Community Dentistry and Oral Epidemiology

WILEY

COMMUNITY DENTISTRY AND ORAL EPIDEMIOLOGY

ORIGINAL ARTICLE OPEN ACCESS

Using Propensity Score Subclassification to Estimate the Population-Average Causal Effect of Temporomandibular Dysfunction Experience on Oral Health-Related Quality of Life Among Australian Adults

Kamal Hanna[1] | Ninuk Hariyani[1,2] | Gloria Mejia[1] | Lisa Jamieson[1] | David S. Brennan[1]

[1]Australian Research Centre for Population Oral Health (ARCPOH), Adelaide Dental School, The University of Adelaide, Adelaide, South Australia, Australia | [2]Department of Dental Public Health, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

Correspondence: Kamal Hanna (kamal.hanna@adelaide.edu.au)

Received: 23 January 2024 | **Revised:** 15 January 2025 | **Accepted:** 17 January 2025

Funding: Data collection of the National Survey of Adult Oral Health (NSAOH) 2004–06 wave has received support for data collection from the National Health and Medical Research Council (NHMRC) via Project Grant #299060. Additionally, the NSAOH data collection has received support from the Australian Government Department of Health and Ageing, the Australian Institute of Health and Welfare (AIHW), the Australian Dental Association (ADA), Colgate Oral Care, and the Centre for Disease Control and Prevention—United States of America.

Keywords: TMDAustralia | causal inference | National Survey of adult Oral health | OHRQoL | oral health-related quality of life | PATE | population average treatment effect | propensity score subclassification | temporomandibular dysfunction

ABSTRACT

Background: Temporomandibular dysfunction (TMD) experience might impair oral health-related quality of life (OHRQoL). Causal inference using population-based cross-sectional data is challenging given the potential for bias. Propensity Score Subclassification (PS-Subclassification) provides a tool to mitigate confounding bias. The aim of this study was to estimate the Population-Average Treatment Effect (PATE) of having TMD experience among Australian adults on OHRQoL using PS-Subclassification and statistically estimated Minimally Important Differences (MID-S).

Method: Australia's National Survey of Adult Oral Health (NSAOH) 2004–06 data were used which included a Computer Assisted Telephone Interview, mailed questionnaire and oral epidemiological examination. Data included demographics, socioeconomics, caries experience (DMFT index), periodontitis, TMD experience using the TMD Diagnostic Criteria Question, the Oral Health Impact Profile (OHIP-14) and perceived stress. Analysis steps included: (1) generating propensity scores (PS) for TMD experience probability using causal model-derived confounders while incorporating survey design elements; (2) PS-Subclassification and weighting; (3) assessing common support and group balance and (4) estimating the PATE for TMD experience on OHIP-14 overall and domains scores using complex samples GLM.

Results: Of the 4063 NSAOH participants, 397 with TMD and 3656 without TMD were included in PS-Subclassification (all data were used) and shared common support for their PS and established adequate covariate balance (SMD < 0.2). Experiencing TMD had higher OHIP-14 total scores ($B = 3.498$, 95% CI: 2.218–4.778) with a small MID-S (Cohen's $F^2 = 0.03$). TMD experience impaired all OHIP-14 domains ($p < 0.05$) with physical pain and psychological domains among the highest impaired OHIP-14 domains with a small MID-S.

Abbreviations: ATE, average treatment effect; CATI, computer-assisted telephone interview; DMFT, decayed, missing and filled teeth; NSAOH, National Survey of Adult Oral Health; OHIP-14, oral health impact profile-short form; OHRQoL, oral health-related quality of life; PATE, population average treatment effect; PS, propensity scores; PSS-14, 14-Item Perceived Stress Scale; QoL, quality of life; SMD, standardised mean difference; TMD, temporomandibular dysfunction.

Conclusion: TMD experience impaired the overall OHRQoL measured by the OHIP-14 among Australian adults with a small MID-S. Physical pain and psychological domains were among the highest impaired OHRQoL domains with a small MID-S. Clinicians and policymakers might consider these findings to support TMD screening and patient-centred management.

1 | Introduction

Temporomandibular dysfunction (TMD) is a group of degenerative musculoskeletal disorders that affects the morphology and function of the masticatory system and it is estimated to affect almost 10% of the Australian adult population [1]. TMD patients often have one or more symptoms which might cause impairments in the biomedical and psychosocial aspects of how the individual perceives their well-being concerning oral health [2], which is referred to as Oral Health-related Quality of Life (OHRQoL) [3]. Several causal pathways might be proposed to explain how TMD might impair the multi-dimensional perception of the impact on their OHRQoL. Most TMD patients report orofacial pain which lowers the patient's perception of their OHRQoL [4, 5]. Severe orofacial pain was observed in another oral condition to impair daily functions including work or study [6]. Additionally, severe orofacial pain resulted in mood change [6], and consequently, having psychosocial impacts on OHRQoL. Another causal pathway for the perceived psychosocial impacts in TMD patients might be attributed to the audible click or grating [7]. Also, TMD patients often reported trismus and chewing difficulty as it is known to affect the individual's perception of their OHRQoL [8]. Such trismus and chewing difficulty in TMD patients might lead them to change their food consistency to overcome these issues or it might be part of the conservative TMD management [7], which might impair food enjoyment. Understanding the causal effect of TMD experience on the dimensions and overall OHRQoL (known as the Average Treatment Effect ATE) among Australian adults (known as population-average effect) might enable better assessment of clinical relevance and facilitate effective patient-centred management approach. Such understanding of the population-average causal effect of TMD experience on OHRQoL among Australian adults is currently limited in the available literature.

Estimating the population-average causal effect of TMD on OHRQoL using population-based data is challenging considering the potential confounding bias which might distort the causal relationship between TMD and OHRQoL. While study designs such as randomised controlled trials might balance allocation probability and distribution of confounders across exposure groups this approach may be unethical to apply under some circumstances [9]. Accordingly, researchers developed the propensity score (PS) methods to balancing the probability of being exposed as well as confounders in non-randomised observational studies [10, 11]. Population-based cohort studies that estimate the causal effect of TMD on OHRQoL are challenging as they require significant resources and costs associated with participants' recruitment and retention over some time with a sufficient sample size [12]. Therefore, researchers have increasingly used PS techniques with cross-sectional data to overcome the ethical considerations related to experimental study design or the significant cost/resources needed for cohort studies where it is reported to be successful in estimating the population-average causal effect [9, 13]. Among the PS techniques, the PS subclassification method is argued to estimate the average treatment effect that can be generalised to the target population with the least bias in the estimated effect when compared with other PS methods such as matching and weighting [11]. Such an average treatment effect generalizable to the target population is known as the population average treatment effect (PATE).

This study aimed to use the PS subclassification method to get an insight into estimating, among the Australian adult population, the PATE of having TMD experience on the overall OHRQoL and OHRQoL domain scores as well as statistically estimating the clinical relevance of such effects.

2 | Methods

2.1 | Determining Important Confounders Using a Knowledge-Based TMD and OHRQoL Causal Model

The initial step in getting an insight into the TMD population-average effect on OHRQoL was to determine confounders that were related to TMD and OHRQoL using a knowledge-based causal model rather than relying on confounders of convenience -an approach recommended by Petersen, van der Laan [12]. Accordingly, a TMD and OHRQoL model was developed using the R package 'dagitty' [14], based on previous publications [15–20], and research conducted by the research team [1, 21], which was presented in Figure 1.

