Egyptian Informatics Journal (2016) 17, 251–263

Cairo University

Egyptian Informatics Journal

www.elsevier.com/locate/eij
www.sciencedirect.com

FULL-LENGTH ARTICLE

Application of fuzzy consensus for oral pre-cancer and cancer susceptibility assessment

Satarupa Banerjee [a,*], Aishwaryaprajna [b], Debjani Chakraborty [c], Amita Giri [d], Ranjan Ghosh [e], Badal Chandra Sarkar [e], Jyotirmoy Chatterjee [a]

[a] *School of Medical Science and Technology, Indian Institute of Technology, Kharagpur, 721302, India*
[b] *School of Mathematics, Techno India University, Kolkata 700091, India*
[c] *Department of Mathematics, Indian Institute of Technology, Kharagpur, 721302, India*
[d] *Department of Pathology, North Bengal Medical College and Hospital, Darjeeling 734432, India*
[e] *Department of Oral and Maxillofacial Pathology, North Bengal Dental College and Hospital, Darjeeling 734432, India*

Received 1 January 2015; accepted 5 September 2015
Available online 3 March 2016

KEYWORDS

Fuzzy rulebase;
Oral health;
Oral habit;
Susceptibility assessment;
Oral pre-cancer and cancer

Abstract Health questionnaire data assessment conventionally relies upon statistical analysis in understanding disease susceptibility using discrete numbers and fails to reflect physician's perspectives and missing narratives in data, which play subtle roles in disease prediction. In addressing such limitations, the present study applies fuzzy consensus in oral health and habit questionnaire data for a selected Indian population in the context of assessing susceptibility to oral pre-cancer and cancer. Methodically collected data were initially divided into age based small subgroups and fuzzy membership function was assigned to each. The methodology further proposed the susceptibility to oral precancers (viz. leukoplakia, oral submucous fibrosis) and squamous cell carcinoma in patients considering a fuzzy rulebase through *If-Then* rules with certain conditions. Incorporation of similarity measures using the Jaccard index was used during conversion into the linguistic output of fuzzy set to predict the disease outcome in a more accurate manner and associated condition of the relevant features. It is also expected that this analytical approach will be effective in devising strategies for policy making through real-life questionnaire data handling.

* Corresponding author. Tel.: +91 3222 281225; fax: +91 3222 282221.
E-mail addresses: satarupa@smst.iitkgp.ernet.in, satarupabando@gmail.com (S. Banerjee).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

ELSEVIER | Production and hosting by Elsevier

1. Introduction

India experiences one of the highest incidence rates of oral cancer globally [1]. Oral cancer is the leading cancer type in men and the third most common cancer in women [2]. In India, oral cancer is usually detected at advanced stages and the five year survival rate for advanced oral cancer is very low [3], posing an important public health challenge. Hence the early detection of

http://dx.doi.org/10.1016/j.eij.2015.09.005
1110-8665

these diseases is urgency. The most important risk factors for these cancers are tobacco and betel quid/areca nut use in some regions of Asia, including India [4,5], where betel quid/areca nut is commonly chewed with smokeless tobacco (SLT) [6]. Oral carcinogenesis is a multistep phenomenon, which often progresses from intermediate oral precancers [e.g. leukoplakia (OLK), oral submucous fibrosis (OSF)], to carcinoma in situ and then to malignant oral squamous cell carcinoma (OSCC). The OLK is defined by white plaques of questionable risk having excluded (other) known diseases or disorders that carry no increased risk of cancer [7]. Annually, approximately 2–3% of oral lesions show malignant transformation. Long term follow-up studies clinically suggest that OLK with severe dysplasia is more susceptible to transformation into OSCC [8]. OSF is a chronic, premalignant condition, characterized by progressive submucosal fibrosis inside oral cavity [9]. Despite differences in their origin, all precancers converge into oral squamous cell carcinoma rapidly.

In our study, the two oral precancers were chosen due to their definite cause–effect relation with the tobacco and related product, areca nut and related materials. Cigarette smoking was found to be considered as one of the major factors in the etiology of OLK [10], while for OSF, the major etiology of the disease considered was chewing areca nut [9].

1.1. Literature review for rulebase construction

Consumption of tobacco is the major cause of death and disability worldwide. It is obtained from Nicotiana tabacum. When broadly classified, tobacco is either smoked or SLT "a large variety of commercially or non-commercially available products and mixtures that contain tobacco as the principal constituent and are used either orally or nasally without combustion" [11]. Many of the components present in tobacco are mutagenic. The use of SLT varies by age, sex, ethnicity and socioeconomic status, both within and among countries [11]. Both the prevalence and severity of tobacco-related oral lesions demonstrate a dose–response relationship with the amount, frequency and duration of SLT exposure. The chronic exposure can lead to OLK [12]. SLT use in the United States has been associated with an increased risk for oral cancer in a dose–response fashion [13].

In Indian population role of Bidi and SLT is well known [14]. When smoking tobacco is to be considered, in a meta analysis, Bidi smoking had higher odds ratio (OR) than cigarette smoking [15]. So Bidi smoking was considered more harmful than cigarette smoking. Cigarette smoke condensate was found to enhance matrix metalloproteinases (MMPs), MMP-2 and MMP-9 expression and thus increase collagen degradation which ultimately increases chance of metastasis in cancer patients [16]. Considering the synergistic effect, if any person was exposed to both of the smoking agents on daily basis, they were considered to have maximum susceptibility for cancer.

Study suggested role of betel quid without tobacco consumption for oropharynx and esophagus cancer. As OSCC was considered, we took account of chewing habit as potentially carcinogenic agent. In a study it was considered that, irritation caused due to high frequency of chewing may cause oral precancer (OSF) even without tobacco [17]. Even dose response relationship in frequency and duration of betel quid chewing without tobacco was found to have an elevated risk for OSF [18]. Even chewing betel quid with or without tobacco itself is considered as an independent cancer causing factor [19].

