.

versa.[6,7] Whereas pain has clear negative impact on social activities,[8] the influence of depression on social behavior is complex.[9] OFP patients also present with sleep problems, sleep apnea, and excessive daytime sleepiness (EDS).[10-12] These sleep-related issues have significant burden on several quality-of-life domains including daily life functioning and social life.[13] In addition, sleep apnea showed significant associations with depression and anxiety.[14,15]

Since current evidence has pointed toward the strong associations among OFP, psychological distress, and impaired sleep health,[10,14] it is hypothesized that a number of individuals are living with comorbidity of the aforementioned symptoms and experiencing much difficulty in work, social, and leisure functioning. Machine learning has shown promises in identifying hidden relationships between variables as well as extracting distinct patterns from unlabeled data.[16] The aims of this study were to explore the use of unsupervised machine learning method, the k-medoids clustering, in partitioning the population based on the report of oral pain, psychological distress, and sleep problems and to compare the demographic and socio-economic characteristics as well as levels of work/social/leisure activities between clusters.

MATERIALS AND METHODS

This is a cross-sectional study and complies with strengthening the reporting of observational studies in epidemiology guidelines.[17]

DATASET

This study was based on the publicly available dataset collected during the National Health and Nutrition Examination Survey (NHANES).[18] Thus, ethical approval from an Institutional Review Board was not needed. Briefly, the NHANES is a yearly survey targeting non-institutionalized civilian residents of the USA and aims to determine the prevalence of major diseases and their risk factors. In 2017–2018, 9254 participants completed the NHANES interview.[18]

VARIABLES

Individual participants fulfilling inclusion criteria for analyses in the present study were those who (1) were 30–79 years of age, (2) provided basic demographic and socio-economic data including sex, race/ethnicity, educational level, and ratio of family income to poverty, and (3) completed questionnaires assessing oral health-related problems, psychological characteristics, sleep problems, and functioning in work, social, and leisure domains. Participants who refused to answer the questionnaire items or responded as "don't know" were excluded.

Oral health-related problems were determined by assessing the frequency of (1) oral pain and (2) difficulty doing usual jobs due to oral problems in the past year. Participants answered by choosing the following: never (0), hardly ever (1), occasionally (2), fairly often (3), and very often (4). Numbers in the parentheses indicate the score given to each answer.

Psychological characteristics were examined based on levels of depression and anxiety. The NHANES utilized Patient Health Questionnaire-9 (PHQ-9) to assess the degree of depression. Each item can be scored from 0 (not at all) to 3 (nearly every day). The total PHQ-9 score is categorized into five groups: minimal (0–4), mild (5–9), moderate (10–14), moderately severe (15–19), and severe (20–27).[19] Anxiety was assessed by asking a question on how often participants felt worried, nervous, or anxious. Possible responses included: never (0), a few times a year (1), monthly (2), weekly (3), and daily (4).

Sleep problems included the frequency of (1) sleep apnea and (2) EDS. Participants were asked how often they snorted, gasped, or stopped breathing while asleep. Responses to the question included: never (0), rarely—1–2 nights/week (1), occasionally—3–4 nights/week (2), and frequently—5 or more nights/week (3). EDS was assessed on how often participants felt overly sleepy during the day. Responses were never (0), rarely—1 time/month (1), sometimes—2–4 times/month (2), often—5–15 times/month (3), and almost always—16–30 times/month (4).

Regarding work activities, participants were asked whether a physical, mental, or emotional problem (1) kept them from working at a job and (2) limited the kind or the amount of work they could do. Responses to these items were either "yes" or "no." Participants were also asked to rate their difficulties in (1) participating in social activities and (2) doing things to relax at home. Possible responses included: no difficulty (1), some difficulty (2), much difficulty (3), and unable to do (4).

DATA PREPROCESSING

Python version 3.8 was used for preprocessing data and implementing the clustering algorithm. Five variables, including oral pain, depression, anxiety, sleep apnea, and EDS, were selected for cluster analysis. Since the variables differed in their score ranges, they were transformed to be within the domain [0,1] using the minimum–maximum scaling method.