2.2 | Data Source

To conduct this study, data were sourced from the National Survey of Adult Oral Health (NSAOH) 2004–06 wave [22]. The NSAOH is a three-staged random stratified clustered representative sample of Australia's adult residents aged 15 years or over who resided in a household with access to a telephone line that was listed in Australia's 'Electronic White Pages' [23], with their corresponding postcode. The study is a cross-sectional design where data were collected between July 2004 and September 2006 with more details can be found in Slade, Spencer & Roberts-Thomson [24]. The NSAOH data used in this study comprised three datasets: Computer Assisted Telephone Interview (CATI), Self-Complete Questionnaire and a Standardised Oral Epidemiological Examination.

2.3 | NSAOH Participants' Selection

The CATI interviewer randomly selected an adult aged 15 years or older present in the household and the potential participant was invited to join the study. Additionally, after participants completed the CATI interview, they were invited to participate in a mailed questionnaire and the oral epidemiological examination and if agreed were contacted to complete the other

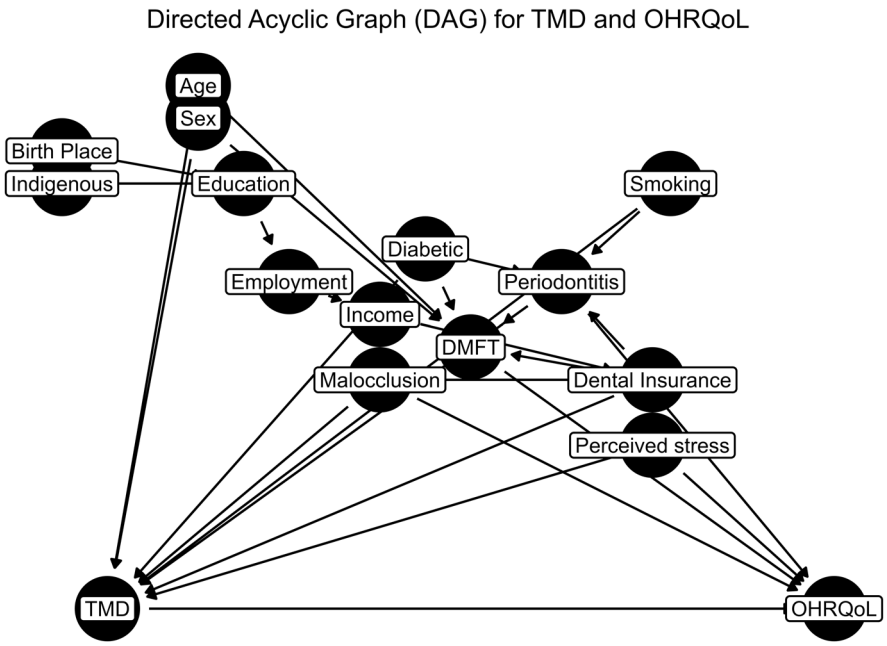

FIGURE 1 | Causal model for TMD experience (exposure) impact on OHRQoL (outcome) (DMFT, decayed, missing and filled teeth; OHRQoL, oral health-related quality of life; TMD, temporomandibular dysfunction).

NSAOH components [18, 25]. It is worth noting that participants who agreed to complete the oral epidemiological examination had to be assessed for eligibility to be included using a medical questionnaire as per the adopted oral examination protocol [26], such as rheumatic fever, endocarditis, bleeding disorders, joint replacement within the past 3-months, etc. In the analysis, this study included participants aged 18 years old or above with complete records across the three NSAOH datasets.

2.4 | Ethical Considerations

NSAOH has been granted ethical approval from the Low-Risk Human Research Ethics Committee, the University of Adelaide (approval number: H-001-2004). All NSAOH participants provided informed consent which was verbal for the CATI interview and written for the self-complete questionnaire and the standardised oral epidemiological examination. Further, NSAOH oral examiners appropriately referred participants with concerning medical or dental conditions to receive any necessary professional care.

2.5 | Variables Included in Our Analysis

2.5.1 | Outcome Variable: The Oral Health Impact Profile—Short Form (OHIP-14)

The OHIP-14 was included in the mailed self-complete questionnaire. The OHIP-14 is a 14-item self-reported questionnaire and it is a shorter version of the original 49-item OHIP. The OHIP-14 consists of seven domains with each domain represented by two items where the OHIP-14 domains are: functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap. Items for the OHIP-14 are scored on a 5-point scale ranging from 0 for 'Never' to 4 for 'Very often' with the reference period of the OHIP-14 as 'over the past 12 months'. OHIP-14 total scores range from 0 to 56 with a higher score meaning higher oral health impacts. While the OHIP-14 measures the adverse impacts of oral conditions on the individual's perceived wellbeing, it was used by Slade, Foy, Shugars & Phillips [27], as an OHRQoL measure. The OHIP-14 became a popular measure of OHRQoL and the most common OHRQoL measurement tool used to investigate TMD management modalities' impact on OHRQoL [2].

2.5.2 | Exposure Variable TMD Status Measured by the Diagnostic Criteria Question for TMD

The mailed self-complete questionnaire included the Diagnostic Criteria Question for TMD which was found to have satisfactory sensitivity and specificity in predicting clinical diagnosis for TMD [28]. The questionnaire consists of seven items in two domains: pain (three questions) and symptomatic TMD (four questions) which represent functional disturbance. The criterion adopted for the presence of TMD was recording a positive response for one or more of the pain items and a positive response in one or more of the symptomatic TMD items are in line with the clinical diagnosis of TMD used in previous research [29], and recommended by Sanders & Slade [19], and were used by researchers in recent studies [1, 21], where a confirmatory factor analysis was conducted and provided validity of its two-domain structure.

2.5.3 | Confounding Variables Derived From the Developed TMD and OHRQoL Model

2.5.3.1 | Demographics and Socio-Economic Confounders. The CATI interview collected information about

the participants' age (years), sex (male/female), indigenous status (yes/no), birthplace (Australia or overseas), annual total household income which was grouped at $<$\$60 000 or $\geq$ $<$\$60 000, have private dental insurance (yes/no), highest educational attainment which was grouped as University/College, Certificate/TAFE or Other and occupation which were grouped as professionals/admin/managerial, intermediate or manual workers.

2.5.3.2 | Oral Health Confounders. *Self-reported static malocclusion*: The mailed self-complete questionnaire asked the participants whether their 'Teeth feel do not fit properly together' and the response was yes/no. The study has used this question as a proxy for static malocclusion.

The number of decayed, missing and filled teeth (DMFT) as well as periodontitis status were obtained during the standardised dental examination which was conducted by 30 dentists who received training to comply with the oral examination protocol [30], with median inter-examiner reliability found to be 0.85. A self-illuminated dental mirror and a periodontal probe (with 2 mm marking) were used to conduct the oral examination.

The number of missing permanent teeth for any reason and not replaced by a removable or fixed prosthesis was recorded. To record the number of teeth with untreated decay, examiners used visual inspection with the aid of the dental mirror and periodontal probe where the latter was used to gently check the surface texture for the five surfaces of the present permanent teeth. The number of teeth with untreated decay was computed by summing the number of present teeth with untreated decay lesions on one or more of the examined surfaces. The number of filled teeth was obtained using visual inspection of the five surfaces of the present teeth. The number of filled teeth was computed by summing the number of present teeth with one filling or more.

To assess the periodontitis status, recession and probing depth were collected by oral examiners to compute clinical attachment loss (CAL). Measurements were obtained from three points: mesiobuccal, mid-buccal and distobuccal for all teeth present except for the third molars. Periodontitis status was assessed using the US Centre for Disease Control and Prevention case definition of periodontitis [31]. For this study, a binary variable (Yes/No) was created for the presence of 'moderate or severe periodontitis' where moderate periodontitis was defined as the presence of ≥ 2 sites of inter-proximal CAL of ≥ 4 mm, and severe periodontitis was defined as the presence of ≥ 2 sites of inter-proximal CAL of ≥ 6 mm.