Though the Indian scenario is unknown, age and cultural background are important variables influencing oral health-related quality of life. Poor oral health is another major underlying cause of carcinogenesis beside the tobacco and betel quid/areca consumption habits. The younger age groups showed an increase in the proportion of individuals free from caries and restorations. Again globally, poor oral health among older people has been particularly evident in high levels of tooth loss, dental caries experience, and the prevalence rates of periodontal disease, xerostomia and oral precancer/cancer [20]. In India, there are many things to do for upgrading oral health awareness [21]. Again a direct relation was noted between the favorable dental health awareness, attitude, oral hygiene behavior, and socioeconomic status in the Indian population [22]. Even poor oral health was considered as an independent causative condition of OSCC [23]. The study suggested that educational level influences the oral conditions and should be considered in assessing risk, and in planning appropriate preventive measures as health literacy also has impact on oral health [24,25]. Therefore, oral health literacy and oral hygiene can be taken into account while understanding the malignant potentiality and susceptibility to oral precancers and cancer. Level of schooling in Indian education system was used in flexible manner to define oral health education in this study.

1.2. Fuzzy logic and epidemiology

Fuzzy logic is used widely to interpret uncertain knowledge present in a system and includes vague human assessment in problems which are not considered in any conventional computing methods. It can also be considered as an approach of computing with words as linguistic language is always preferred for expressing opinions. The beauty of real-life application of fuzzy logic lies in the precision in meaning of an outcome and getting an idea of a complex system with tolerance of imprecision [26].

In modeling problems, words are often led to use predicates in natural languages to represent incomplete information in a flexible way. The information may be quantifiable due to its nature, and can be stated only in linguistic terms. So to quantify the linguistic expressions here fuzzy numbers must have to be defined. Defining fuzzy numbers allows modeling of complex systems using a higher level of abstraction originating from defined knowledge and experience. A fuzzy number $\tilde{A}$ [27] is a convex normalized fuzzy set defined on the universe of discourse $\Re$ [the set of all real numbers) with a piecewise continuous membership function and bounded support.

Here, a fuzzy number $\tilde{A}$ was used as a triangular fuzzy number (TFN) denoted by $\tilde{A} = (m, \beta, \gamma)$ if its membership function is of the following form

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 - \frac{m-x}{\beta} & \text{if } m - \beta \leqslant x \leqslant m \\ 1 - \frac{x-m}{\gamma} & \text{if } m \leqslant x \leqslant m + \gamma \end{cases} \quad [28]$$

$\tilde{A}$ may also be represented by $(\underline{a}, a, \bar{a})$ where $\underline{a} = m - \beta$, $a = m$ and $\bar{a} = m + \gamma$ denote the left point, center and right point of $\tilde{A}$.

In epidemiological data analysis, there often remains an uncertainty in the form of individual exposure estimates [29] and increasing knowledge gaps [30]. When real-life situations are taken into account, it is often seen that the number of daily consumption of addictive product is not always fixed, or illiterate as well as people without proper awareness often cannot provide specific information on year of onset of addictive habit. In turn, they can provide information on a tentative number or year. Prediction of chance of occurrence or the prevalence of a disease is also an uncertain situation [31]. This impreciseness can be well interpreted applying the consensus of fuzzy logic in the epidemiological data analysis. Similarity measures [32] have been proposed recently to measure the degree of similarity in fuzzy sets for better understanding the output of the dataset. Conventional statistical analysis of clinicoepidemiological data does not consider such human perception for information extraction from massive data to get a crisp solution as does the fuzzy logic [33]. The most used method for questionnaire data analysis is a logistic regression technique which can describe the comparative relationship between the response variable and the predictor variables. Data pooling, cleaning, stratification, etc., are needed before data analysis in the conventional procedure and are also a comparatively simpler process and easier to interpret, but the outcome cannot help in the prediction of real-life scenario considering the complex vagueness of situations. Therefore introduction of the notion of logical fuzzy *If-Then* rule to understand such a complex process can provide a better information in the decision making process for a certain range of uncertainty [34].

Thus, this study, intends to apply a fuzzy rule – base for better prediction of malignancy or pre-malignancy susceptibility viz. OLK, OSF and OSCC other oral disease from as well as mathematical validation of consideration of physician's assumptions and conclusions of previous epidemiological studies in disease prediction chances assigning fuzzy rulebase. It would further help the health caregiver to predict chances of disease occurrence and public health policy makers in public health prevention efforts.

2. Methods

2.1. Subject selection

Study population of Terrain and Duars region, the northern region of West Bengal, India, was chosen in this study. However this population is known for its cultural diversity of the people in this area, but not well studied in respect of their oral health. The different social communities of the region include Nepali, Bhutia, Mech, Rajbanshi, Lepcha, Rava, Drupka and Sherpa. Multi-ethnicity is unique in this dataset.

2.2. Data collection strategy

A cross-sectional study of three months (February, March and December 2013) was performed at the North Bengal Dental College and Hospital (NBCDH) in two phases. 938 patients (512 males and 426 females) age 18 years onward who attended the outpatient department of the hospital for treatments related to oral health were interviewed using a pretested predesigned and structured oral health habit related questionnaire. The data were collected on age, gender, education level, the presence of oral lesions, alcohol drinking, tobacco smoking type and frequency, tobacco chewing type and frequency, areca nut and leaves chewing frequency and brushing habits. Prior written informed consent from all the patients and by the concerned authorities was taken (from the NBCDH

Table 1 Fuzzy scale for the input variables (in 0–1).

Variable	Linguistic scale	Fuzzy scale	Variable	Linguistic scale	Fuzzy scale
Oral health literacy (X1) [based on schooling level in Indian education system]	Bad Poor Satisfactory Good Very good	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)	*Smoking type* (X5) [based on materials used like Beedi and cigarette]	Low Medium High	(0, 0, 0.50) (0, 0.50, 1) (0.50, 1, 1)
Oral hygiene (X2) [assessed on brushing modality like with toothpaste, powder, kalamanjan, lalmanjan, daantan, sand, gul, etc.]	Bad Poor Satisfactory Good Very good	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)	*Smoking frequency* (X6) [frequency of consumption of products divided into five ways, i.e. occasional, less than 2/day, 2–5/day, 5–10/day and more than 10/day]	Very Low Low Moderate High Very High	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)
SLT type (X3) [based on materials used like Khaini, Gutkha, Zarda, Nassi and Gudaku and related products]	Very Low Low Moderate High Very High	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)	*Chewing habit type* (X7) [based on the product used like betel quid, Areca leaves, and area nut consumption]	Low Medium High	(0, 0, 0.50) (0, 0.50, 1) (0.50, 1, 1)
SLT frequency (X4) [frequency of consumption of products divided into five ways, i.e. occasional, less than 2/day, 2–5/day, 5–10/day and more than 10/day]	Very Low Low Moderate High Very High	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)	*Chewing habit frequency* (X8) [frequency of consumption of products divided into five ways, i.e. occasional, less than 2/day, 2–5/day, 5–10/day and more than 10/day]	Very low Low Moderate High Very High	(0, 0, 0.25) (0, 0.25, 0.50) (0.25, 0.50, 0.75) (0.50, 0.75, 1) (0.75, 1, 1)

Institutional Ethical Committee of dated 07.05.2012)s for the study. Histopathology was performed with incision biopsies collected from the patients provisionally diagnosed with oral precancers and cancers for confirmation.