K-MEDOIDS CLUSTERING

K-medoids, a variant of k-means, which is appropriate for variables in ordinal scale,[20] was chosen for clustering in the present study. The k-medoids algorithm from the

scikit-learn-extra module (https://scikit-learn-extra.readthedocs.io/en/latest/) was implemented on the five selected variables. Briefly, initial medoids representing each cluster were selected using the k-medoids++ approach. Each data point (also referred to as object) was assigned to its nearest medoid based on the Euclidean distance. The sum of the distances from all objects to their designated medoids was calculated. An object with the least total distance to others within the same cluster was chosen as a new medoid. The steps were then repeated by re-assigning each object to a new nearest medoid until the sum of the distances did not decrease further.[21]

The optimum number of clusters (k) was determined using the elbow method.[20] The k-medoids algorithm was run on different values of k, ranging from 1 to 15. The sum of the distances from all objects to their medoids (also referred to as inertia) was calculated for each k. The inertia was then plotted against respective k and the optimum value of k was determined graphically. The point of inflection on the graph where the inertia started to decrease slowly and linearly was chosen.

STATISTICAL ANALYSIS

Statistical comparisons between clusters were performed in IBM SPSS Statistics Version 25.0 (IBM Corp., Armonk, NY, USA). The distribution of categorical variables between clusters was assessed using Pearson's χ^2 test. One-way analysis of variance and Tukey's *post hoc* test were used to compare inter-cluster age and ratio of family income to poverty. To compare ordinal variables between clusters, the Kruskal–Wallis H test followed by multiple comparisons and adjustment of the significance level with the Dunn–Bonferroni test was performed. Statistical significance was considered at $P<0.05$.

RESULTS

CHARACTERISTICS OF THE INCLUDED PARTICIPANTS

A total of 1613 participants in the NHANES 2017–2018 cohort fulfilled the inclusion criteria. Seven

thousand six hundred forty-one participants had missing data and were excluded from further analyses. Slightly higher proportion of men was included. The mean age of all participants was 61.7 years. A majority of participants were non-Hispanic Whites. More than half had associate degrees or above. The mean ratio of family income to poverty was 2.5 ± 1.6 [Table 1]. The average levels of oral pain, depression, anxiety, and sleep apnea were minimal to none. EDS was sometimes experienced by the participants with the median of 2–4 times a month [Table 2]. Approximately 46% of the participants reported having problems limiting work capacity and 30.8% were unable to work regularly. Most participants had no difficulty doing jobs due to oral problems, attending social events, and enjoying home leisure activities [Table 3].

CLUSTER ANALYSIS

The elbow method indicated five as the most appropriate number of clusters [Figure 1]. Figure 2 gives overview cluster profiles based on the five symptoms (oral pain, depression, anxiety, sleep apnea, and EDS) used for clustering. Statistical comparisons of each symptom between clusters are summarized in Table 2.
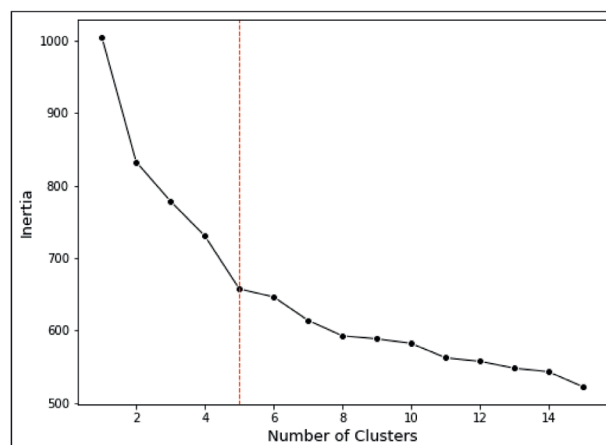

Figure 1: Visualization of the elbow method showing the optimal number of clusters (k=5)
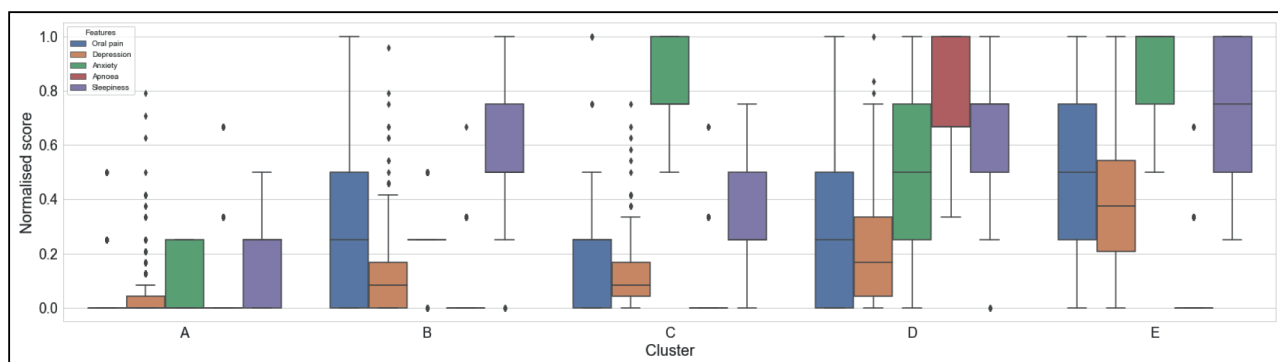

