Heliyon 10 (2024) e25035

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Sociodemographic, biological, and timing characteristics of dental caries and fluorosis using mixed-type cluster analysis on 12-year-olds in Ho Chi Minh city, 1989–2019

Hung Trong Hoang [a], Nam Cong-Nhat Huynh [a,*], Trang Thi-Ngoc Tran [a], Minh Duc Nguyen [b], Eugenio D. Beltrán-Aguilar [c]

[a] Faculty of Odonto-Stomatology, University of Medicine and Pharmacy at Ho Chi Minh City, Viet Nam
[b] Odonto-Maxillo-Facial Hospital in Ho Chi Minh City, Viet Nam
[c] WHO Collaborating Center, Department of Epidemiology and Health Promotion, New York University, USA

A R T I C L E I N F O

Keywords:
Dental caries
Dental fluorosis
Water fluoridation
Mixed-type cluster analysis

A B S T R A C T

Objectives: In 1990, Ho Chi Minh City started Community water fluoridation (CWF) at 0.7 ppm F, and in 2000, it was adjusted to 0.5 ppm F. Here, we analyzed dental caries and fluorosis data in Ho Chi Minh City to explore commonalities associated with CWF among 12-year-old children.
Methods: Dental caries and fluorosis data were collected in 1989, 2003, 2012, and 2019 (N = 4773). Trained dentists scored dental caries using the WHO detection criteria and fluorosis using Dean's Fluorosis Index. We used these data and the k-prototypes method by the R package to identify clusters of participants with shared clinical and water fluoride levels.
Results: We used datasets 1 (4773 participants) and 2 (4194 participants, missing fluorosis data in 1989). K-prototypes analysis identified three clusters in each dataset. Cluster 1, with 80 % of the sample at 0.5 ppm F area characterized by low caries and fluorosis scores. Cluster 2 with 60 % of the sample non-fluoridated area had high caries and low fluorosis scores. Cluster 3, with 75 % of the sample in 0.7 ppm area, had low caries but borderline high fluorosis scores.
Conclusion: Identifying three clusters based on clinical and environmental scores supports the decision to fluoridate the water to prevent caries (0–0.7 ppm) and the shift from 0.7 to 0.5 ppm to keep the caries preventive effect while reducing the risk of fluorosis.
Clinical significance: Our results support the effectiveness of CWF in preventing dental caries and the appropriateness of changing the F concentration to reduce the risk of fluorosis while maintaining its effectiveness.

1. Introduction

Ho Chi Minh City (HCMC) is on the northern edge of the Mekong Delta in Southern Vietnam. The Saigon River crosses the city and is the primary source of drinking water. The city has a tropical climate, hot and humid, year-round. HCMC spreads over 2090 km^2 and is divided into 24 districts. With over nine million people, HCMC is Vietnam's largest city and Southeast Asia's major commercial,

* Corresponding author. Faculty of Odonto-Stomatology, University of Medicine and Pharmacy at Ho Chi Minh City, 652 Nguyen Trai, W11, D5, 749000, Ho Chi Minh City, Viet Nam.
E-mail address: namhuynh@ump.edu.vn (N.C.-N. Huynh).

https://doi.org/10.1016/j.heliyon.2024.e25035
Received 6 July 2023; Received in revised form 18 January 2024; Accepted 18 January 2024
Available online 20 January 2024

tourist, cultural, and scientific hub [1,2].

The 1999 National Oral Health Survey [2] reported that 84% of 12-year-old children in HCMC had dental caries or its sequelae. The mean DMFT was 3.4, despite school-based interventions (oral hygiene instruction with fluoride toothpaste). The need for more effective programs for caries prevention led the City to consider two complementary approaches: 1) to expand the school-based oral care programs which started in the early 1980s to all kindergartens and primary schools in HCMC, and 2) to implement a community water fluoridation program as recommended by WHO experts and local oral health authorities. Community Water Fluoridation in HCMC started in early 1990 [3–7].

HCMC's municipal water supply does not cover all districts; thus, the city is divided into two distinct fluoridated and non-fluoridated areas (Fig. 1). According to the HCMC Centers for Disease Control Agency report, between 1990 and 2000, the average fluoride concentration in tap water in the fluoridated districts was 0.7 ppm F, and in the non-fluoridated areas (well water) was below 0.1 ppm F [8,9]. In June 2000, the People's Committee of HCMC adjusted the fluoride concentration in tap water to 0.5 ppm F [10].

Cluster analysis identifies groups of individuals sharing common values from a set of parameters, thus providing an ecological view of the data [11]. Cluster analysis identifies meaningful conglomerates of people with similar epidemiological profiles [12]. Often, cluster analysis is used with quantitative data. In this study, we applied the k-prototypes method in the R statistical package to analyze mixed-type data, including numeric variables (DMFT index) and categorical variables (fluoride concentration, Dean's index, sex, year), in their association with diagnosis variables (dental caries and enamel fluorosis). In this study, we explored cohorts of 12-year-old children during 30 years of CWF in HCMC, Vietnam, identified by fluoride exposure and sociodemographic variables.

Since 1990, numerous local studies have reported the CWF's effectiveness in preventing and controlling dental caries and monitoring dental fluorosis [13]. After 30 years of implementation, the opportunity presented to evaluate the changes in fluoride dosage on program effectiveness and risk of fluorosis using contemporary methodological approaches.