2.5.3.3 | Health Behaviours and General Health Confounders. The CATI interview collected information on being a current smoker (yes/no) as well as if a doctor said you have diabetes (yes/no). Moreover, the mailed self-complete questionnaire included the 14-item Perceived Stress Scale which was developed by Cohen, Kamarck & Mermelstein [32], as a measure of subjective stress and consists of 14 items representing how unpredictable, uncontrollable, and overloaded participants feel in their life. The reference period of the PSS-14 is 'last month'. Items for the PSS-14 are scored on a 5-point Likert-like scale ranging from 0 for 'Never' to 4 for 'Very often'. The PSS-14 consists of two subscales: Perceived distress (negative subscale) and Perceived control (positive subscale). The perceived distress subscale is the sum score for items 1,2,3,8,11,12 and 14 where higher scores represent higher perceived distress. The perceived control subscale is the sum score of items 4,5,6,7,9,10 and 13 where a higher score represents higher perceived control (coping).

2.6 | Data Analysis

Data analysis was carried out using R version 4.4.1 [33], and RStudio version 2024.09.0 Build 375 [34], with the application of several key packages including the MatchIt package version 4.5.5 [35]. The PS subclassification and weighted analysis was conducted to estimate the effect of TMD experience on OHRQoL among the Australian adult population, as measured by the OHIP-14 total score and domain scores. NSAOH participants who completed the three components of the NSAOH were combined and initially included and then, rows with missing values in any of the covariates were excluded to ensure a complete dataset. The analysis included the following steps: (1) PS for TMD experience were estimated using a complex samples logistic regression model, incorporating the survey package. This regression model included the identified confounders, and interaction terms were used to capture the effect modification of having private dental insurance on oral health status covariates. The population weights, cluster and stratification elements of NSAOH were used in computing PS to account for the complex survey design and ensure the representativeness of the sample to the Australian adult population sampled by this dataset; (2) PS Subclassification was performed using the MatchIt package (method = 'subclass') [36], with an aim to estimate the Average Treatment Effect (ATE) via using estimand = 'ATE'. PS subclassification involved dividing the dataset into several subclasses based on the estimated PS. To determine the optimal number of subclasses, an interactive approach was employed, testing between 5 and 15 subclasses to find the best balance between TMD exposed and unexposed groups, as determined by standardised mean differences (SMD) and the covariate balance summary generated by the MatchIt package. Additionally, the Matchit package created a new weight variable post-PS Subclassification controlled by the s.weights and estimand = 'ATE' functions [36]. The new weights accounted for the population weight from the original survey data, multiplied by the subclassification weights [37], to ensure adequate reweighting. These new weights provided an accurate reflection of each subclass's contribution to the overall population, enabling the generalisation of the findings [11]; (3) Assessing common support of PS across TMD exposed and unexposed groups. Weighted boxplots and density plots [38], were created for PS to visualise the common support of PS before and after subclassification, ensuring adequate PS overlap between TMD exposed and unexposed groups; (4) Assessing group balance: The MatchIt package generated a balance summary which included covariates' mean and computing the SMD estimated by Cohen's d effect size [39], evaluated for both the original data and the PS subclassification averaged across subclasses. Categorical variables were included by creating dummy variables for each category thus, allowing a more meticulous assessment of covariate balance across subclasses. The SMD is the most commonly used balance diagnostic in PS methods where SMD above the threshold of 0.2 suggests covariates' imbalance [40]; (5) Estimating the causal PATE of TMD experience in PS

Subclassification on OHRQoL using OHIP-14 overall score using complex samples *t*-test as well as on the OHIP-14 overall score and domain scores using adjusted complex samples general linear model using the survey package which accounted for strata, clusters and the new computed weight variable that ensured obtaining a single marginal effect that can be generalised to the Australian adult population (PATE) [11, 36]. To estimate the clinical relevance, this study used a statistical approach to estimate the minimally important difference (MID-S) of TMD impact on OHIP-14 scores based on Cohen's f^2 which is an appropriate measure of effect size considering it could estimate the effect size from multivariable hierarchical models [41]. Cohen's f^2 was calculated using the following equation: $f^2 = R^2_{AB} - R^2_A / 1 - R^2_{AB}$ [41] where R^2_{AB} was the OHIP-14 variance explained in multivariable linear regression by TMD experience together with a set of confounders and R^2_A is the OHIP-14 variance explained by confounders in the multivariable general linear regression (excluding the TMD experience). This study used Cohen's f^2 index to determine the magnitude of the effect size where $f^2 \geq 0.02$, $f^2 \geq 0.15$, *and* $f^2 \geq 0.35$ represent small, medium and large effect sizes, respectively [41, 42]. A supplementary sensitivity analysis was performed to assess the robustness of the estimated causal effect of TMD exposure on OHIP-14 overall and domains score if an unmeasured confounder was introduced [43]. Two hypothetical confounders were introduced one with a low and another with a moderate correlation with the exposure and the outcome using R commands. The generated confounders were assessed for correlation with both TMD status and OHIP-14 and then, complex samples GLMs were performed for OHIP-14 total and domain scores in adjusted models that included the chosen confounders and each of the hypothetical confounders.

3 | Results

3.1 | Propensity Score Subclassification and Balance Summary

This study initially included 4078 Australian adults who completed the three NSAOH components and after exclusion of participants who had missing values, we included 4053 Australian adults in our PS subclassification analysis. In the original data, 397 (9.8%) had TMD (exposed group). Propensity scores for TMD status was computed using logistic regression with survey design (strata, clusters and weights) and added interaction terms for having private dental insurance with oral status elements (decayed, missing and filled teeth as well as 'Teeth feel do not fit properly together'). Using an interactive approach in determining the optimal number of subclasses that achieved the optimal covariate balance using a range of subclasses from 5 to 15 subclasses, having five subclasses achieved optimal balance of covariates and PS across TMD status groups. The generated PS subclassification included 397 with TMD and 3656 without TMD where there was no discarded participants as a result of PS subclassification. During the PS subclassification process, the Matchit package computed a new weights variable for PATE estimation post-PS Subclassification. Table 1 showed adequate covariate balance in PS subclassification across TMD status groups demonstrated by SMD values being < 0.2 compared with covariate imbalance in the original data.

3.2 | Assessment of Common Support of the Predicted Probability (Propensity Score) for Having TMD Across the PS Subclassification TMD Groups

PS of TMD status across the PS Subclassification compared with original data were assessed. The weighted boxplot and density plots (Figure 2) revealed that TMD status groups post-PS Subclassification shared common support for PS compared with original data. Further, among the Australian adult population, the predicted probability (PS) in the original data in those with TMD (mean = 0.173, 95% CI: 0.150–0.197) and those without (mean = 0.0903, 95% CI: 0.0865–0.0941) was different ($t_{(326)} = 6.92$, $p < 0.001$, SMD = 0.7565) whereas in PS subclassification data, the PS in those with TMD (mean = 0.108, 95% CI: 0.0917–0.125) and those without TMD (mean = 0.096, 95% CI: 0.092–0.100) was not different ($t_{(326)} = 1.4414$, $p = 0.150$, SMD = 0.135) indicating the PS subclassification achieved adequate balance of TMD status predicted probability in a complex samples *t*-test.

3.3 | Estimating the PATE of Having TMD on OHIP-14 Total Scores

Among the Australian adult population represented by the PS Subclassification data, the total OHIP-14 scores in those with TMD (Mean = 10.544, 95% CI: 8.786–12.302) compared to those without TMD (Mean = 6.750, 95% CI: 6.298–7.201) were significantly higher ($t_{(326)} = 4.2$, $p < 0.001$, Cohen's $d = 0.02$) in a complex samples *t*-test. In a fully-adjusted complex samples general linear model for OHIP-14 total scores, experiencing TMD caused higher OHIP-14 scores ($B = 3.498$, 95% CI: 2.218–4.7776) as showed in Table 2. Additionally, the effect size of PATE of having TMD on OHIP-14 total score was calculated using Cohen's F^2 and found it to be of a small MID-S (Cohen's $F^2 = 0.03$).