Figure 2: Comparisons of overall cluster profiles

Table 1: Comparisons of demographic and socio-economic characteristics

	Whole sample (n=1613)	Cluster A (n=397)	Cluster B (n=503)	Cluster C (n=233)	Cluster D (n=266)	Cluster E (n=214)	Statistics
Sex, n (% within cluster)							
Male	840 (52.1)	243 (61.2)	252 (50.1)	106 (45.5)	151 (56.8)	88 (41.1)	P <0.001
Female	773 (47.9)	154 (38.8)	251 (49.9)	127 (54.5)	115 (43.2)	126 (58.9)	χ^2 (4) = 30.74
Age, mean ± SD	61.7 ± 11.3	65.4 ± 7.96[a]	63.7 ± 9.9[a]	59.5 ± 12.9[b]	58.9 ± 12.1[b,c]	56.2 ± 13.1[c]	P <0.001 F(4) = 36.72
Race/ethnicity, n (%)							P <0.001 χ^2 (20) = 66.6
Mexican American	155 (9.6)	42 (10.6)	49 (9.7)	25 (10.7)	22 (8.3)	17 (7.9)	
Other Hispanic	156 (9.7)	46 (11.6)	46 (9.1)	21 (9)	25 (9.4)	18 (8.4)	
Non-Hispanic White	652 (40.4)	113 (28.5)	202 (40.2)	105 (45.1)	118 (44.4)	114 (53.3)	
Non-Hispanic Black	389 (24.1)	108 (27.2)	129 (25.6)	48 (20.6)	65 (24.4)	39 (18.2)	
Non-Hispanic Asian	167 (10.4)	66 (16.6)	51 (10.1)	22 (9.4)	21 (7.9)	7 (3.3)	
Other races	94 (5.8)	22 (5.5)	26 (5.2)	12 (5.2)	15 (5.6)	19 (8.9)	
Educational level, n (%)							P <0.001 χ^2 (16) = 46.57
Less than 9th grade	134 (8.3)	47 (11.8)	37 (7.4)	19 (8.2)	15 (5.6)	16 (7.5)	
9th–11th grade	177 (11)	38 (9.6)	50 (9.9)	24 (10.3)	31 (11.7)	34 (15.9)	
High school graduate	409 (25.4)	101 (25.4)	128 (25.4)	50 (21.5)	70 (26.3)	60 (28)	
Associate degree	544 (33.7)	102 (25.7)	185 (36.8)	84 (36.1)	92 (34.6)	81 (37.9)	
College graduate or above	349 (21.6)	109 (27.5)	103 (20.5)	56 (24)	58 (21.8)	23 (10.7)	
Ratio of family income to poverty, mean ± SD	2.5 ± 1.6	2.8 ± 1.6[d]	2.6 ± 1.6[d]	2.6 ± 1.6[d]	2.4 ± 1.6[d]	1.8 ± 1.3	P <0.001 F(4) = 13.3

SD = standard deviation. The same superscript letters indicate no statistical differences between clusters

Table 2: Comparisons of oral pain, psychological distress, and sleep problems

	Whole sample	Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Statistics
Oral pain							P<0.001 H(4) = 395.5
Median (IQR)	1 (0–2)	0 (0-0)	1 (0–2)[a]	1 (0–1)	1 (0–2)[a]	2 (1–3)	
10th–90th percentiles	0–2	0–1	0–2	0–2	0–3	0–4	
Depression (PHQ-9)							P <0.001 H(4) = 463.8
Median (IQR)	2 (0–6)	0 (0–1)	2 (0–4)[b]	2 (0.5–4)[b]	4 (1–8)	9 (5–13)	
10th–90th percentiles	0–10	0–4	0–7.6	0–8	0–13	3–17.5	
Anxiety							P <0.001 H(4) = 1078.2
Median (IQR)	1 (0.5–3)	0 (0–1)	1 (1-1)	3 (3–4)	2 (1–3)	4 (3–4)	
10th–90th percentiles	0–4	0–1	0–2	2–4	0–4	3–4	
Sleep apnea							P <0.001 H(4) = 968.5
Median (IQR)	0 (0–1)	0 (0-0)[c,d]	0 (0-0)[d]	0 (0-0)[c,d]	2 (2–3)	0 (0-0)[c]	
10th–90th percentiles	0–2	0–1	0–1	0–1	1–3	0–1	
Daytime sleepiness							P <0.001 H(4) = 796.3
Median (IQR)	2 (1–3)	0 (0–1)	2 (2–3)[a]	1 (0.5–2)[b]	3 (2–3)	3 (2–4)	
10th–90th percentiles	0–3	0–2	1–3	0–2	1–4	2–4	