2. Materials and methods

2.1. Working data preparation

Data were collated from oral examinations of 12-year-individuals residing in Ho Chi Ming City, Vietnam, who participated in 1989, 2003, 2012, and 2019 oral health surveys and corresponding to some of the following historical events.

- 1989 Pre-fluoridation baseline data.
- 1990 Initiation of CWF with a fluoride concentration of 0.7. No data.
- 2000 Adjustment of fluoride concentration to 0.5 ppm. No data.
- 2003 Survey. Twelve-year-old children were exposed to CWF at 0.7 ppm during their lifetime.
- 2012 Survey. Twelve-year-old children were exposed to CWF at 0.5 ppm during their lifetime, first cohort.
- 2019 Survey. Twelve-year-old children were exposed to CWF at 0.5 ppm during their lifetime, second cohort.

The Odonto-Maxillo-Facial Hospital's Ethics Board approved the study (No. 542/BVRHM, 16 Nov 2017). All participants provided signed consent forms.

Each survey collected sociodemographic characteristics, e.g., sex and normative data for dental caries and fluorosis. All surveys used the WHO detection criteria to measure cavitated caries lesions at the tooth (suffix T) and surface (suffix S) level, from which we calculated the number of decayed (D), missing (M), and filled (F) teeth or surfaces.

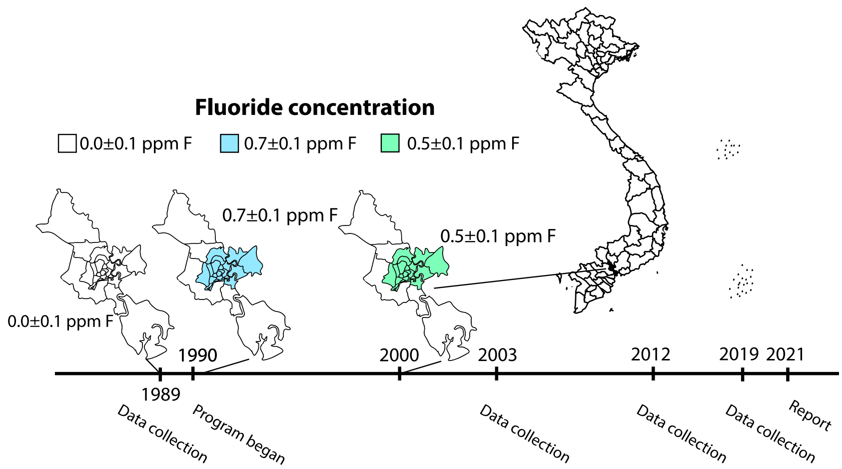

Fig. 1. Diagram of 30 years of water fluoridation in Ho Chi Minh City. Baseline data were collected in 1989. CWF started in 1990 at 0.7 ppm F, then reduced to 0.5 ppm F in 2000. Subsequent data on dental caries and fluorosis were collected in 2003, 2012, and 2019.

Dental fluorosis was estimated with the Dean's Fluorosis Index (DFI). Dean's criteria were applied to all the teeth using the recommended six-level scale: unaffected (0), questionable (0.5), very mild (1), mild (2), moderate (3), and severe (4) [14]. Examiners used natural but not direct light. Per Dean's recommendation, we determined the degree of fluorosis severity based on the two teeth showing the most advanced signs of fluorosis [14]. Also, following Dean's recommendations, we estimated population prevalence by including children with scores "very mild" or higher [14,15]. We used the Community Fluorosis Index (CFI) to measure whether fluorosis levels were of public health importance.

The working datasets had drinking water fluoride concentrations from the HCMC CDC reports in 1990, 2000, 2003, 2012, and 2019. The same selection criteria, sampling methods, and the WHO Oral Health Surveys Basic Methods were used in all surveys. At least one investigator participated in all surveys to maintain integrity in detection criteria. Trained and standardized examiners collected all dental caries and fluorosis data. The 2003 and 2012 surveys included the same examiners. Furthermore, the primary examiner in 2003 and 2012 was the standard examiner for the 2019 survey [16,17].

2.1.1. Demographic comparison

The Chi-Square test was used to compare the educational levels of each parent across three categories: completion of primary school, completion of secondary school, and completion of high school, for each survey.

2.1.1.1. Dataset preparation. Two working datasets were prepared: Dataset 1 for dental caries analysis included 4773 participants with data from 1989, 2003, 2012, and 2019 surveys (Fig. S1A). Dataset 2 for dental caries and fluorosis analysis included 4194 individuals from 2003, 2012, and 2019 (Fig. S1A). Dataset 2 did not include 1989 data because dental fluorosis was not assessed in that survey.

We used dichotomous outcomes (true/false) for dental caries and fluorosis to evaluate the clusters. Boxplots, bar plots, mosaic plots, and flow diagrams were used for illustration using ggplot 2 (v3.3) and ggalluvial (v0.12) in the R statistical package. Log10-transformation and scaling were applied to all numeric variables. The variables included in Dataset 1 were DS, DMFS, DT, DMFT, gender, fluoride concentration (0, 0.5, and 0.7 ppm), and survey year. The variables included in Dataset 2 were DS, DMFS, DT, DMFT, gender, fluoride concentration (0, 0.5, and 0.7 ppm), survey year, and Dean's Fluorosis Index (normal, questionable, very mild, mild, moderate, and severe).