2.3. Data analysis strategy

2.3.1. Conventional statistical analysis

Clinicoepidemiological data of the studied subjects were analyzed statistically using SPSS version 17 for risk estimation analysis and primary selection of features to be used in application of fuzzy consensus. During this analysis, each input variable between patients with and without oral lesions was compared using Pearson's χ^2 test [35] and the cutoff significance was established at $p < 0.01$. At 95% confidence interval OR was also calculated. The value of the OR if is greater than 1, indicates toward increased risk, whereas if less than 1 indicates their protective nature.

2.3.2. Defining fuzzy numbers and rulebase generation

In a fuzzy decision making system, defining membership function and fuzzy inference rule to map linguistic variable from

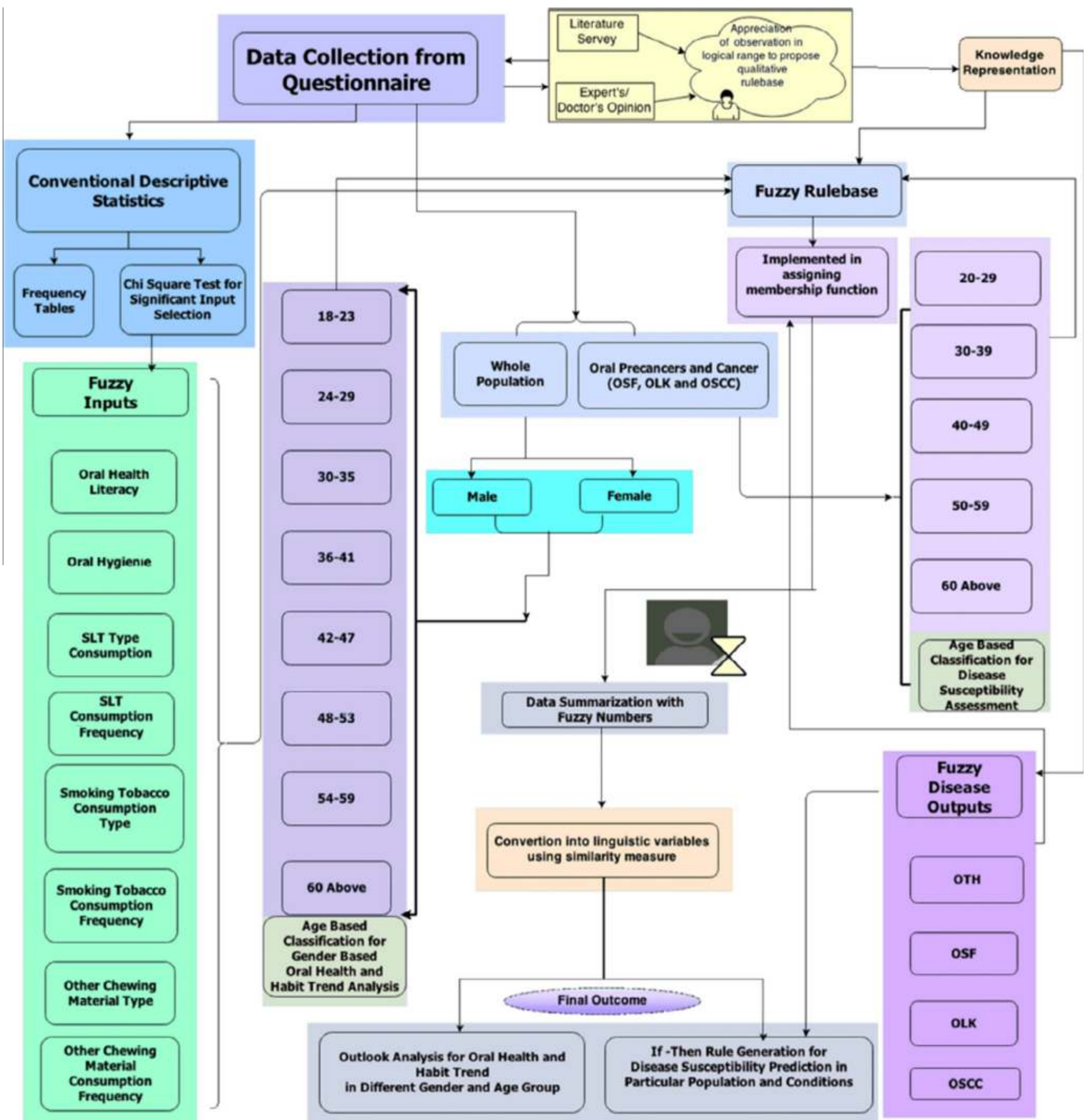

Figure 1 Proposed methodology on application of fuzzy consensus in assessing oral precancer and cancer susceptibility.

numeric data for fuzzy reasoning is important [36]. During the fuzzy rulebase generation eight parameters were considered viz. oral health literacy (X1), oral hygiene (X2) on brushing habit, SLT type (X3), SLT frequency (X4), smoking type (X5), smoking frequency (X6), other chewing habit type (X7) and other chewing habit frequency (X8). The rulebase was prepared using physician's intuition and information extracted from the literature review [11–25]. The rulebase appended in the Appendices A–F was used for assigning fuzzy membership function through rank ordering in male and female patients separately assuming that outlook toward oral health is different for each gender. Each gender was then again separated into eight age groups in years i.e., 18–23, 24–29, 30–35, 36–41, 42–47, 48–53, 54–59 and 60 above for age associated trend analysis. The fuzzy scale of output variable for other oral diseases (OTH) assigned was (0, 0, and 0.33), OSF was (0, 0.33, 0.67), OLK was (0.33, 0.67, 1) and for OSCC it was (0.67, 1, 1) according to increasing susceptibility to malignant potentiality [37]. Input variables were quantitatively described through intuition with their assigned membership functions in a fuzzy scale of 0–1 in Table 1.