IQR = interquartile range. The same superscript letters indicate no statistical differences between clusters

Cluster A (24.6%) showed low scores in all variables and was named the healthy cluster. Cluster B (31.2%) displayed characteristics similar to the whole sample and was named the norm cluster. Compared with the healthy cluster, the norm cluster had significantly higher levels of oral pain, depression, anxiety, and EDS (P<0.001). Cluster C (14.4%) had isolated high anxiety level and was named the anxiety cluster. Cluster D (16.5%) reported the most frequent sleep apnea episodes (P<0.001) with moderate levels of other comorbid conditions and was named the apnea-comorbid cluster. Despite having a similar level of oral pain to the norm cluster, the apnea-comorbid cluster expressed significantly higher levels of depression, anxiety, and EDS (P<0.001). Cluster E (13.3%) reported the highest level of oral pain (P<0.001) along with depression (P<0.001), anxiety (P<0.05), and EDS (P<0.05) and was named the pain-comorbid cluster.

Table 3: Comparisons of functioning in work/social/leisure domains

	Whole sample	Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Statistics
Unable to work, *n* (%)							*P* <0.001
Yes	497 (30.8)	60 (15.1)	130 (25.8)	75 (32.2)	96 (36.1)	136 (63.6)	χ^2 (4) = 163
No	1116 (69.2)	337 (84.9)	373 (74.2)	158 (67.8)	170 (63.9)	78 (36.4)	
Limited in amount of work, *n* (%)							*P* <0.001
Yes	752 (46.6)	109 (27.5)	213 (42.3)	111 (47.6)	155 (58.3)	164 (76.6)	χ^2 (4) = 154.4
No	861 (53.4)	288 (72.5)	290 (57.7)	122 (52.4)	111 (41.7)	50 (23.4)	
Difficulty doing job due to oral problems							*P* <0.001
							H(4) = 122.4
Median (IQR)	0 (0-0)	0 (0-0)[a]	0 (0-0)[b]	0 (0-0)[a,b]	0 (0-0)[b]	0 (0-1)	
10th–90th percentiles	0–1	0-0	0–1	0–1	0–1	0–2.5	
Difficulty attending social event							*P* <0.001
							H(4) = 211.1
Median (IQR)	1 (1-1)	1 (1-1)	1 (1-1)[c]	1 (1-1)[c]	1 (1–2)	2 (1–2)	
10th–90th percentiles	1–2	1-1	1–2	1–2	1–2	1–3	
Difficulty with home leisure activity							*P* <0.001
							H(4) = 75.1
Median (IQR)	1 (1-1)	1 (1-1)[d]	1 (1-1)[d,e]	1 (1-1)[d,e]	1 (1-1)[e]	1 (1–1.25)	
10th–90th percentiles	1-1	1-1	1-1	1-1	1–2	1–2	

IQR = interquartile range. The same superscript letters indicate no statistical differences between clusters

COMPARISONS OF DEMOGRAPHIC AND SOCIO-ECONOMIC CHARACTERISTICS

The anxiety, the apnea-comorbid, and the pain-comorbid clusters had significantly lower mean age than the healthy and the norm clusters (*P*<0.001). The χ^2 test of independence revealed significant difference in the distributions of sex, race/ethnicity, and educational levels between clusters (*P*<0.001). Higher proportion of men was observed in the healthy cluster and the apnea-comorbid cluster, whereas women were present in higher proportion among the anxiety cluster and the pain-comorbid cluster [Table 1].

Regarding race/ethnicity, 28.5% of the heathy cluster were Whites. The percentage was higher in the anxiety cluster (45.1%), the apnea-comorbid cluster (44.4%), and the pain-comorbid cluster (53.3%). The proportions of Black and Asian individuals in each cluster were opposite to Whites. The highest percentage of Black (27.2%) and Asian (16.6%) individuals was observed in the healthy cluster, whereas the lowest was seen in the pain-comorbid cluster (18.2% and 3.3%, respectively) [Table 1].