3.4 | Estimating PATE of Having TMD on OHIP-14 Domain Scores

Using multivariable multi-variate complex samples general linear models (Table 3), the TMD experience causal PATE on OHIP-14 domains' score was estimated. All OHIP-14 domains were impaired in those who had TMD experience. Physical pain was the highest impacted OHIP-14 domain ($B = 0.805$, 95% CI: 0.433–1.178) as well as the highest observed effect size which was found to be small (Cohen's $F^2 = 0.03$). The second most impacted OHIP-14 domain was psychological discomfort ($B = 0.760$, 95% CI: 0.410–1.110) with a small effect size (Cohen's $F^2 = 0.02$). The third most impacted OHIP-14 domain was psychological disability ($B = 0.615$, 95% CI: 0.3370–0.894) with a small effect size (Cohen's $F^2 = 0.02$). The fourth, fifth, sixth and seventh impacted OHIP-14 domains were Functional Limitation ($B = 0.375$. 95% CI: 0.154–0.595, Cohen's $F^2 = 0.01$), Physical Disability ($B = 0.367$, 95% CI: 0.1711–0.567, Cohen's $F^2 = 0.01$), Social Disability ($B = 0.341$, 95% CI: 0.156–0.526, Cohen's $F^2 = 0.01$) and handicap ($B = 0.233$, 95% CI: 0.069–0.397, Cohen's $F^2 = 0.01$) respectively with very small observed effect sizes.

TABLE 1 | Comparison of balance on key confounders across TMD status groups in the original data and the propensity score subclassification data.

	TMD status (Original data)			TMD status (PS subclassification)		
	Yes ($n = 397$)	No ($n = 3656$)		Yes ($n = 397$)	No ($n = 3656$)	
	Mean	Mean	SMD	Mean	Mean	SMD
Age (years)	38.6085	42.9570	−0.2678	41.2982	42.5928	−0.0797
No. of teeth decayed	1.0182	0.5229	0.2541	0.5547	0.5398	0.0076
No. of teeth filled due to pathology	7.4362	7.9232	−0.0757	7.4075	7.8809	−0.0736
No. of teeth missing for any reason	5.3587	5.7308	−0.0778	5.4785	5.6935	−0.0449
Perceived Stress Scale (PSS-14)						
Perceived Distress Subscale	14.1748	11.7469	0.4483	12.2724	11.9392	0.0615
Perceived Control Subscale	16.2580	17.7618	−0.3171	17.5992	17.6400	−0.0086
The doctor said you have diabetes						
Yes	0.0436	0.0429	0.0035	0.0411	0.0435	−0.0122
No	0.9564	0.9571	−0.0035	0.9589	0.9565	0.0122
Have private dental insurance						
Yes	0.3798	0.4778	−0.1990	0.4389	0.4689	−0.0609
No	0.6202	0.5222	0.1990	0.5611	0.5311	0.0609
Sex						
Female	0.6452	0.4817	0.3342	0.5013	0.4968	0.0092
Male	0.3548	0.5183	−0.3342	0.4987	0.5032	−0.0092
Teeth feel do not fit properly together						
Yes	0.3965	0.1802	0.4917	0.2275	0.1982	0.0668
No	0.6035	0.8198	−0.4917	0.7725	0.8018	−0.0668
Income						
Less than $60 k	0.5955	0.5127	0.1671	0.5102	0.5189	−0.0177
$60 k or more	0.4045	0.4873	−0.1671	0.4898	0.4811	0.0177
Periodontitis: moderate or severe						
Yes	0.1910	0.2183	−0.0676	0.2242	0.2147	0.0235
No	0.8090	0.7817	0.0676	0.7758	0.7853	−0.0235
Current smoker						
Yes	0.2432	0.1413	0.2606	0.1375	0.1478	−0.0263
No	0.7568	0.8587	−0.2606	0.8625	0.8522	0.0263
Employed						
Yes	0.6052	0.6630	−0.1203	0.6526	0.6596	−0.0146
No	0.3948	0.3370	0.1203	0.3474	0.3404	0.0146
Occupation						
Professional/managerial	0.6003	0.6016	−0.0027	0.5966	0.6014	−0.0098

(Continues)

TABLE 1 | (Continued)

	TMD status (Original data)			TMD status (PS subclassification)		
	Yes ($n=397$)	No ($n=3656$)		Yes ($n=397$)	No ($n=3656$)	
	Mean	Mean	SMD	Mean	Mean	SMD
Intermediate	0.2803	0.2522	0.0637	0.3049	0.2550	0.1130
Manual workers	0.1194	0.1462	−0.0791	0.0985	0.1436	−0.1330
Educational attainment						
University/College	0.4631	0.5352	−0.1445	0.4885	0.5293	−0.0819
Certificate/TAFE	0.4632	0.4101	0.1073	0.4587	0.4145	0.0893
Other	0.0736	0.0547	0.0773	0.0528	0.0562	−0.0138
Country of birth						
Australia	0.8071	0.7885	0.0463	0.8069	0.7906	0.0406
Overseas	0.1929	0.2115	−0.0463	0.1931	0.2094	−0.0406
Indigenous						
No	0.9855	0.9897	−0.0385	0.9782	0.9894	−0.1010
Yes	0.0145	0.0103	0.0385	0.0218	0.0106	0.1010

Note: SMD, standardised mean difference as estimated by the weighted balance summary flowing PS Subclassification using the Matchit package in R.

3.5 | Supplementary Sensitivity Analysis

The created hypothetical confounders were assessed for correlation with TMD status and the OHIP-14 total scores. The low correlation hypothetical confounder exhibited correlation ($r=0.1$, $p<0.001$) with TMD status and ($r=0.2$, $p<0.001$) with OHIP-14 total scores. In an adjusted complex samples GLM for OHIP-14 total scores where the model included the identified confounders and the low correlation hypothetical confounder (Table S1), having TMD has resulted in higher OHIP-14 total scores with a small MID-S ($B=3.524$, 95% CI: 2.250–4.798, Cohen's $f^2=0.03$). Additionally, having TMD resulted in higher scores for OHIP-14 domains with highest impacts observed in physical pain ($B=0.821$, 95% CI: 0.453–1.189, Cohen's $f^2=0.03$), psychological discomfort ($B=0.770$, 95% CI: 0.420–1.119, Cohen's $f^2=0.02$) and psychological disability ($B=0.619$, 95% CI: 0.341–0.897, Cohen's $f^2=0.02$) whereas other domains were also impacted ($p<0.005$). The moderate correlation hypothetical confounder exhibited correlation ($r=0.3$, $p<0.001$) with TMD status and ($r=0.4$, $p<0.001$) with OHIP-14 total scores. In an adjusted complex samples GLM for OHIP-14 total scores where the model included the identified confounders and the moderate correlation hypothetical confounder (Table S2), having TMD resulted in higher OHIP-14 total scores with a small MID-S ($B=3.469$, 95% CI: 2.190–4.748, Cohen's $f^2=0.03$). Additionally, having TMD resulted in higher scores for OHIP-14 domains with highest impacts observed in physical pain ($B=0.804$, 95% CI: 0.432–1.177, Cohen's $f^2=0.02$), psychological discomfort ($B=0.760$, 95% CI: 0.410–1.109, Cohen's $f^2=0.02$) and psychological disability ($B=0.611$, 95% CI: 0.332–0.890, Cohen's $f^2=0.02$) whereas other domains were also impacted ($p<0.005$).