2.2. Cluster analysis

We used the mixed variable-type data clustering procedure (ClustMixType v2.0) in the R statistical package (v4.1, RStudio Inc, USA) [18] for cluster analysis. ClustMixType uses the "k-prototypes partitioning clustering" algorithm by Zhexue Huang [19], which applies a simple matching dissimilarity measure to deal with categorical objects, replaces the means of clusters with modes, and uses a frequency-based method to update ways in the clustering process to minimize the clustering cost function. Also, we calculated the Silhouette index to compute and narrow the optimal number of clusters based on k-prototypes [20]. Its application requires the specification of two hyperparameters: the number of clusters k and a second parameter λ that controls the interplay of the different data types for distance computation. The squared Euclidean distance defined the dissimilarity measure of numeric features [19]. Continuous data should be scaled using the Euclidean distance, such as in K-prototypes. We applied transformed and scaled continuous data as input [21]. Outliers were identified by Euclidean distance in which the darkest colored areas correspond to candidates who can be considered outliers. We assigned a certain proportion of cases with the highest distance (5 %), we then isolated these individuals from the population. The outliers were <0.5 %; hence, we did not exclude them.

The k-proto function of the ClustMixType package was applied to compute k-prototypes clustering with k from 1 to 10. In k-prototypes, through probing many values of the number of clusters from 2 to 5 based on the appropriateness of the data and cluster distribution. In our study, we chose k = 3. The k-prototypes algorithm is a partition clustering algorithm that uses input: initial data set X and number of clusters k; output: k-sample objects such that the standard function reaches the minimum value. First, to initialize k-sample objects for X, each object acts as the representative center of each cluster. Secondly, distribute each object in X to each cluster so that they are closest to the sample object in the cluster, and update the sample object for each cluster. After all the objects have been distributed to the clusters, check the similarity of the objects in each cluster with the sample objects. If there is a sample object that is most similar it different from the sample object in the current cluster, move this considered object to the cluster corresponding to the closest sample object and simultaneously update the sample objects for these clusters. Finally, repeat until no objects change after checking all objects. This checking process is repeated until we reach a state where all objects have been assigned to their appropriate cluster.

The clustering results were interpreted using clinical and diagnosis variables (Figs. S1D, E, S2D, E). In Dataset 2, we also calculated the Community Fluorosis Index (CFI). For that purpose, we used standard CFI weights (0–4) and cut-off points to assess public health significance according to Dean's recommendations (negative = CFI<0.4 borderline = 0.4–0.6, slight = 0.6–1.0 is slight, medium = 1.0–2.0, market = 2.0–3.0, and very marked = 3.0–4.0) [22].

We used alluvial diagrams to visualize the groups and clusters. Alluvial diagrams are a type of flow diagram developed to represent changes in network structure [23]. These diagrams can show differences in group composition between conditions, including statistical information, especially the frequency distribution of numeric or categorical variables in absolute values or percentages.

3. Results

3.1. Demographic comparison

Because we used data from different historical times, secular changes are a potential source of bias. The only demographic variable available for such comparison across all surveys was parents' education level. Thus, we compared the available three-level education of each parent (completed primary school, completed secondary school, and completed High school) in each survey. We found no statistically significant differences in the father's and mother's level of education between fluoridated and non-fluoridated in each survey. However, the proportion of higher-education parents increased with time.

3.2. K-prototypes cluster analysis identified three clusters for dental caries in dataset 1

The distribution of dental caries indices by dental caries (true/false) and fluoride concentration (0, 0.5, and 0.7 ppm) are shown in Fig. 2A and B. As expected, children living in non-fluoridated areas had higher DMFS, DS, DMFT, and DT scores. In contrast, children in districts exposed to 0.5 and 0.7 ppm F in drinking water had lower disease scores. Scores at 0.5 and 0.7 ppm F concentrations were similar (Fig. 2B).

The algorithm identified three clusters in Dataset 1. Clusters 1 and 2 had low DMFS, DS, DMFT, and DT scores in contrast with Cluster 3, with high dental caries scores (Fig. 2C). Furthermore, while all children in Cluster 3 had dental caries, most caries-free children (>75 %) were in Cluster 1; Cluster 2 was in between (Fig. 2D). We added detailed prototypes for each cluster in Figs. S1B and C. Regarding fluoride concentration, around 80 % of children in Cluster 1 lived in Districts with 0.5 ppm F. Approximately 75 % of children in Cluster 2 lived in Districts with 0.7 ppm F. Sixty percent of children in Cluster 3 lived in no-fluoridated Districts (Fig. 2D). From a different point of view, 75 % of children in Cluster 1 lived in Districts with 0.7 ppm, and 60 % of children in Cluster 3 lived in non-fluoridated Districts (Fig. 2E).

Alluvial diagrams showed the same characteristics for the three clusters (Fig. 3). By fluoride concentration, most children from non-fluoridated areas had dental caries. In contrast, children from 0.5 to 0.7 ppm F areas did not (Fig. 3A and B). Cluster 1 consisted of 80 %