2.3.3. Summarization of data

The data were summarized in terms of frequencies with respect to each age group and gender. During the process, initially the frequency of patients was sorted out according to each rule. Then the relative frequency of the type was multiplied by corresponding fuzzy scale. Each fuzzy number, thus obtained was then added for each group to get the final assigned fuzzy number which summarizes the condition of the group in a quantitative manner.

For the patients with oral precancer and cancer, the data were summarized in a slightly different way. 41 patients diagnosed with OSF, OLK and OSCC were considered to prepare the fuzzy disease expert system. There the age group was divided into five classes in years (i.e. 20–29, 30–39, 40–49, 50–59 and 60 years and above). The rulebase used for age group 18–23 of the whole population (WH) was used for group 20–29 (DS) with the disease. Similarly, rules for 30–35 (WH) were used for 30–39 (DS), 42–47 (WH) for 40–49 (DS), 54–59 (WH) for 50–59 (DS) while 60 and above remained same. The assigned fuzzy numbers were provided in Appendix H.

The resulting fuzzy numbers were then used for further decision making process. As linguistic interpretation of the mid value of the generated fuzzy set was found to be ambiguous in few cases, similarity measures through the Jaccard index have been introduced in this study. The linguistic outputs, which were found to be changed after considering similarity measures, were shown in bold font in Tables 4a, 4b and 5 for males, females and diseased patients respectively and from which *If-Then* rules to predict disease susceptibility in certain conditions were finally obtained. Fig. 1 depicts the workflow of the proposed methodology for application of fuzzy consensus for oral precancer and cancer susceptibility assessment.

3. Result and discussion

3.1. Descriptive statistics of the whole population

In assessing the addictions reported to have association with oral carcinogenesis of respondents, detailed observations were documented in Table 2. Variables having statistical significance with the disease outcome are depicted in Table 3 and were further considered for data summarization.

In the whole population, there is substantial evidence of a relationship between daily smoking tobacco (Pearson chi square 36.440, Likelihood Ratio 28.936), use of SLT (Pearson chi square 53.431, Likelihood Ratio 51.147), betel nut chewing (Pearson chi square 9.469, Likelihood Ratio 8.617) as well as betel leaves (Pearson chi square 8.249, Likelihood Ratio 7.376) consumption and the presence of oral lesions. Low literacy rates (49.68% people were with education up to 8th standard of Indian education system) in conjunction with debilitating addictive habits were also found to be associated with oral precancers and OSCC occurrence in this area. Daily brushing habit and higher education showed a protective effect on oral lesion occurrence. However, daily alcohol intake did not show any statistical correspondence with disease prevalence in this population. Interestingly, oral precancers and cancer were found to be more prevalent in people of the rural area of Darjeeling district, in the Bengali Hindu community in this study.

3.2. Interpretation of the summarized data

Most of the current literatures focus on institution based studies regarding incidence and prevalence of different oral lesions

Table 2 Descriptive statistics of the whole dataset.

Variables	Frequency	Percentage
Gender		
Female	426	45.4
Male	512	54.6
Education		
Illiterate	158	16.8
Primary	115	12.3
Secondary	148	15.8
Higher Secondary	157	16.7
Graduate	150	16.0
Higher than Graduate	17	1.8
Brushing		
Yes	927	98.8
No	11	1.2
Smoking tobacco		
Yes	184	19.6
No	754	80.4
Smokeless tobacco		
Yes	268	28.6
No	670	71.4
Alcohol		
Yes	109	11.6
No	829	88.4
Betel nut		
Yes	259	27.6
No	679	72.4
Betel leaves		
Yes	224	23.9
No	714	76.1
Lesion present	52	5.5
No lesions	886	84.5

where risk estimation is performed for individual causative factors [38,39]. Therefore, this study endeavored to find a meaningful association among multiple clinico-epidemiological parameters concerned to oral health literacy, oral habits and hygiene with oral disease susceptibility in different age groups. In this approach, conventional risk estimation tool was used for primary feature extraction only. The methodical fuzzy mapping of features and linguistic conversion of assigned membership functions, using Jaccard index presented in tables (Tables 4a, 4b and 5), was adopted for understanding such associations by trend analysis of the dataset. Here disease susceptibility was assessed using constructed *If-Then* rules in particular conditions. The rules were also implied for predicting oral complications other than oral precancers and cancers (OTH) too.

In the context of appreciating feature summarization against the disease conditions, Tables 4a, 4b and 5 were obtained through the linguistic conversion of fuzzy numbers using the Jaccard index for male, female and diseased patients respectively. All eight features (viz. X1–X8) were collated. The rules are constructed in such a manner that one of that may be explained in the following way – for males, in the age group of 18–23 years, *If* the oral health and habit were interpreted as satisfactory oral health literacy and very good oral hygiene

Table 3 Significant variables obtained from the chi-square test.

Variable	Type of variable (range)	Meaning of variable	Result of chi square test (chi square, degree of freedom, *p* value)	OR	95% confidence interval
Age	Dichotomous (0/1)	Age more/less than 40	30.899, 1, 0.000	5.430	2.809, 10.498
Education	Dichotomous (0/1)	Education more than/less than standard 8th	8.417, 1, 0.004	0.419	0.229, 0.776
Tobacco smoking	Dichotomous (0/1)	Yes/No	36.440, 1, 0.000	5.015	2.834, 8.873
Tobacco chewing	Dichotomous (0/1)	Yes/No	53.431, 1, 0.000	7.742	4.119, 14.549
Areca nut chewing	Dichotomous (0/1)	Yes/No	9.469, 1, 0.002	2.374	1.349. 4.17
Betel leaves chewing	Dichotomous (0/1)	Yes/No	8.249, 1, 0.004	2.279	1.282, 4.053
Brushing	Dichotomous (0/1)	Yes/No	20.191, 1, 0.000	0.096	0.027, 0.338

Table 4a Oral health and habit trend obtained through linguistic conversion of fuzzy numbers using Jaccard index in male.

Age group	Oral health literacy	Oral hygiene	SLT types consumption	SLT consumption frequency	Smoking tobacco type	Smoking tobacco consumption frequency	Other chewing material type	Other chewing material frequency	Disease chances
18–23	Satisfactory	Very Good	Very Low	Very Low	Low	Very Low	Low	Very Low	OTH
24–29	Satisfactory	Good	**Very Low**	Very Low	Low	Very Low	Low	Very Low	OTH
30–35	Poor	Good	Low	Very Low	Low	Very Low	Low	Very Low	OTH
36–41	Poor	Good	Low	**Low**	Medium	Moderate	Low	**Low**	OTH
42–47	Bad	Good	**Low**	Moderate	**Low**	**Low**	Medium	Low	OTH
48–53	Bad	**Good**	Moderate	Moderate	Medium	Moderate	**Low**	Low	OTH
54–59	Bad	Good	**Moderate**	Moderate	**Low**	**Low**	Medium	Low	OTH
60 Above	Bad	Good	Moderate	Moderate	Medium	Moderate	**Low**	**Low**	**OSF**

Table 4b Oral health and habit trend obtained through linguistic conversion of fuzzy numbers using Jaccard index in female.