4 | Discussion

In this study, a PS Subclassification tool was used to mitigate confounding bias by balancing confounders among those experiencing TMD and those who do not, as well as balancing the predicted probability of experiencing TMD in the Australian adult population. Consequently, insights into estimating the PATE of TMD experience on OHRQoL among the Australian adult population were obtained. Further, this study estimated the effect size as a statistical approach to the Minimally Important Difference (MID-S). This study found that the PS Subclassification tool enabled an adequate balance of TMD experience predicted probability and the distribution of knowledge-based confounders driven from the developed TMD and OHRQoL causal model across TMD status groups in the PS Subclassification dataset created following the application of PS Subclassification process. Therefore, this study adequately mitigated confounding bias on measured covariates. Further, this study found that, among the Australian adult population, TMD experience has significantly impaired the overall OHRQoL measured by the total OHIP-14 scores and has a small MID-S. Additionally, on estimating the causal PATE of TMD experience on OHRQoL aspects as measured by the OHIP-14 domains, this study found that all of the OHIP-14 domains were significantly impaired in those who experienced TMD. Physical pain was the most impaired OHRQoL domain in those with TMD experience with small MID-S followed by psychological discomfort and psychological disability respectively with a small MID-S whereas physical disability, functional limitations, social disability and handicap were also impaired with very small MID-S. The supplementary sensitivity analysis showed that the observed PATE of having TMD

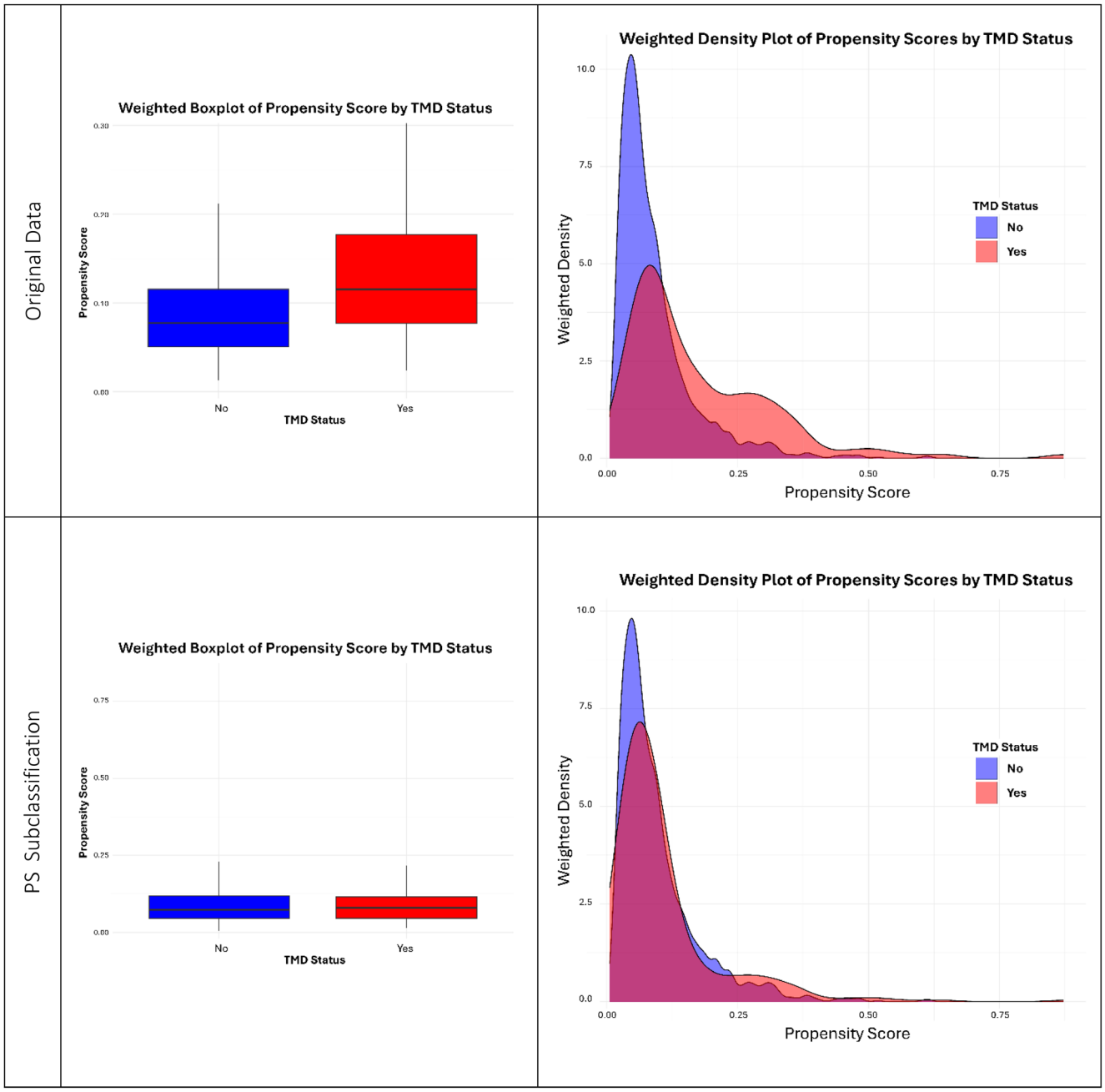

FIGURE 2 | Sharing common support of propensity score by TMD status in the original data and the PS subclassification using a weighted boxplot (on the left side) and a weighted density plots (on the right side).

on OHIP-14 overall and domain scores did not substantially change when low or moderate correlation hypothetical confounders were introduced to the adjusted model.

To the authors' knowledge, this study is one of the pioneer studies to estimate the population-average causal effect of having TMD experience on OHRQoL using PS subclassification. The PS Subclassification approach used in this study achieved an adequate balance of TMD experience predicted probability and the measured confounders considering the reported PS Subclassification balance diagnostics using the SMD [40]. This approach might provide an opportunity to mimic randomised controlled trials in mitigating the effect of observed confounder bias on TMD causal effect estimation using cross-sectional data. Using a PS framework with cross-sectional data in estimating causal effect was recommended to overcome barriers such as ethical considerations that prevent using experimental studies and the costs/resources needed for cohort studies [9, 13]. When the PS Subclassification tool was used in this study, identification of confounders was needed in generating PS as well as balancing confounders across the TMD status groups in the PS Subclassification dataset. Confounders were selected using a knowledge-based causal model [1, 15–21], an approach recommended by Petersen, van der Laan [12]. Accordingly, it was believed that this study mitigated the confounding bias on the measured confounders which, arguably, made PATE estimates reliable.

TABLE 2 | Complex samples GLM estimating TMD experience population average treatment effect on OHIP-14 overall scores.

Parameter	Estimate	95% CI	
		Lower	Upper
TMD=Yes	3.498	2.218	4.7776
Diabetic = No	−1.6102	−3.6129	0.3926
Private dental insurance = No	0.8986	0.2402	1.5571
Sex = Male	−0.2923	−0.9977	0.4131
Total household income			
≥ < $60 000	−1.2269	−1.9822	−0.4715
< $60 000	0.000[a]		
Employed = No	−0.5355	−1.3572	0.2862
Currently smoker = No	0.1452	−0.9302	1.2205
Occupation			
Manual workers	−2.1450	−3.1644	−1.1256
Intermediate	−0.9075	−1.8844	0.0694
Professionals/ admin/managerial	0.00[a]		
Teeth felt do not fit properly together = No	−3.9587	−4.7851	−3.1322
Periodontitis: moderate or severe = No	−1.5427	−2.3395	−0.7458
Education			
Certificate/TAFE	0.6984	−0.0757	1.4726
Other	0.4440	−0.9956	1.8836
University/College	0.00[a]		
Birthplace = Overseas	1.6295	0.7838	2.4751
Indigenous = Yes	1.7044	−2.0668	5.4755
No. of decayed teeth	1.0436	0.6157	1.4714
No. of missing teeth	0.2985	0.2209	0.3760
No. of filled teeth	0.1977	0.1343	0.2611
Age (Years)	−0.0785	−0.1087	−0.0483
Perceived Stress Scale (PSS-14)			
PSS Distress	0.3468	0.2662	0.4274
PSS Control	−0.1465	−0.2259	−0.0672

Note: Effect Size for TMD experience: Cohen's $f^2 = 0.03$.
[a]Set to zero because this parameter is redundant.