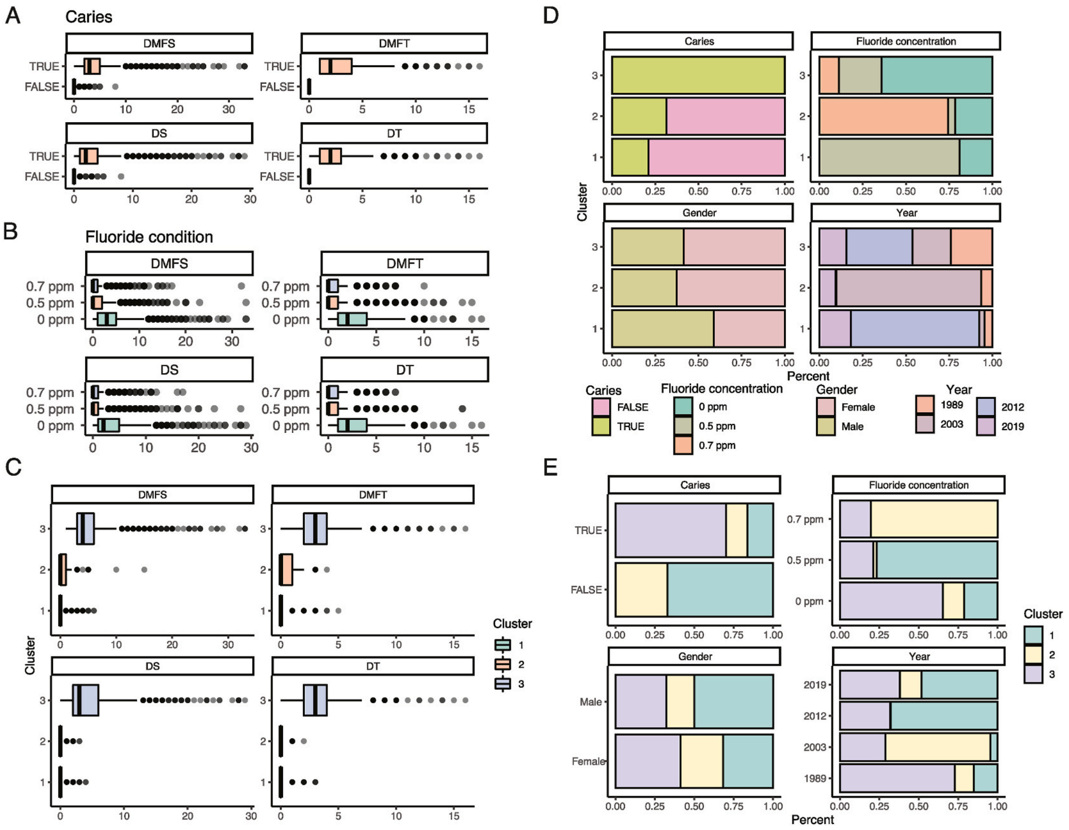

Fig. 2. Data exploration and clustering results for dental caries in dataset 1. A) DMFS, DMFT, DS, DT index distributions of caries (true) and caries-free (false) children. B) DMFS, DMFT, DS, DT index distributions of 0, 0.5, and 0.7 ppm F area subjects. C) DMFS, DMFT, DS, and DT index distributions of 3 clusters. D) Caries diagnosis, fluoride concentration, gender, and year ratio of 3 clusters. E) Cluster distribution of categorical variables (caries diagnosis: caries/true and caries-free/false, fluoride concentration, gender, and year of data collection).

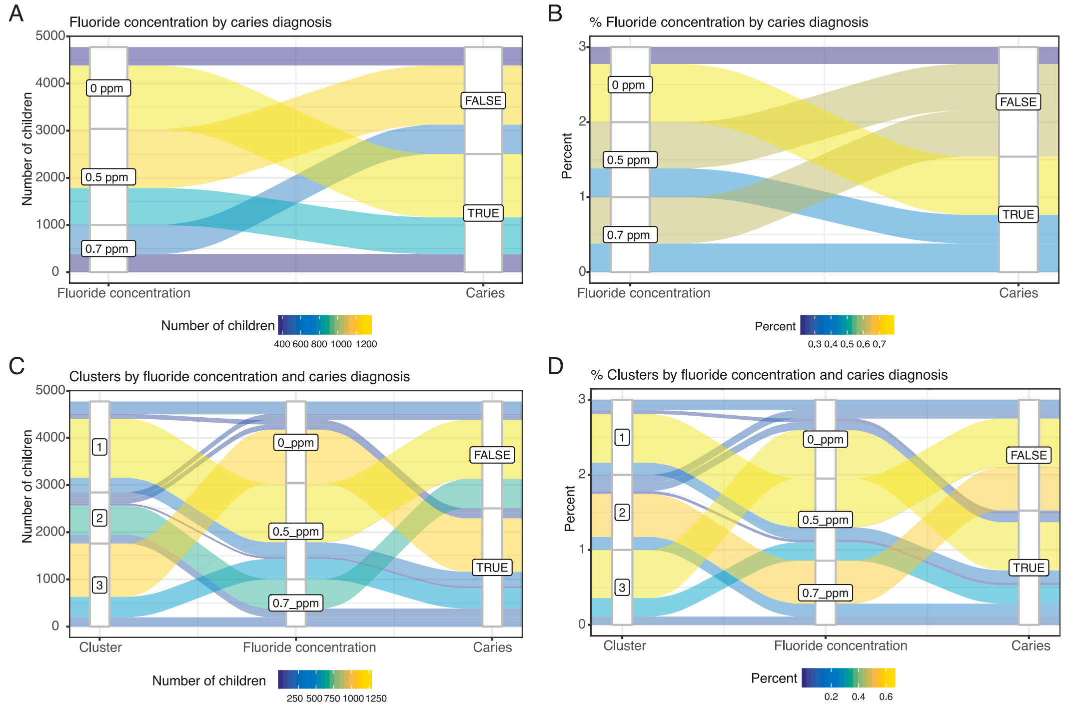

Fig. 3. Alluvial diagrams of fluoride concentration and clusters for dental of caries in dataset 1. A) Fluoride concentration distribution by caries diagnosis: caries/true and caries-free/false (presented as the absolute number of subjects). B) Fluoride concentration distribution by caries diagnosis: caries/true and caries-free/false (presented as percentage). C) Cluster distribution by fluoride concentration and caries diagnosis: caries/true and caries-free/false (presented as the absolute number of subjects). B) Cluster distribution by fluoride concentration and caries diagnosis: caries/true and caries-free/false (presented as a percentage).