The highest percentage of individuals with college degree or above was observed in the healthy cluster (27.5%), while such individuals constituted about one-tenth of the pain-comorbid cluster. In addition, the pain-comorbid cluster comprised the highest proportion of individuals who had not finished high school (23.4%). The ratio of family income to poverty was significantly lower in the pain-comorbid cluster, compared with others (*P*<0.001) [Table 1].

COMPARISONS OF DIFFICULTY LEVELS IN PERFORMING WORK, SOCIAL, AND LEISURE ACTIVITIES

The proportions of individuals with functional limitations regarding work differed between clusters (*P*<0.001). Approximately 15% of the healthy cluster reported that they were unable to work due to a physical, mental, or emotional problem. The proportions of such individuals in the anxiety and the apnea-comorbid clusters were more than twice as seen in the healthy cluster. The highest percentage (63.6%) was observed in the pain-comorbid cluster. Slightly over three-fourth in the pain-comorbid cluster and more than half of the apnea-comorbid cluster also reported limitations in the amount of work they could do. The degree of difficulty in doing jobs due to oral problems was low in all clusters. Nevertheless, the pain-comorbid cluster reported significantly higher difficulty than others (*P*<0.001). The pain-comorbid cluster had significantly higher difficulties in attending social events and participating in home leisure activities than the rest of the clusters (*P*<0.001). Compared with the norm cluster, the apnea-comorbid cluster showed significantly greater difficulty in attending social event (*P*<0.001), but not with home leisure activities [Table 3].

DISCUSSION

This is the first study that employed unsupervised machine learning to identify patterns of symptoms regarding oral pain, psychological distress, and sleep problems among the US nationally representative sample. These symptoms have been previously shown to be coexistent.[4,5,10,22] In the present study, the *k*-medoids algorithm partitioned the participants into five distinct clusters, confirming the associations among such symptoms.

COMORBIDITY OF ORAL PAIN, PSYCHOLOGICAL DISTRESS, AND SLEEP PROBLEMS

EDS is not uncommon as the study participants had experienced EDS 2–4 times/month on average. This

is supported by a previous study reporting that only three-tenth of the U.S. population had appropriate sleep duration.[23] In the present study, the degree of oral pain was found parallel to that of EDS as can be seen across all clusters. The finding aligns with a previous study showing that individuals with pain had poorer sleep quality and more severe daytime dysfunction than healthy subjects.[10] Another population-based longitudinal study also reported persistent pain as an independent predictor of EDS onset.[24]

Sleep apnea was most frequently reported in the apnea-comorbid cluster, whereas it was almost never reported in other clusters. The apnea-comorbid cluster also manifested varying degrees of oral pain, psychological distress, and EDS. Such observation supports the theory that sleep apnea might pose as a risk factor of these comorbid symptoms. This is in accord with findings from a large cohort study showing that poor sleep quality preceded pain.[22] It has been suggested that sleep apnea is associated with a cascade of events involving sleep deprivation, intermittent hypoxia, and generation of inflammatory mediators, resulting in pain hypersensitivity.[25,26]

The highest degrees of oral pain and psychological distress were evident in the pain-comorbid cluster, supporting previous findings that a strong association exists between various OFP and psychological distress.[4,6,10,27] Depression shares common pathophysiology with pain, involving neurotransmitters, noradrenaline, and serotonin. The exact relationship between depression and pain is still equivocal as to whether depressed individuals are prone to develop pain or vice versa.[28] EDS severity was also highest in the pain-comorbid cluster. In addition to pain, depression and anxiety were previously reported as strong predictors of EDS.[29]

Through the use of unsupervised machine learning, individuals with isolated high anxiety were identified. As anxiety was previously shown as a risk factor of pain,[30] high degree of pain was expected in this cluster. Surprisingly, the anxiety cluster presented with minimal degree of oral pain. Such observation is likely explained by a previous study showing that persistent OFP was associated with trait anxiety but not state anxiety.[31] Hence, it can be speculated that the anxiety cluster predominantly included individuals with state anxiety, whereas those with trait anxiety might make up the majority of the pain-comorbid cluster. This hypothesis is further supported by recent studies showing that trait anxiety was associated with depression and poor sleep quality.[32,33]

DEMOGRAPHIC AND SOCIO-ECONOMIC CHARACTERISTICS

Higher proportion of women was present in the anxiety and the pain-comorbid clusters, which is in line with findings from previous studies.[6,14,34] Women demonstrated lower pain threshold[35] and greater likelihood of developing depression and anxiety.[36,37] Earlier findings found that sleep apnea had a male predilection,[22,38] complying with the present study in which men dominated the apnea-comorbid cluster.