There are several PS methods described in the medical literature [44]. The PS Subclassification method used in this study was reported to be suitable to estimate the PATE while using all data similar to inverse probability of treatment weighting (IPTW) PS method [44], and is argued to have lower bias in effect estimation and be robust to slight mis-specification in the PS model compared with IPTW [11, 44]. The Matchit package in R [33, 35], offered simplicity to perform PS Subclassification and obtain balance summary diagnostics as an overall, weighted average balance across the subclasses. When performing the PS Subclassification, an interactive approach was used to determine the optimal number of subclasses, rather than the commonly used five subclasses described in the literature [44], based on which number of subclasses achieved the best balance on PS and confounders. To generalise the Average Treatment Effect estimates to the Australian adult population, the survey design elements (weights, clusters and strata) were included in computing PS [11]. Further, the s.weights and estimand = 'ATE' functions in the Matchit package controlled the computation of a new weight variable during the PS Subclassification process [36], which maintained the representativeness of the PS Subclassification dataset to the original population sampled by NSAOH. Considering the new weight variable was included in the effect estimation post-PS Subclassification along with survey design elements (strata and clusters), a single marginal effect of TMD experience on OHIP14 total scores as well as each of the OHIP-14 domain scores were estimated which could be generalizable to the Australian adult population—PATE [11, 36]. Also, the study used estimates of effect size including Cohen's *d* and Cohen's f^2 which are not confounded by the sample size.

The principal finding of this study was that TMD experiences caused a significant impairment in the overall OHRQoL among the Australian adult population which was consistent with a recent systematic review and meta-analysis of small-sized studies [2]. This might be attributed to the presence of oro-facial pain [4, 5], trismus and chewing difficulties [8], which might lead to diet modification as well as the grating/clicking of the TMD [7], and thus, impair the physical and psychological aspects of OHRQoL. In measuring the OHRQoL, this study used OHIP-14 which is a popular generic measure of OHRQoL in TMD research [2]. While this study observed a small MID-S of TMD experience on OHRQoL, other disease-specific OHRQoL measures for TMD such as the OHIP-TMD [45], might show a different MID-S than that was observed in this study.

In this study, all OHRQoL domains measured by OHIP-14 were impacted in those with TMD experience in the Australian adult population suggesting the broad extent of how TMD experience impaired OHRQoL aspects. Further, this study found that physical pain, psychological discomfort and psychological disability were the highest impaired OHRQoL domains as well as having the highest MID-S exhibiting a small MID-S whereas the rest of OHIP-14 domains showed very small MID-S. This finding might offer a preliminary insight into the observed use of non-steroidal anti-inflammatory drugs (NSAID) and selective serotonin reuptake inhibitors (SSRIs) in some studies investigating a pharmacological therapy for TMD patients however, further research is needed given the limitations of the available studies [46]. Given the observed negative impact of TMD on OHRQoL, health policymakers and clinicians might consider screening for TMD in routine dental care since findings from comparable countries reported

TABLE 3 | Multivariate complex samples general linear model to estimate the population average treatment effect of TMD experience on OHIP-14 domains.

OHIP-14 domains		Parameter estimates[a]				Effect size[b]
				95% Confidence interval		
		Parameter	Estimate	Lower	Upper	Cohen's f^2
1	Functional limitation	TMD=Yes	0.375	0.154	0.595	0.01
2	Physical pain	TMD=Yes	0.805	0.433	1.178	0.03
3	Psychological discomfort	TMD=Yes	0.760	0.410	1.110	0.02
4	Physical disability	TMD=Yes	0.367	0.1711	0.567	0.01
5	Psychological disability	TMD=Yes	0.615	0.3370	0.894	0.02
6	Social disability	TMD=Yes	0.341	0.156	0.526	0.01
7	Handicap	TMD=Yes	0.233	0.069	0.397	0.01

[a]Model: OHIP-14: domains + TMD + Diabetic + Private dental insurance + Sex + Total household income + Employed + Smoker + Occupation + Teeth felt do not fit properly together + Periodontitis + Education + Birthplace + Indigenous + AGE + No. of filled teeth + No. of decayed teeth + No. of missing teeth + Perceived Stress Scales: PSS Distress + PSS Control.
[b]Effect Size for TMD experience where $R^2_{AB} = (R^2$ for the model with TMD) and $R^2_A = (R^2$. for the model without TMD) Cohen's $f^2 = R^2_{AB} - R^2_A / 1 - R^2_{AB}$ where $f^2 \geq 0.02$, $f^2 \geq 0.15$, and $f^2 \geq 0.35$ represent small, medium and large effect sizes, respectively.

that TMD is undetected in general dental practice [47], where it is estimated that a minority of patients (3%–7%) seek a professional's advice for their TMD [48], compared with the reported TMD prevalence in the population. Early detection of TMD might provide an opportunity for early management which might improve patients' oral health status and consequently, their OHRQoL.

This study has some limitations. The study relied on a questionnaire-based assessment of TMD experience considering that the NSAOH standardised oral epidemiological examination protocol did not include clinical evaluation of TMD experience [30], and the questionnaire version used varied from the TMD questions adopted by the International Network of Orofacial Pain and Related Disorders Methodology [49]. However, the used questionnaire has demonstrated sensitivity and specificity with the clinical diagnosis of TMD [19], and in a previous study [1], a confirmatory factor analysis was conducted and showed the validity of its two-domain structure as well as it was used in several studies [1, 19, 21]. Another limitation might be the temporality of the estimated causal effect of TMD on OHRQoL considering the cross-sectional nature of the NSAOH data used. However, the reference period for the OHIP-14 was 'in the past 12 months.' Also, the exposure is believed to have preceded the outcome considering the chronic nature of TMD [50]. Additionally, cross-sectional studies have been increasingly used to estimate causal effect using the PS methods [9, 13]. Another limitation might be attributed to bias due to unmeasured confounders which is an inherent limitation of all observational studies. However, it is assumed that random NSAOH sampling and PS subclassification that achieved confounders' balance might reduce their impact on causal estimates reported in this study. Moreover, this study included a range of confounders which were unlikely to be included in a single study and supported by the high quality of the NSAOH data. To address the potential bias due to unmeasured confounding, the study included a supplementary sensitivity analysis where hypothetical low and moderate correlation confounders were introduced in the adjusted regression model [43], with qualitatively close findings of TMD impairment of OHRQoL suggesting the reliability of the reported causal estimates. Another limitation of this study was attributed to the age of this dataset warranting caution in generalising the findings of this study to the current Australian adult population given the change of their characteristics and socio-economic status over time. However, the PS Subclassification balance diagnostics on TMD predicted probability and measured confounders including the population characteristics and socio-economic status added to the internal validity of the reported causal estimates along with the sensitivity analysis which showed qualitatively close findings mitigating the potential impact of unmeasured confounding thus, making this study useful in understanding how TMD experience might impair OHRQoL in the Australian adult population sampled in this dataset. PS-based studies tend to utilise older datasets [51], to highlight the applicability of newly developed statistical techniques and might support the future application in estimating TMD population-average causal effect on OHRQoL given the current limitation in the available dental literature. Besides, this study can form baseline findings which might be used in future comparative studies with subsequent NSAOH waves given it was not done before among this population.