0.5 ppm F and caries-free subjects. Cluster 2 consisted of 75 % 0.7 ppm F and caries-free subjects. Cluster 3 consisted of 60 % fluoride-free and participants with dental caries (Fig. 3C and D). It is worth noticing that children living in fluoridation areas (0.5 and 0.7 ppm F) had a similar percentage of children with no dental caries or sequelae (Fig. 3D). These children were in Clusters 1 and 2, respectively.

Overall, cluster 1 had low DMFS, DS, DMFT, and DT indices, high caries-free prevalence, and almost all participants lived in a 0.5 ppm F area in 2012. Cluster 2 had low DMFS, DS, DMFT, and DT indices, high caries-free prevalence, and almost all participants lived in a 0.7 ppm F area in 2003. Cluster 3 had high DMFS, DS, DMFT, and DT indices, low caries-free prevalence, and almost all participants lived in a non-fluoridated area at all times. Notably, 0.5 and 0.7 ppm F areas (mostly in Clusters 1 and 2) had the same effect on caries-free outcomes. In summary, Cluster 1 was characterized by caries-free children living in a 0.5 ppm F area. Cluster 2 was also characterized by caries-free children who lived in a 0.7 ppm area. Cluster 3 included children with dental caries who lived in non-fluoridated areas.

3.3. K-prototypes cluster analysis identified three clusters for dental fluorosis in dataset 2

Fig. 4A–S2A, B, and C show the distribution of indices on dental caries and fluorosis for Dataset 2.

The algorithm identified three clusters in Dataset 2. Dataset 2 had 479 fewer observations than Dataset 1. We repeated the cluster analysis for dental caries and found no difference from the values reported in the previous section (Fig. 4). We included categorical variables (Dean's index, fluoride concentration, gender, year) and diagnosis variables (caries, fluorosis) in Fig. 4C and D. Cluster 1 showed the lowest fluorosis prevalence (20 % in the cluster had very mild or more dental fluorosis), followed by Cluster 2 (25 %) and Cluster 3 (30 %). Dean's index distribution also showed that most children in Cluster 1 had no fluorosis, while Cluster 3 had the highest moderate values (about 6 %). Cluster 2 included participants with no and very mild DFI scores. Consistently, 50 % of children in Cluster 1 lived in non-fluoridated districts; 85 % of participants in Cluster 2 lived in Districts with 0.5 ppm F; and 80 % of participants in Cluster 3 lived in Districts with 0.7 ppm F in the drinking water.

Alluvial diagrams also expressed the same characteristics of the three clusters (Fig. 5A and B). Children from 0 ppm F areas had dental caries and were fluorosis-free, while children from 0.5 to 0.7 ppm F areas did not (Fig. 5A and B). Cluster 1 contained >50 % of 0 ppm F, caries, and fluorosis-free subjects (Figs. S2F and G). Cluster 2 comprised 0.5 ppm F, caries-free, and fluorosis-free subjects (60 %). Cluster 3 comprised less than 40 % of 0.5 ppm F, caries-free, and fluorosis-free subjects.

The Community Fluorosis Index (CFI) for Clusters 1, 2, and 3 were 0.29, 0.33, and 0.54, respectively. Thus, based on the Dean's