Age group	Oral health literacy	Oral hygiene	SLT types consumption	SLT consumption frequency	Smoking tobacco type	Smoking tobacco consumption frequency	Other chewing material type	Other chewing material frequency	Disease chances
18–23	**Good**	Very Good	Very Low	Very Low	Very Low	Very Low	Low	Very Low	OTH
24–29	Satisfactory	Good	Very Low	Very Low	Very Low	Very Low	Low	Very Low	OTH
30–35	Bad	Good	Very Low	Very Low	Very Low	Very Low	Low	Very Low	OTH
36–41	Bad	Good	Very Low	Very Low	Very Low	Very Low	Low	**Low**	OTH
42–47	Bad	Good	**Very Low**	Low	Very Low	Very Low	Medium	Low	OTH
48–53	Bad	Good	Low	Low	Very Low	Very Low	Low	Very Low	OTH
54–59	**Poor**	**Good**	Low	Low	Very Low	Very Low	Medium	Low	OTH
60 above	Bad	Poor	Low	Low	Very Low	Very Low	Medium	Medium	OTH

Table 5 Oral health and habit trend analysis obtained through linguistic conversion of fuzzy numbers using the Jaccard index in OSF, OLK and OSCC patients with probable disease susceptibility.

Age group	Oral health literacy	Oral hygiene	SLT types consumption	SLT consumption frequency	Smoking tobacco type	Smoking tobacco consumption frequency	Other chewing material type	Other chewing material frequency	Disease chances
20–29 (M & F)	Satisfactory	Very Good	**Low**	**Moderate**	**Medium**	Low	Low	Very Low	OSF
30–39 (M & F)	Bad	**Satisfactory**	Medium	Moderate	Low	Very Low	Medium	Low	**OLK**
40–49 (M & F)	Bad	Good	High	High	Medium	Moderate	Medium	**Moderate**	OLK
50–59 (M & F)	Bad	Satisfactory	Medium	Moderate	Medium	Moderate	Medium	Low	OLK
60 Above (M & F)	Bad	Satisfactory	High	High	High	High	Medium	Moderate	OSCC
60 above (Male only)	Bad	Poor	Low	Low	Very Low	Very Low	Medium	Medium	**OSF**

and very low SLT type consumption and very low SLT consumption frequency and low smoking tobacco type and very low smoking tobacco consumption frequency and low other chewing material type and very low other chewing material frequency *Then* they were susceptible to oral complications other than precancers and cancers (OTH). Interpretation of each row to be read in the Tables 4a, 4b and 5 similarly portrayed the oral health and habit scenario of the population in a compartmentalized manner, along with the disease prediction rules for males, females and diseased patients for different age groups separately.

When the trend between male and female population of different age groups (Tables 4a and 4b) was compared in case of oral health literacy, the outlook was found to be considerably varied. It also demonstrated age based feature classification in a more precise manner. Only female of 18–23, 30–41 and 54–59 years depicted better oral health literacy than males, and the concept of more education is connoted in males than female in Indian population [40]. In this population, SLT type and consumption frequency were comparatively lower in females than males which also support the previous findings [41]. In females, the habit of smoking tobacco consumption in terms of both the types and frequency was found to be very low, while in males, there was an increase from age 42 years (Table 4a). The Areca nut and leaves chewing habit type and frequency when considered, the addiction was found to be more in females [6], which was further elaborated in Table 4a depicting most of the linguistic outcomes of X6 and X7 against males were almost low, whereas the trend in females was mostly medium. Furthermore it elucidated the deterioration of females' oral hygiene after the age of 60 years (Table 4b).

Though OSF is associated with areca nut intake in dose dependent manner [17], this study highlighted the necessity of fuzzy approach in assessing its critical association of other factors too. In this regard present evaluation unveiled, that a patient within the age group of 20–29 years, even if areca leaves and nut associated chewing material consumption type and their frequency were low and very low respectively, becomes susceptible to OSF when other addictive habits (i.e. X3–X6) were present. This finding supports the concept of addictive interaction model in oral precancer susceptibility. Again smoking of tobacco is known to be associated with OLK [10]. Present findings further provided new information to aid the *If-Then* fuzzy rule for onset of such pre-cancer from 40 years where from 36 years smoking type and their frequency were increased in males (Tables 4a and 5). Increase in deleterious oral habits such as SLT consumption frequency, was also found in males from age group of 36 years (Table 5). In case of OSCC prevalence, the present study demonstrated an association of high smoking tobacco with SLT consumption, poor oral hygiene, bad oral health literacy and high age (Table 5) and can be corroborated with the findings of the other studies [13,23].

Hence, the proposed oral pre-cancer, cancer and other oral diseases susceptibility assessment methodology with embedded fuzzy analytical dimensions depicted the association of multiple clinico-epidemiological parameters (viz. oral health and literacy as well as addictive oral habits) in simple linguistic terms which not only were useful for clinical users but also carried translational values.

4. Conclusion

Fuzzy rule-base approach has been utilized for value addition to the findings from conventional statistical approach in defining particular association between significant clinicoepidemiological parameters and their plausible impact on disease output in a particular dataset. Low literacy rates in conjunction with debilitating addictive habits were found to be important underlying reasons for oral precancers and OSCC occurrence in the studied population. Further, oral health and habits' trend analysis through fuzzy *If-Then* rule demonstrated gender based differences in the awareness outlook in different age groups. Chances of disease susceptibility in certain condition can also be predicted by the proposed methodology. The novelty of the proposed approach relies upon consideration of uncertainty of conditions associated with disease occurrence and incorporation of physician's intuition in real-life situations, in contrast to conventional statistical method which predicts disease chances in rigid quantitative values. The new dimension of questionnaire data handling involved population of specific demography, and same methodology can be implemented for other demographic conditions as well. It is also first of its kind and can help clinicians and policy makers in adopting interventions and habit preventing strategies.

Conflict of interest

The authors declare no conflict of interest.