The pain-comorbid cluster had significant lower mean age than the healthy cluster. Our finding is thus in agreement with previous studies, showing that the prevalence of OFP declined with increasing age.[2,39]

The highest proportion of Whites was seen in the pain-comorbid cluster, whereas the lowest was in the healthy cluster. The reverse was observed for Asians, whose proportion was 5-fold in the healthy cluster, compared with the pain-comorbid cluster. Controversy exists regarding the cross-cultural differences in pain perception. In previous studies, Asians were reported to have greater pain sensitivity than Whites.[40,41] Such association was mediated via psychological factors as Asians also demonstrated higher levels of depression and anxiety. Participants in the previous study also had a lower age range, 18–44 years.[41] In addition, Asian youth reported higher parental pressure, which was correlated with psychological distress.[40] More research is needed to explore the levels of pain and psychological distress among Asians at different stages of life.

Regarding the socio-economic status, the pain-comorbid cluster had less proportion of individuals with college degrees and lower mean ratio of family income to poverty than other clusters. This finding emphasizes the association among OFP, psychological distress, and low socio-economic status as shown in the previous studies.[30,42,43] Therefore, policy should be focussed on this cohort in order to improve public health as a whole and prevent subsequent burden on healthcare systems.

WORK, SOCIAL, AND HOME LEISURE ACTIVITIES

In the present study, the pain-comorbid cluster had significantly higher severity in regard to work, social, and home leisure activities than other clusters. Hence, the null hypothesis was rejected.

The percentage of individuals with work incapacity/limitation was lowest in the healthy cluster and increased as the severity of reported symptoms increased. Difficulty in doing job due to oral problems was minimal in all clusters; the same was observed for difficulty attending social events and home leisure activities. Still, the pain-comorbid cluster presented with the highest levels of difficulties in all domains. Previous studies investigating OFP patients also found that individuals with higher severity of pain, psychological distress, and sleep problems had greater functional limitation as well

as physical and social disability than those with mild conditions.[10,11]

LIMITATIONS

Some limitations can be appreciated from the present study. First, all variables used for clustering are to be regarded as symptoms rather than clinical diagnoses. Secondly, the types and origin of OFP could not be determined due to the non-specific questionnaire used in the NHANES. Such detailed classification, however, is not feasible for large survey studies. The heterogeneity of OFP may thus pose as a confounding factor. Psychological distress and sleep-related problems were reported to be associated with certain subtypes of OFP,[27] but this relationship may not be existent for other subtypes. Findings with non-significant association, which are often not reported, are therefore important and should be encouraged. In the present study, the answers to the questionnaire were based on the frequency of pain and might at best be interpreted that those who scored high were likely to have persistent pain. Our finding is thus logical and in line with prior studies, confirming robust associations among persistent pain, psychological distress, and sleep problems.[4,6,12] Thirdly, based on the cross-sectional study design, definite causality between oral pain and other comorbidities cannot be concluded.

CONCLUSION

Through the use of unsupervised machine learning, the clusters with comorbidity of oral pain, psychological distress, and sleep problems have emerged. The major characteristics of the comorbid clusters included mean age below 60 years, White, and low levels of education and family income. Work, social, and leisure activities were also significantly impaired. In light of these findings, these individuals with oral pain and comorbid conditions necessitate public health intervention as their compromised productivity and social interaction could potentially lead to negative impacts not just on healthcare systems but also on economy.

CONFLICTS OF INTEREST

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

AUTHORS CONTRIBUTIONS

Nontawat Chuinsiri—Study conception, data retrieval and analysis, data interpretation, and manuscript writing.

ETHICAL POLICY AND INSTITUTIONAL REVIEW BOARD STATEMENT

This study was based on the publicly available dataset from the National Health and Nutrition Examination Survey 2017–2018, so ethical approval from an Institutional Review Board was not required.

PATIENT DECLARATION OF CONSENT

The U.S. Centre of Disease Control and Prevention's National Centre for Health Statistics Ethics Review Board approved the oral health data collection protocols, and all survey participants provided written informed consent (Protocol number: 2018-01, effective beginning October 26, 2017).