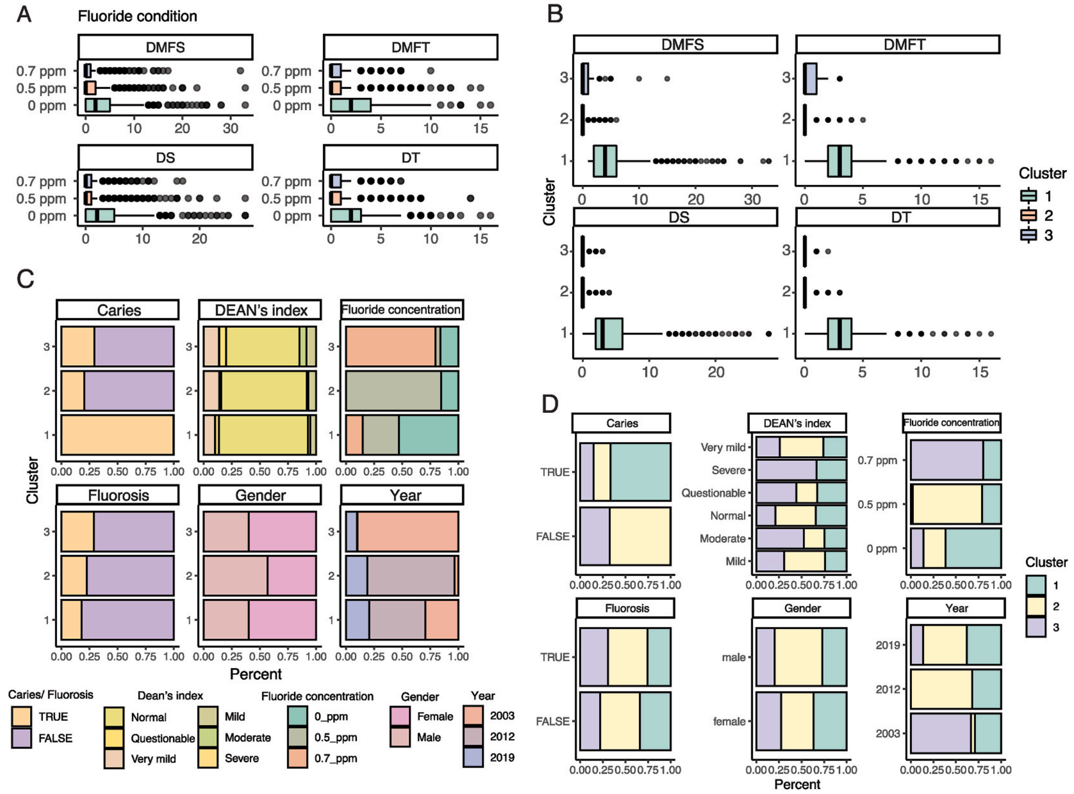

Fig. 4. Data exploration and clustering results for dental fluorosis in dataset 2. A) DMFS, DMFT, DS, DT index distributions of 0, 0.5, and 0.7 ppm F area subjects. B) DMFS, DMFT, DS, and DT index distributions of 3 clusters. C) Caries diagnosis: caries/true and caries-free/false, fluorosis diagnosis, Dean's index, fluoride concentration, gender, and year distributions of 3 clusters. D) Cluster distribution of categorical variables (caries diagnosis: caries/true and caries-free/false, fluorosis diagnosis, Dean's index, fluoride concentration, gender, and year of data collection).

criteria, the CFI in Cluster 3 (0.4–0.6) was of borderline public health importance.

In summary, we identified three clusters. One cluster included children living in non-fluoridated areas, having higher levels of dental caries and lower fluorosis prevalence and severity. A contrasting cluster included children living in Districts with 0.5 ppm F areas, having lower levels of dental caries and levels of dental fluorosis that are not of public health significance according to Dean's criteria. Finally, a third cluster, included children living in Districts with 0.7 ppm F, also showed lower levels of dental caries than the first cluster but with levels of fluorosis that were of borderline public health significance.

4. Discussion

The water fluoridation program in HCMC started in January 1990 [3–7]. However, seven out of 17 districts cannot access the public water system (Fig. 1). We labeled these districts as non-fluoridated even though there may be a halo effect when their residents consumed goods manufactured in fluoridated areas.

Since its inception, the program has been monitored for dental caries and fluorosis using the non-fluoridated areas as control. As part of the monitoring, a high prevalence of dental fluorosis was reported among 8-year-old children living in fluoridated areas in 1998 [17]. Consequently, in 2000, the local health authorities decided to lower the F recommended levels to 0.5 ppm F. Hong Kong used the same justification to reduce the original 1 ppm F to 0.7 and then to 0.5 ppm [8,9]. When considering the risks and benefits, water fluoridation is a suitable community-based approach for delivering fluoride because it is effective (benefits), and the main side effect (risk) at the recommended concentrations is dental fluorosis in its milder forms. Also, water fluoridation benefits the entire community regardless of socioeconomic status. Community water fluoridation is an effective dental public health intervention [24] and has been identified by the U.S. CDC as one of the ten most successful public health interventions of the 20th century [25].

Our analysis identified three clusters in dental caries and fluorosis. One cluster had a high prevalence and severity of dental caries while low in similar measures of dental fluorosis. The participants in this cluster live in non-fluoridated Districts. Another cluster had a lower prevalence and severity of dental caries and a higher prevalence of dental fluorosis but of no public health concern. These participants consumed drinking water at 0.5 ppm F. The third cluster also had a lower prevalence and severity of dental caries, but the prevalence and severity of dental fluorosis were higher and at borderline public health concern levels. These children consumed drinking water at 0.7 ppm F. The contrast of clusters 2 and 3 suggests that the preventive effect on dental caries was maintained after