On the other hand, this study has several strengths. This study used the PS Subclassification tool to achieve adequate balance in the predicted probability of TMD experience and measured confounders across TMD exposed and unexposed groups in the PS Subclassification dataset in an attempt to mitigate confounding bias and have a reliable estimation of the population average causal effect on OHRQoL—an approach that is limited in the current TMD and OHRQoL literature. Estimating the causal population average effect of TMD on OHRQoL might be limited as conducting RCT is unethical and the cohort studies have difficulties in recruiting sufficient sample size or retaining participants over a long duration which requires substantial resources and is associated with significant costs. Moreover, this study selected the confounders used in PS computation

and subclassification balance summary from the developed knowledge-based TMD and OHRQoL model rather than using confounders of convenience with the added advantage of PS subclassification being robust to slight mis-specification in PS model [44]. The PS subclassification approach employed in this study used all participants included in the original dataset along with inclusion of survey design elements in PS computation and post-PS Subclassification weighting which enabled generalizability of the reported estimates and therefore reporting the PATE of TMD experience on OHRQoL. This study estimated that TMD experience impaired the overall OHRQoL as well as impaired all OHRQoL domains measured by OHIP-14 suggesting the broad extent of how TMD impaired OHRQoL which might need the clinicians and health policymakers' attention to provide screening and thus, early detection and management to improve OHRQoL of the adult population. This study revealed that the physical pain and psychological aspects of OHRQoL were the highest impaired OHRQoL domains which might offer a preliminary insight into the observed use of pharmacological approaches by researchers in TMD management where potential benefit on improving OHRQoL of TMD patients was reported but, further research is needed.

5 | Conclusion

This study found that having TMD experience impaired OHRQoL with a small MID-S suggesting its negative impact on Australian adults' wellbeing sampled by the NSAOH dataset used in this study. Further, having TMD experience impaired all OHRQoL domains measured by the OHIP-14 suggesting the broad extent of how having TMD experience negatively impacted on OHRQoL. This might suggest the need for clinicians and health policymakers to adopt screening for TMD to enable early detection and management which might improve OHRQoL—an end outcome of healthcare services as there is some evidence that TMD experience might be under-detected in comparable healthcare systems. Also, physical pain and psychological OHRQoL domains (small MID-S) were the most highly impaired by TMD experience. This finding might offer a preliminary insight into the reported efficacy of NSAID and selective serotonin reuptake inhibitors (SSRIs) in TMD management warranting the need for further research on patient-centred pharmacological management approaches for TMD.

Author Contributions

Kamal Hanna: conceptualization, data curation, methodology, formal analysis, visualisation, writing – original draft, writing-review and editing; **Ninuk Hariyani:** provided intellectual contribution and revised the manuscript; **Gloria Mejia:** provided intellectual contribution, methodology, writing, review and editing; **Lisa Jamieson:** provided intellectual contribution and revised the manuscript; methodology, writing, review and editing; **David S. Brennan:** supervision, conceptualization, methodology, writing, review and editing.

Acknowledgements

The authors acknowledge the effort made by the Australian Research Centre for Population Oral Health, Adelaide Dental School, the University of Adelaide for NSAOH data collection and management. The principal author dedicates this piece of research to Dr. Rahul Nair who inspired him to carry out this research. Open access publishing facilitated by The University of Adelaide, as part of the Wiley - The University of Adelaide agreement via the Council of Australian University Librarians.

Ethics Statement

Ethical approval for the NSAOH was obtained from the University of Adelaide Human Research Ethics Committee with approval number: H-001-2004.

Consent

The consent form obtained from the NSAOH's participants indicates the use of collected data in research publication however, no participant's identifying information will be published. There is no embargo placed on publishing research from this dataset.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data analysis output file is available upon request from the Australian Research Centre for Population Oral Health, Adelaide Dental School, the University of Adelaide.

References

1. K. Hanna, R. Nair, N. Amarasena, J. Armfield, and D. Brennan, "Temporomandibular Dysfunction Is Associated With Oral Health-Related Quality of Life: An Australian National Study," *BMC Oral Health* 21, no. 432 (2021): e1–e13.

2. Z. Qamar, A. M. S. Alghamdi, N. K. B. Haydarah, et al., "Impact of Temporomandibular Disorders on Oral Health-Related Quality of Life: A Systematic Review and Meta-Analysis," *Journal of Oral Rehabilitation* 50, no. 8 (2023): 706–714.

3. M. T. John, "Foundations of Oral Health-Related Quality of Life," *Journal of Oral Rehabilitation* 48, no. 3 (2020): 355–359, https://doi.org/10.1111/joor.13040.

4. S. Shueb, D. Nixdorf, M. John, B. Alonso, and J. Durham, "What Is the Impact of Acute and Chronic Orofacial Pain on Quality of Life?," *Journal of Dentistry* 43, no. 10 (2015): 1203–1210.

5. I. Oghli, T. List, N. Su, and B. Häggman-Henrikson, "The Impact of Oro-Facial Pain Conditions on Oral Health-Related Quality of Life: A Systematic Review," *Journal of Oral Rehabilitation* 47, no. 8 (2020): 1052–1064.

6. K. Hanna, P. Sambrook, J. Armfield, and D. Brennan, "Exploring and Modelling Impacts of Third Molar Experience on Quality of Life: A Real-Time Qualitative Study Using Twitter," *International Dental Journal* 67, no. 5 (2017): 272–280.

7. M. Fletcher, J. Piecuch, and S. Lieblich, "Temporomandibular Joint Disease: Anatomy and Pathophysiology of the Temporomandibular Joint," in *Peterson's Principles of Oral and Maxillofacial Surgery* (People's Medical Publishing House USA Ltd (PMPH), 2004).

8. D. S. Brennan, A. J. Spencer, and K. F. Roberts-Thomson, "Tooth Loss, Chewing Ability and Quality of Life," *Quality of Life Research* 17, no. 2 (2008): 227–235.

9. C. Mason, C. Sabariego, Đ. M. Thắng, and J. Weber, "Can Propensity Score Matching Be Applied to Cross-Sectional Data to Evaluate Community-Based Rehabilitation? Results of a Survey Implementing the WHO's Community-Based Rehabilitation Indicators in Vietnam," *BMJ Open* 9, no. 1 (2019): e022544.

10. P. C. Austin, "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies," *Multivariate Behavioral Research* 46, no. 3 (2011): 399–424.

11. E. H. Dugoff, M. Schuler, and E. A. Stuart, "Generalizing Observational Study Results: Applying Propensity Score Methods to Complex Surveys," *Health Services Research* 49, no. 1 (2014): 284–303.

12. M. L. Petersen and M. J. van der Laan, "Causal Models and Learning From Data: Integrating Causal Modeling and Statistical Estimation," *Epidemiology (Cambridge, Mass.)* 25, no. 3 (2014): 418–426.

13. K. Sekine, R. R. Carandang, K. I. C. Ong, A. Tamang, and M. Jimba, "Identifying the Causal Effect of Child Marriage on Unmet Needs for Modern Contraception and Unintended Pregnancy in Nepal: A Cross-Sectional Study Using Propensity Score Matching," *BMJ Open* 11, no. 4 (2021): e043532.

14. J. Textor, B. van der Zander, M. S. Gilthorpe, M. Liskiewicz, and G. T. Ellison, "Robust Causal Inference Using Directed Acyclic Graphs: The R Package 'dagitty'," *International Journal of Epidemiology* 45, no. 6 (2016): 1887–1894.

15. G. Slade, "Assessing Change in Quality of Life Using the Oral Health Impact Profile," *Community Dentistry and Oral Epidemiology* 26, no. 1 (1998): 52–61.