DATA AVAILABILITY STATEMENT

The dataset used in the current study is publicly available from the CDC/National Centre for Health Statistics at the following URL: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017.

REFERENCES

1. Deolia SG, Kela KS, Sawhney IM, Sonavane PA, Nimbulkar G, Reche A. Evaluation of oral health care seeking behavior in rural population of Central India. J Family Med Prim Care 2020;9:886-91.
2. Horst OV, Cunha-Cruz J, Zhou L, Manning W, Mancl L, DeRouen TA. Prevalence of pain in the orofacial regions in patients visiting general dentists in the Northwest Practice-Based Research Collaborative in Evidence-Based Dentistry research network. J Am Dent Assoc 2015;146:721-8.e3.
3. Häggman-Henrikson B, Liv P, Ilgunas A, Visscher CM, Lobbezoo F, Durham J, et al. Increasing gender differences in the prevalence and chronification of orofacial pain in the population. Pain 2020;161:1768-75.
4. Bäck K, Hakeberg M, Wide U, Hange D, Dahlström L. Orofacial pain and its relationship with oral health-related quality of life and psychological distress in middle-aged women. Acta Odontol Scand 2020;78:74-80.
5. Chuinsiri N, Jitprasertwong P. Prevalence of self-reported pain-related temporomandibular disorders and association with psychological distress in a dental clinic setting. J Int Med Res 2020;48:300060520951744.
6. Kim JY, Kim YS, Ko I, Kim DK. Association between burning mouth syndrome and the development of depression, anxiety, dementia, and Parkinson disease. JAMA Otolaryngol Head Neck Surg 2020;146:561-9.
7. List T, Jensen RH. Temporomandibular disorders: Old ideas and new concepts. Cephalalgia 2017;37:692-704.
8. Dueñas M, Ojeda B, Salazar A, Mico JA, Failde I. A review of chronic pain impact on patients, their social environment and the health care system. J Pain Res 2016;9:457-67.
9. Steger MF, Kashdan TB. Depression and everyday social activity, belonging, and well-being. J Couns Psychol 2009;56:289-300.

10. Natu VP, Yap AU, Su MH, Irfan Ali NM, Ansari A. Temporomandibular disorder symptoms and their association with quality of life, emotional states and sleep quality in South-East Asian youths. J Oral Rehabil 2018;45:756-63.

11. Smith MT, Wickwire EM, Grace EG, Edwards RR, Buenaver LF, Peterson S, *et al*. Sleep disorders and their association with laboratory pain sensitivity in temporomandibular joint disorder. Sleep 2009;32:779-90.

12. Almoznino G, Benoliel R, Sharav Y, Haviv Y. Sleep disorders and chronic craniofacial pain: Characteristics and management possibilities. Sleep Med Rev 2017;33:39-50.

13. Waldman LT, Parthasarathy S, Villa KF, Bron M, Bujanover S, Brod M. Understanding the burden of illness of excessive daytime sleepiness associated with obstructive sleep apnea: A qualitative study. Health Qual Life Outcomes 2020;18:128.

14. Kmeid E, Nacouzi M, Hallit S, Rohayem Z. Prevalence of temporomandibular joint disorder in the Lebanese population, and its association with depression, anxiety, and stress. Head Face Med 2020;16:19.

15. Garbarino S, Bardwell WA, Guglielmi O, Chiorri C, Bonanni E, Magnavita N. Association of anxiety and depression in obstructive sleep apnea patients: A systematic review and meta-analysis. Behav Sleep Med 2020;18:35-57.

16. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: A practical introduction. BMC Med Res Methodol 2019;19:64.

17. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. BMJ 2007;335:806-8.

18. National Center for Health Statistics. National Health and Nutrition Examination Survey: NHANES 2017-2018 2020. Available from: https://www.cdc.gov/nchs/nhanes/index.htm. [Last accessed 2020 Dec 30].

19. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001;16:606-13.

20. Lee M, Lee S, Park J, Seo S. Clustering and characterization of the lactation curves of dairy cows using K-medoids clustering algorithm. Animals (Basel) 2020;10:1-14.

21. Park H-S, Jun C-H. A simple and fast algorithm for K-medoids clustering. Exp Syst Appl 2009;36:3336-41.

22. Wu JH, Lee KT, Kuo CY, Cheng CH, Chiu JY, Hung JY, *et al*. The association between temporomandibular disorder and sleep apnea—A nationwide population-based cohort study. Int J Environ Res Public Health 2020;17:1-13.