Appendix A. Rulebase for oral health literacy

Age group 18–23	If Schooling is Il	Then Education is Bad
Age group 18–23	If Schooling is M or Pr	Then Education is Poor
Age group 18–23	If Schooling is S	Then Education is Satisfactory
Age group 18–23	If Schooling is HS	Then Education is Good
Age group 18–23	If Schooling is G or HG	Then Education is Very Good
Age group 24–29	If Schooling is Il	Then Education is Bad
Age group 24–29	If Schooling is M or Pr	Then Education is Poor
Age group 24–29	If Schooling is S or HS	Then Education is Satisfactory
Age group 24–29	If Schooling is G	Then Education is Good
Age group 24–29	If Schooling is HG	Then Education is Very Good
Age group 30–35	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 30–35	If Schooling is HS	Then Education is Poor
Age group 30–35	If Schooling is G	Then Education is Satisfactory
Age group 30–35	If Schooling is HG	Then Education is Good
Age group 36–41	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 36–41	If Schooling is HS	Then Education is Poor
Age group 36–41	If Schooling is G or HG	Then Education is Satisfactory
Age group 42–47	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 42–47	If Schooling is HS or G	Then Education is Poor
Age group 42–47	If Schooling is HG	Then Education is Satisfactory
Age group 48–53	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 48–53	If Schooling is HS	Then Education is Poor
Age group 48–53	If Schooling is G or HG	Then Education is Satisfactory
Age group 54–59	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 54–59	If Schooling is HS	Then Education is Poor
Age group 54–59	If Schooling is G or HG	Then Education is Satisfactory
Age group 60 above	If Schooling is Il, Pr, M, S	Then Education is Bad
Age group 60 above	If Schooling is HS	Then Education is Poor
Age group 60 above	If Schooling is G or HG	Then Education is Satisfactory

Il = Illiterate, Pr = Primary, Med = Upto 8th Standard, S = Secondary, HS = Higher Secondary, G = Graduate, HG = Higher than Graduate.

Appendix B. Rulebase for oral hygiene

Age group 18–23	If Brushing with Pa	Then Oral Hygiene Very Good
Age group 18–23	If Brushing with Po	Then Oral Hygiene Good
Age group 18–23	If Brushing with KM, LM	Then Oral Hygiene Medium
Age group 18–23	If Brushing with Dantan	Then Oral Hygiene Poor
Age group 18–23	If Brushing with Other	Then Oral Hygiene Very Poor
Age group 24–29	If Brushing with Pa	Then Oral Hygiene Good
Age group 24–29	If Brushing with Po, KM, LM	Then Oral Hygiene Medium
Age group 24–29	If Brushing with Dantan	Then Oral Hygiene Poor
Age group 24–29	If Brushing with Other	Then Oral Hygiene Very Poor
Age group 30–35	If Brushing with Pa	Then Oral Hygiene Good
Age group 30–35	If Brushing with Po	Then Oral Hygiene Medium
Age group 30–35	If Brushing with KM, LM, Dantan	Then Oral Hygiene Poor
Age group 36–41	If Brushing with Others	Then Oral Hygiene Very Poor
Age group 36 to 41	If Brushing with Pa	Then Oral Hygiene Good
Age group 36 to 41	If Brushing with Po, KM, LM	Then Oral Hygiene Medium
Age group 36 to 41	If Brushing with Dantan	Then Oral Hygiene Poor

Age group 36 to 41	If Brushing with Others	Then Oral Hygiene Very Poor
Age group 42 to 47	If Brushing with Pa	Then Oral Hygiene Good
Age group 42 to 47	If Brushing with Po, KM, LM, Dantan	Then Oral Hygiene Poor
Age group 42 to 47	If Brushing with Others	Then Oral Hygiene Very Poor
Age group 48 to 53	If Brushing with Pa	Then Oral Hygiene Good
Age group 48 to 53	If Brushing with Po, KM, LM, Dantan, Others	Then Oral Hygiene Very Poor
Age group 54 to 59–	If Brushing with Pa	Then Oral Hygiene Good
Age group 54–59	If Brushing with Po, KM, LM, Dantan, Others	Then Oral Hygiene Very Poor
Age group 60–more	If Brushing with Pa	Then Oral Hygiene Good
Age group 60–more	If Brushing with Po, KM, LM, Dantan, Others	Then Oral Hygiene Very Poor

Pa = Toothpaste, Po = Toothpowder, KM = Kalamanjan, LM = Lalmanjan, Dantan = Tree stems, Other = Sand, oil, salt, etc.

Appendix C. Rulebase for SLT type consumption

Age group 18–23	If smokeless Type of 4 agent types or more like Khaini, Gutkha, Zarda, Gundi, Nassi, Guraku etc Tobacco leave consumption	Then susceptibility very high
Age group 18–23	If smokeless any 3	Then susceptibility high
Age group 18–23	If smokeless any 2	Then susceptibility moderate
Age group 18–23	If smokeless any 1	Then susceptibility low
Age group 18–23	If smokeless occasional	Then susceptibility very low
Age group 24–29	If smokeless of 4 types or more	Then susceptibility very high
Age group 24–29	If smokeless any 3, 2	Then susceptibility high
Age group 24–29	If smokeless 1	Then susceptibility moderate
Age group 24–29	If smokeless occasional	Then susceptibility low
Age group 30–35	If smokeless all 4	Then susceptibility very high
Age group 30–35	If smokeless any 3, 2	Then susceptibility high
Age group 30–35	If smokeless 1	Then susceptibility moderate
Age group 30–35	If smokeless occasional	Then susceptibility low
Age group 36–41	If smokeless all 4, 3	Then susceptibility very high
Age group 36–41	If smokeless 2	Then susceptibility high
Age group 36–41	If smokeless 1	Then susceptibility moderate
Age group 36–41	If smokeless occasional	Then susceptibility low
Age group 42–47	If smokeless 4, 3, 2	Then susceptibility very high
Age group 42–47	If smokeless 1	Then susceptibility high
Age group 42–47	If smokeless occasional	Then susceptibility moderate
Age group 48–53	If smokeless 4, 3, 2	Then susceptibility very high
Age group 48–53	If smokeless 1	Then susceptibility high
Age group 48–53	If smokeless occasional	Then susceptibility moderate
Age group 54–59	If smokeless 4, 3, 2	Then susceptibility very high
Age group 54–59	If smokeless 1	Then susceptibility high
Age group 54–59	If smokeless occasional	Then susceptibility moderate
Age group 60 and above	If smokeless 4, 3, 2	Then susceptibility very high
Age group 60 and above	If smokeless 1	Then susceptibility high
Age group 60 and above	If smokeless occasional	Then susceptibility moderate