16. L. Al-Harthi, M. Cullinan, J. Leichter, and W. Thomson, "The Impact of Periodontitis on Oral Health-Related Quality of Life: A Review of the Evidence From Observational Studies," *Australian Dental Journal* 58, no. 3 (2013): 274–277.

17. K. S. Lee, N. Jha, and Y. J. Kim, "Risk Factor Assessments of Temporomandibular Disorders via Machine Learning," *Scientific Reports* 11, no. 1 (2021): 19802.

18. C. L. Cruz, K. C. Lee, J. H. Park, and A. I. Zavras, "Malocclusion Characteristics as Risk Factors for Temporomandibular Disorders: Lessons Learned From a Meta-Analysis," *Journal of Oral Diseases* 2015 (2015): 302646.

19. A. Sanders and G. Slade, "Gender Modifies the Effect of Perceived Stress on Orofacial Pain Symptoms: National Survey of Adult Oral Health," *Journal of Orofacial Pain* 25, no. 4 (2011): 317–326.

20. J. Goergen, R. S. A. Costa, F. S. Rios, et al., "Oral Conditions Associated With Oral Health Related Quality of Life: A Population-Based Cross-Sectional Study in Brazil," *Journal of Dentistry* 129 (2023): 104390.

21. K. Hanna, R. Nair, J. Armfield, and B. Brennan, "Temporomandibular Dysfunction Among Working Australian Adults and Association With Workplace Effort-Reward Imbalance," *Community Dental Health* 37 (2020): 253–259.

22. Australian Institute of Health and Welfare, "National dental data sources," 2017, accessed July 1, 2017, http://www.aihw.gov.au/dental/data-sources/.

23. Electronic white pages. Australia on Disc 2004.

24. G. Slade, A. Spencer, and K. Roberts-Thomson, *Australia's Dental Generations: The National Survey of Adult Oral Health 2004–06* (Dental statistics and research series, 2007).

25. A. Chisnoiu, A. Picos, S. Popa, et al., "Factors Involved in the Etiology of Temporomandibular Disorders - a Literature Review," *Clujul Medical* 88, no. 4 (2015): 473–478.

26. The University of Adelaide, "The National Survey of Adult Oral Health 2004/05: Protocol for Oral Epidemiological Examinations. Adelaide: Australian Research Centre for Population Oral Health," 2005.

27. G. Slade, S. Foy, D. Shugars, C. Phillips, and R. P. White, Jr., "The Impact of Third Molar Symptoms, Pain, and Swelling on Oral Health-Related Quality of Life," *Journal of Oral and Maxillofacial Surgery* 62, no. 9 (2004): 1118–1124.

28. D. Locker and G. Slade, "Association of Symptoms and Signs of TM Disorders in an Adult Population," *Community Dentistry and Oral Epidemiology* 17, no. 3 (1989): 150–153.

29. A. Sanders, G. Slade, K. Carter, and J. Stewart, "Trends in Prevalence of Complete Tooth Loss Among Australians, 1979—2002," *Australian and New Zealand Journal of Public Health* 28, no. 6 (2004): 549–554.

30. L. Do, A. Ellershaw, J. Harford, et al., "Unit DSaR, Ed. Dental Statistics and Research Series," in *The National Survey of Adult Oral Health 2004-2006*, vol. 43 (Australian Institute of Health and Welfare, 2008).

31. R. Page and P. Eke, "Case Definitions for Use in Population-Based Surveillance of Periodontitis," *Journal of Periodontology* 78, no. 7 Suppl (2007): 1387–1399.

32. S. Cohen, T. Kamarck, and R. Mermelstein, "A Global Measure of Perceived Stress," *Journal of Health and Social Behavior* 24, no. 4 (1983): 385–396.

33. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024).

34. *RStudio: Integrated Development Environment for R* (Posit, PBC, 2024).

35. Package 'MatchIt': Nonparametric Preprocessing for Parametric Causal Inference [computer program]. Version 4.5.5 2023.

36. r-project.org, "method_subclass {MatchIt}. R Documentation," accessed Oct 10, 2024, https://search.r-project.org/CRAN/refmans/MatchIt/html/method_subclass.html.

37. R. J. Desai, K. J. Rothman, B. T. Bateman, S. Hernandez-Diaz, and K. F. Huybrechts, "A Propensity-Score-Based Fine Stratification Approach for Confounding Adjustment When Exposure Is Infrequent," *Epidemiology (Cambridge, Mass.)* 28, no. 2 (2017): 249–257.

38. M. S. Ali, R. H. H. Groenwold, and O. H. Klungel, "Best (But Oft-Forgotten) Practices: Propensity Score Methods in Clinical Nutrition research1, 2, 3," *American Journal of Clinical Nutrition* 104, no. 2 (2016): 247–258.

39. *Practical Meta-Analysis Effect Size Calculator* (George Mason University, 2001).

40. N. Larrain and O. Groene, "Improving the Evaluation of an Integrated Healthcare System Using Entropy Balancing: Population Health Improvements in Gesundes Kinzigtal," *SSM—Population Health* 22 (2023): 101371.

41. A. Selya, J. Rose, L. Dierker, D. Hedeker, and R. Mermelstein, "A Practical Guide to Calculating Cohen's f^2, a Measure of Local Effect Size, From PROC MIXED," *Frontiers in Psychology* 3, no. 111 (2012): e1–e6.

42. J. Cohen, *Statistical Power Analysis for the Social Science*, 2nd ed. (Lawrence Erlbaum Associates, 1988).

43. S. Greenland, "Basic Methods for Sensitivity Analysis of Biases," *International Journal of Epidemiology* 25, no. 6 (1996): 1107–1116.

44. S. Deb, P. C. Austin, J. V. Tu, et al., "A Review of Propensity-Score Methods and Their Use in Cardiovascular Research," *Canadian Journal of Cardiology* 32, no. 2 (2016): 259–265.

45. P. L. Yule, J. Durham, H. Playford, et al., "OHIP-TMDs: A Patient-Reported Outcome Measure for Temporomandibular Disorders," *Community Dentistry and Oral Epidemiology* 43, no. 5 (2015): 461–470.

46. D. Fischoff and S. Spivakovsky, "Are Pharmacological Treatments for Oro-Facial Pain Effective?," *Evidence-Based Dentistry* 19, no. 1 (2018): 28–29.

47. A. Lövgren, C. M. Visscher, B. Häggman-Henrikson, F. Lobbezoo, S. Marklund, and A. Wänman, "Validity of Three Screening Questions (3Q/TMD) in Relation to the DC/TMD," *Journal of Oral Rehabilitation* 43, no. 10 (2016): 729–736.

48. M. Murphy, R. MacBarb, M. Wong, and K. Athanasiou, "Temporomandibular Disorders: A Review of Etiology, Clinical Management, and

Tissue Engineering Strategies," *International Journal of Oral & Maxillofacial Implants* 28, no. 6 (2013): e393-414.

49. International Network of Orofacial Pain and Related Disorders Methodology, "International Network of Orofacial Pain and Related Disorders Methodology," 2021, accessed May 15, 2021, https://ubwp.buffalo.edu/rdc-tmdinternational/other-instruments-2/.

50. K. C. Srivastava, D. Shrivastava, Z. A. Khan, et al., "Evaluation of Temporomandibular Disorders Among Dental Students of Saudi Arabia Using Diagnostic Criteria for Temporomandibular Disorders (DC/TMD): A Cross-Sectional Study," *BMC Oral Health* 21, no. 1 (2021): 211.

51. I. Novitzky-Basso, C. Patriquin, S. M. Linn, et al., "Propensity Score Matching Analysis Comparing the Efficacy and Steroid Tapering Benefit of Extracorporeal Photopheresis to Best Available Therapy in Third-Line or Beyond Treatment for Chronic GvHD," *Transplantation and Cellular Therapy* 29, no. 12 (2023): e771–e773.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.