23. Cash RE, Beverly Hery CM, Panchal AR, Bower JK. Association between sleep duration and ideal cardiovascular health among US adults, National Health and Nutrition Examination Survey, 2013-2016. Prev Chronic Dis 2020;17:E43.

24. Jaussent I, Morin CM, Ivers H, Dauvilliers Y. Incidence, worsening and risk factors of daytime sleepiness in a population-based 5-year longitudinal study. Sci Rep 2017;7:1372.

25. Ryan S, McNicholas WT. Intermittent hypoxia and activation of inflammatory molecular pathways in OSAS. Arch Physiol Biochem 2008;114:261-6.

26. Charokopos A, Card ME, Gunderson C, Steffens C, Bastian LA. The association of obstructive sleep apnea and pain outcomes in adults: A systematic review. Pain Med 2018;19:69-75.

27. Alrashdan MS, Alkhader M. Psychological factors in oral mucosal and orofacial pain conditions. Eur J Dent 2017;11:548-52.

28. Sheng J, Liu S, Wang Y, Cui R, Zhang X. The link between depression and chronic pain: Neural mechanisms in the brain. Neural Plast 2017;2017:9724371.

29. Chen L, Luo C, Liu S, Chen W, Liu Y, Li Y, *et al*. Excessive daytime sleepiness in general hospital nurses: Prevalence, correlates, and its association with adverse events. Sleep Breath 2019;23:209-16.

30. Kindler S, Samietz S, Houshmand M, Grabe HJ, Bernhardt O, Biffar R, *et al*. Depressive and anxiety symptoms as risk factors for temporomandibular joint pain: A prospective cohort study in the general population. J Pain 2012;13:1188-97.

31. Honda Y, Handa T, Fukuda KI, Koukita Y, Ichinohe T. Comparison of risk factors in patients with acute and chronic orofacial pain. Anesth Prog 2018;65:162-7.

32. Labad J, Salvat-Pujol N, Armario A, Cabezas Á, Arriba-Arnau A, Nadal R, *et al*. The role of sleep quality, trait anxiety and hypothalamic-pituitary-adrenal axis measures in cognitive abilities of healthy individuals. Int J Environ Res Public Health 2020;17:1-18.

33. Wang T, Li M, Xu S, Liu B, Wu T, Lu F, *et al*. Relations between trait anxiety and depression: A mediated moderation model. J Affect Disord 2019;244:217-22.

34. Reiter S, Emodi-Perlman A, Goldsmith C, Friedman-Rubin P, Winocur E. Comorbidity between depression and anxiety in patients with temporomandibular disorders according to the research diagnostic criteria for temporomandibular disorders. J Oral Facial Pain Headache 2015;29:135-43.

35. Ostrom C, Bair E, Maixner W, Dubner R, Fillingim RB, Ohrbach R, *et al*. Demographic predictors of pain sensitivity: Results from the OPPERA study. J Pain 2017;18:295-307.

36. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. Psychol Bull 2017;143:783-822.

37. Asher M, Aderka IM. Gender differences in social anxiety disorder. J Clin Psychol 2018;74:1730-41.

38. Rezaeitalab F, Moharrari F, Saberi S, Asadpour H, Rezaeetalab F. The correlation of anxiety and depression with obstructive sleep apnea syndrome. J Res Med Sci 2014;19:205-10.

39. Macfarlane TV, Blinkhorn AS, Davies RM, Kincey J, Worthington HV. Oro-facial pain in the community: Prevalence and associated impact. Community Dent Oral Epidemiol 2002;30:52-60.

40. Rowell LN, Mechlin B, Ji E, Addamo M, Girdler SS. Asians differ from non-Hispanic Whites in experimental pain sensitivity. Eur J Pain 2011;15:764-71.

41. Kim HJ, Greenspan JD, Ohrbach R, Fillingim RB, Maixner W, Renn CL, *et al*. Racial/ethnic differences in experimental pain sensitivity and associated factors—Cardiovascular responsiveness and psychological status. PLoS One 2019;14:e0215534.

42. Joury E, Bernabe E, Gallagher JE, Marcenes W. Burden of orofacial pain in a socially deprived and culturally diverse area of the United Kingdom. Pain 2018;159:1235-43.

43. Srivastava S, Purkayastha N, Chaurasia H, Muhammad T. Socioeconomic inequality in psychological distress among older adults in India: A decomposition analysis. BMC Psychiatry 2021;21:179.