Appendix D. Rulebase for smoking tobacco type consumption

Age group 18–23	If smoking cig	Then low
Age group 18–23	If smoking Be	Then Medium
Age group 18–23	If smoking Both	Then High
Age group 24–29	If smoking cig	Then low
Age group 24–29	If smoking Be	Then Medium
Age group 24–29	If smoking Both	Then High
Age group 30–35	If smoking Be, Cig	Then Medium
Age group 30–35	If smoking Both	Then High
Age group 36–41	If smoking Be, Cig	Then Medium

(*continued on next page*)

Age group 36–41	If smoking Both	Then High
Age group 42–47	If smoking Be, Cig	Then Medium
Age group 42–47	If smoking Both	Then High
Age group 48–53	If smoking Be, Cig	Then Medium
Age group 48–53	If smoking Both	Then High
Age group 54–59	If smoking Cig	Then Medium
Age group 54–59	If smoking Be, Both	Then High
Age group 60 above	If smoking Cig	Then Medium
Age group 60 above	If smoking Be, Both	Then High

Be = Beedi, Cig = Cigarette.

Appendix E. Rulebase for other chewing product consumption

Age group 18–22	If Chewing BL	Then low
Age group 18–23	If Chewing AN	Then Medium
Age group 18–23	If Chewing Both	Then High
Age group 24–29	If Chewing BL	Then low
Age group 24–29	If Chewing AN	Then Medium
Age group 24–29	If Chewing Both	Then High
Age group 30–35	If Chewing AN, BL	Then Medium
Age group 30–35	If Chewing Both	Then High
Age group 36–41	If Chewing AN, BL	Then Medium
Age group 36–41	If Chewing Both	Then High
Age group 42–47	If Chewing AN, BL	Then Medium
Age group 42–47	If Chewing Both	Then High
Age group 48–53	If Chewing AN, BL	Then Medium
Age group 48–53	If Chewing Both	Then High
Age group 54–59	If Chewing BL	Then Medium
Age group 54–59	If Chewing AN, Both	Then High
Age group 60 above	If Chewing BL	Then Medium
Age group 60 above	If Chewing AN, Both	Then High

AN = Areca Nut, BL = Betel Leaves.

Appendix F. Rulebase for intake frequency of SLT, smoking tobacco and other chewing products

Age group 18–23	If frequency Occassional (occ)	The intake is Very low
Age group 18–23	If frequency less than 2	The intake is low
Age group 18–23	If frequency less than 5	The intake is Medium
Age group 18–23	If frequency less than 10	The intake is high
Age group 18–23	If frequency more than 10	The intake is very high
Age group 24–29	If frequency occ	The intake is low
Age group 24–29	If frequency less than 2 or 5	The intake is Medium
Age group 24–29	If frequency less than 10	The intake is high
Age group 24–29	If frequency more than 10	The intake is very high
Age group 30–35	If frequency occ, less than 2	The intake is low
Age group 30–35	If frequency less than 5	The intake is Medium
Age group 30–35	If frequency less than 10	The intake is high
Age group 30–35	If frequency more than 10	The intake is very high
Age group 30–33	If frequency occ, less than 2	The intake is Medium
Age group 30–33	If frequency less than 5 or 10	The intake is high
Age group 30–33	If frequency more than 10	The intake is very high
Age group 36–41	If frequency occ, less than 2	The intake is Medium
Age group 36–41	If frequency less than 5	The intake is high
Age group 36–41	If frequency less than 10 or more than 10	The intake is very high
Age group 42–47	If frequency occ, less than 2, 5, 10	The intake is high
Age group 42–47	If frequency more than 10	The intake is very high
Age group 48–53	If frequency occ less than 2, 5, 10	The intake is high
Age group 48–53	If frequency more than 10	The intake is very high
Age group 54–59	If frequency occ, less than 2, 5, 10	The intake is high

Age group 54–59	If frequency more than 10	The intake is very high
Age group 60 and above	If frequency occ less than 2, 5, 10	The intake is high
Age group 60 and above	If frequency more than 10	The intake is very high

Appendix G. Assigned membership functions for the fuzzy sets of the input and output