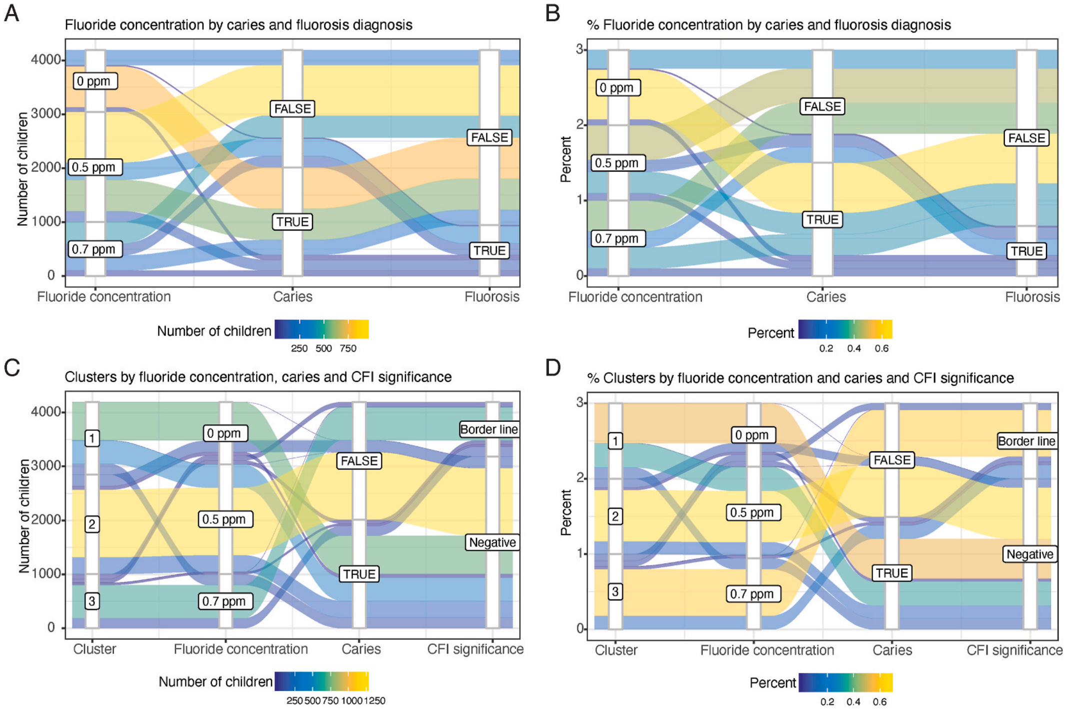

Fig. 5. Alluvial diagrams of fluoride concentration and clusters for dental caries in dataset 2. A) Fluoride concentration distribution by caries and fluorosis diagnosis (presented as an absolute number of subjects). B) Fluoride concentration distribution by caries and fluorosis diagnosis (presented as a percentage). C) Cluster distribution by fluoride concentration, caries diagnosis: caries/true and caries-free/false, and CFI significance (presented as an absolute number of subjects). B) Cluster distribution by fluoride concentration, caries diagnosis: caries/true and caries-free/false, and CFI significance (presented as percentage).

reducing the fluoride concentration from 0.7 ppm F to 0.5 ppm F while keeping the risk of dental fluorosis below public health concerns. Thus, our results confirm the effectiveness of the CWF program that started in 1990 and support the decision to reduce the recommendation to 0.5 ppm to reduce the risk of fluorosis. Furthermore, our results support the expansion of CWF into other Vietnam areas with comparable water systems [26,27].

Methodologically, our study was a cluster analysis that included categorical and numeric variables. Thus, our choice for the k-prototypes algorithm is an approach suitable for chronic diseases like dental caries.

Our working datasets combined different 12-year-old cohorts over 30 years. These cohorts have various and not necessarily similar live-course events during their formative years, which may affect our results. Indeed, the HMCM City population changed demographically and economically from 1989 to 2019. For example, the population increased from 3988 million to 8993 million, and the median income rose to 6417 USD. Dental caries has strong sociodemographic and commercial determinants [28–31], which may have contributed to changes in dental caries in the same 30 years. On the other hand, water fluoridation levels and fluoridation areas were very consistent during the same period. Unfortunately, we cannot assess these secular effects in our study because the study design in 1989 did not include parameters to control for these effects on a long-term basis, and the statistical methodology we used in our research was unavailable.

Our study has the following limitations. First, we focused on 12-year-old children, which may limit the generalizability of the findings to other age groups. However, 12-old-year children are a WHO age indicator [15]. In addition, we cannot assure that all 12-year-old participants in each cluster lived in the same place during their lifespan because residential history was not included in the survey study designs. The fluoride concentration was allocated to all participants in the same fluoridated group and assumed to be constant with continued use of tap water, which may not be the case. However, this limitation applies to all community fluoridation studies.

Another limitation resides in the k-prototype approach, which requires a priori knowledge of the datasets, and the number of clusters is unknown [19]. Also, it is difficult to determine the key variables affecting clustering.

Our study has the following strengths: first, we studied CWF over 30 years in one analytical approach instead of multiple cross-sectional evaluations. Thus, it supplements the published information regarding CWF effectiveness in HCMC and provides insights into temporal changes. Second, it included a large sample of 4773 12-year-old children, increasing the statistical power and enhancing the reliability of the findings. Third, we used a novel methodological approach to understand the associations between variables and outcome measures comprehensively. Fourth, we relied on systematically collected dental caries and fluorosis data at four landmark moments in HCMC CWF history.

5. Conclusion

Mixed-type clustering analysis identified three population groups linked by dental caries and fluorosis prevalence and severity, water fluoridation levels, and sex for 30 years of monitoring and supports benefits and risk-controlled interventions in HCMC. Thus, our research builds on previous studies on the effectiveness of CWF in preventing dental caries and the appropriateness of reducing the recommended levels to reduce the risk of fluorosis while maintaining its effectiveness.

Data availability statement

The scripts, and additional materials are available at https://github.com/namhuynhnc/HCMC-dental-caries-and-fluorosis-. The raw data will be available from the authors upon reasonable request.

CRediT authorship contribution statement

Hung Trong Hoang: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Trang Thi-Ngoc Tran:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Formal analysis, Data curation. Nam Cong-Nhat Huynh: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Formal analysis. **Minh Duc Nguyen:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Eugenio D. Beltrán-Aguilar:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Department of Dental Public Health, Faculty of Odonto-Stomatology, University of Medicine and Pharmacy at Ho Chi Minh City, and the Ho Chi Minh City Hospital of Odonto-Stomatology for supporting this study. The study belongs to the scientific research and technology project "Oral Health of Ho Chi Minh City Residents and Related Factors," which has been approved and funded by the Scientific and Technological Council of the Ho Chi Minh City, Department of Science and Technology under Decision 338/QD-SKHCN dated May 16, 2017.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e25035.

References

[1] T.B. Nguyen, et al., Saigon-ho chi minh city, Cities 50 (2016) 16–27.
[2] T.Q. Vo, Water fluoridation in Ho Chi Minh City for dental caries prevention, in: Vietnam-France Conference 1-1990, 1990. Vietnam-France Conference 1-1990.
[3] T.H.Q. Dao, T.H. Hoang, Analysis of Risk Factors for Dental Fluorosis Among Primary School Students in District 5, Ho Chi Minh City. Medical Publishing House, Proceeding of Scientific Research Papers on Odonto-Stomatology), 2001, pp. 197–209.
[4] T.H.Q. Dao, T.H. Hoang, D.T. Tran, Dental Caries Status of 12 and 15-Year-Old Children after 12 Years of Water Fluoridation in Ho Chi Minh City. Medical Publishing House, Proceeding of Scientific Research Papers on Odonto-Stomatology, 2004, pp. 72–76.
[5] D.T. Tran, et al., Progression of dental caries status in 12-year-old children after 12 years of water fluoridation in Ho Chi Minh City, Journal of Medicine of Ho Chi Minh City 11 (2004).
[6] T.T.H. Nguyen, Prevalence of Dental Fluorosis Among Students in a Primary School in District 5, Ho Chi Minh City. Medical Publishing House, Proceeding of Scientific Research Papers on Odonto-Stomatology, 2001, pp. 52–56.
[7] T.H. Hoang, Enamel fluorosis in 12 and 15 year-old children in HCMC VietNam, J. Dent. Res. 83 (2004) (Special Issue A).
[8] R.W. Evans, E.C. Lo, O.P. Lind, Changes in dental health in Hong Kong after 25 years of water fluoridation, Community Dent. Health 4 (4) (1987) 383–394.
[9] R.W. Evans, Changes in dental fluorosis following an adjustment to the fluoride concentration of Hong Kong's water supplies, Adv. Dent. Res. 3 (2) (1989) 154–160.
[10] Official Letter on Reducing the Fluoride Concentration in Tap Water in Ho Chi Minh City, Ho Chi Minh City People's Committee, 2000.
[11] A. Flynt, N. Dean, A survey of popular R packages for cluster analysis, J. Educ. Behav. Stat. 41 (2) (2016) 205–225.
[12] S. Selinski, K. Ickstadt, Cluster analysis of genetic and epidemiological data in molecular epidemiology, J. Toxicol. Environ. Health 71 (11–12) (2008) 835–844.
[13] T.T. Thuy, et al., Fluoride profiles in premolars after different durations of water fluoridation in Ho Chi Minh City, Vietnam, Arch. Oral Biol. 48 (5) (2003) 369–376.
[14] Moulton, F.R. Fluorine and dental health, American Association for the Advancement of Science, 1942.
[15] W.H. Organization, Oral Health Surveys: Basic Methods, World Health Organization, 2013.
[16] T.H. Hoang, Changes in dental caries among 3- and 5-year old children from 1989 to 2012 in Ho Chi Minh city, Vietnam, Journal of Medicine of Ho Chi Minh City 20 (2) (2016) 263–271.

[17] T.H. Hoang, D.K. Ngo, Changes of dental caries and enamel fluorosis among 8-year-old children after adjusted fluoride level in drinking water in Ho Chi Minh city, Vietnam, Journal of Medicine of Ho Chi Minh City 17 (4) (2013) 222–228.

[18] G. Szepannek, clustMixType: user-friendly clustering of mixed-type data in R, The R Journal 10 (2) (2018) 200–208.

[19] Z. Huang, Extensions to the k-means algorithm for clustering large data sets with categorical values, Data Min. Knowl. Discov. 2 (3) (1998) 283–304.

[20] R. Aschenbruck, G. Szepannek, Cluster validation for mixed-type data, Archives of Data Science, Series A 6 (1) (2020) P02–12.

[21] G. Preud'homme, et al., Head-to-head comparison of clustering methods for heterogeneous data: a simulation-driven benchmark, Sci. Rep. 11 (1) (2021) 4202.

[22] H.S. Horowitz, Indexes for measuring dental fluorosis, J. Publ. Health Dent. 46 (4) (1986) 179–183.

[23] M. Rosvall, C.T. Bergstrom, Mapping change in large networks, PLoS One 5 (1) (2010) e8694.

[24] J.V. Kumar, Is water fluoridation still necessary? Adv. Dent. Res. 20 (1) (2008) 8–12.

[25] Centers for Disease, C. and Prevention, Ten great public health achievements–United States, 1900-1999, MMWR Morb. Mortal. Wkly. Rep. 48 (12) (1999) 241–243.

[26] D. Huang, et al., Maternal and child nutrition and oral health in urban Vietnam, Int. J. Environ. Res. Publ. Health 16 (14) (2019).

[27] T.T. Nguyen, The Relationship between Fluoride Concentration in Drinking Water with Dental Caries and Fluorosis in Vietnamese Children, 2001.

[28] S.A. Fisher-Owens, et al., Influences on children's oral health: a conceptual model, Pediatrics 120 (3) (2007) e510–e520.

[29] R.G. Watt, et al., The lancet oral health series: implications for oral and dental research, J. Dent. Res. 99 (1) (2020) 8–10.

[30] J.L. Rodriguez, et al., Caries risk and social determinants of health: a big data report, J. Am. Dent. Assoc. 154 (2) (2023) 113–121.

[31] L. Jamieson, B. Gibson, W.M. Thomson, Oral health inequalities and the corporate determinants of health: a commentary, Int. J. Environ. Res. Publ. Health 17 (18) (2020).