Age group	Oral health literacy	SLT types consumption	SLT consumption frequency	Smoking tobacco type	Smoking tobacco consumption frequency	Other chewing material type	Other chewing material frequency	Oral hygiene	Y (disease chance)
18–23 (M)	0.3, 0.518, 0.73	0.006, 0.05, 0.097	0.037, 0.081, 0.147	0.013, 0.081, 0.147	0.031, 0.065, 0.116	0.088, 0.206, 0.238	0.022, 0.072, 0.131	0.738, 0.988, 1	0.004, 0.016, 0.349
24–29 (M)	0.304, 0.533, 0.764	0.056, 0.110, 0.164	0.061, 0.117, 0.168	0.009, 0.028, 0.131	0.056, 0.112, 0.205	0.103, 0.224, 0.252	0.040, 0.100, 0.159	0.493, 0.743, 0.995	0.006, 0.009, 0.340
30–35 (M)	0.081, 0.190, 0.440	0.089, 0.177, 0.262	0.097, 0.185, 0.259	0.016, 0.113, 0.194	0.060, 0.073, 0.153	0.113, 0.234, 0.242	0.044, 0.105, 0.161	0.484, 0.734, 0984	0.011, 0.027, 0.354
36–41 (M)	0.055, 0.161, 0.411	0.136, 0.25, 0.356	0.225, 0.331, 0.394	0.060, 0.305, 0.492	0.301, 0.441, 0.475	0.093, 0.237, 0.288	0.102, 0.174, 0.288	0.483, 0.725, 0.975	0, 0.005, 0.339
42–47 (M)	0.008, 0.084, 0.360	0.242, 0.358, 0.454	0.258, 0.373, 0.462	0.023, 0.2, 0.354	0.192, 0.281, 0.354	0.170, 0.362, 0.385	0.204, 0.3, 0.384	0.462, 0.731, 0.962	0.015, 0.026, 0.348
48–53 (M)	0.034, 0.097, 0.347	0.278, 0.420, 0.557	0.313, 0.455, 0.568	0.057, 0.318, 0.523	0.290, 0.420, 0.523	0.136, 0.318, 0.364	0.193, 0.284, 0.364	0.420, 0.631, 0.881	0.038, 0.068, 0.386
54–59 (M)	0.043, 0.091, 0.341	0.256, 0.378, 0.469	0.262, 0.384, 0.488	0.085, 0.293, 0.415	0.232, 0.335, 0.415	0.195, 0.402, 0.414	0.207, 0.311, 0.415	0.463, 0.695, 0.945	0.041, 0.073, 0.394
60 above (M)	0.019, 0.055, 0.305	0.310, 0.454, 0.565	0.319, 0.463, 0.574	0.231, 0.518, 0.574	0.287, 0.430, 0.574	0.148, 0.305, 0.315	0.162, 0.240, 0.315	0.481, 0.722, 0.972	0.123, 0.191, 0.468
18–23 (F)	0.383, 0.657, 0.901	0, 0.003, 0.009	0, 0.003, 0.009	0, 0, 0	0, 0, 0	0.043, 0.136, 0.191	0.018, 0.049, 0.099	0.741, 0.987, 0.991	0, 0, 0.333
24–29 (F)	0.223, 0.443, 0.696	0.024, 0.048, 0.072	0.036, 0.066, 0.090	0, 0, 0	0, 0, 0	0.060, 0.157, 0.205	0.033, 0.087, 0.139	0.463, 0.696, 0.946	0, 0, 0.333
30–35 (F)	0.016, 0.053, 0.303	0.025, 0.053, 0.081	0.03, 0.06, 0.200	0, 0, 0	0, 0, 0	0.075, 0.188, 0.256	0.038, 0.109, 0.178	0.444, 0.675, 0.925	0.004, 0.008, 0.341
36–41 (F)	0.019, 0.051, 0.301	0.051, 0.106, 0.162	0.102, 0.157, 0.204	0, 0, 0	0, 0, 0	0.093, 0.231, 0.278	0.111, 0.181, 0.241	0.50, 0.75, 1	0.006, 0.012, 0.345
42–47 (F)	0.026, 0.057, 0.302	0.188, 0.281, 0.365	0.193, 0.286, 0.375	0, 0.021, 0.042	0.021, 0.031, 0.042	0.156, 0.365, 0.417	0.219, 0.323, 0.417	0.427, 0.666, 0917	0.014, 0.028, 0.361
48–53 (F)	0.006, 0.020, 0.270	0.090, 0.167, 0.243	0.153, 0.230, 0.306	0, 0.070, 0.138	0.069, 0.104, 0.138	0.083, 0.181, 0.194	0.097, 0.146, 0.194	0.458, 0.688, 0.938	0.009, 0.019, 0.361
54–59 (F)	0, 0.026, 0.276	0.145, 0.211, 0.263	0.132, 0.197, 0.263	0, 0, 0	0, 0, 0	0.184, 0.395, 0.421	0.210, 0.316, 0.421	0.421, 0.632, 0.881	0.018, 0.053, 0.368
60 above (F)	0.01, 0.03, 0.28	0.23, 0.34, 0.43	0.23, 0.33, 0.40	0.03, 0.04, 0.04	0.02, 0.03, 0.04	0.2, 0.42, 0.44	0.26, 0.38, 0.48	0.019, 0.38, 0.63	0.053, 0.08, 0.386
20–29	0.25, 0.50, 0.75	0.063, 0.188, 0.313	0.25, 0.375, 0.438	0.125, 0.375, 0.625	0.188, 0.313, 0.5	0, 0, 0	0, 0, 0	0.75, 1, 1	0.25, 0.585, 0.667
30–39	0, 0.05, 0.3	0.25, 0.5, 0.75	0.25, 0.5, 0.7	0.1, 0.2, 0.2	0.1, 0.15, 0.2	0.2, 0.3, 0.6	0, 0.25, 0.5	0.4, 0.6, 0.85	0.267, 0.600, 0.867
40–49	0, 0, 0.25	0.438, 0.688, 0.938	0.562, 0.813, 1	0.188, 0.50, 0.625	0.344, 0.50, 0.625	0.125, 0.375, 0.5	0.25, 0.375, 0.5	0.375, 0.594, 0.844	0.459, 0.792, 0.917
50–59	0.045, 0.091, 0.341	0.318, 0.454, 0.545	0.272, 0.409, 0.545	0.227, 0.545, 0.636	0.364, 0.523, 0.636	0.227, 0.454, 0.454	0.227, 0.341, 0.455	0.363, 0.545, 0.795	0.485, 0.818, 0.939
60 above	0, 0, 0.25	0.481, 0.692, 0.846	0.462, 0.673, 0.846	0.385, 0.808, 0.846	0.5, 0.711, 0.846	0.307, 0.615, 0.615	0.442, 0.596, 0.615	0.308, 0.462, 0.808	0.616, 0.949, 1

Appendix H. ***Best matched similarity measure for linguistic output selection*:**

Age group	Education	smokeless type	smokeless frequency	Smoking type	Smoking frequency	Chewing type	Chewing frequency	Oral hygiene	Y (disease chance)
18–23 (M)	0.955	0.972	0.976	0.905	0.973	0.888	0.976	0.988	0.974
24–29 (M)	0.986	0.967	0.965	0.904	0.972	0.881	0.972	0.984	0.974
30–35 (M)	0.979	0.945	0.945	0.918	0.975	0.869	0.97	0.985	0.974
36–41 (M)	0.973	0.971	0.945	0.712	0.947	0.886	0.951	0.985	0.974
42–47 (M)	0.966	0.94	0.932	0.927	0.956	0.821	0.956	0.983	0.975
48–53 (M)	0.966	0.969	0.976	0.874	0.96	0.841	0.957	0.965	0.968
54–59 (M)	0.968	0.936	0.943	0.874	0.947	0.836	0.958	0.982	0.966
60 above (M)	0.98	0.974	0.976	0.896	0.973	0.837	0.955	0.984	0.936
18–23 (F)	0.966	0.942	0.942	0.83	0.937	0.907	0.972	0.988	0.973
24–29 (F)	0.981	0.962	0.964	0.832	0.937	0.903	0.973	0.982	0.973
30–35 (F)	0.981	0.965	0.988	0.832	0.937	0.906	0.972	0.977	0.974
36–41 (F)	0.981	0.969	0.944	0.832	0.937	0.887	0.935	0.985	0.974
42–47 (F)	0.98	0.959	0.959	0.857	0.953	0.838	0.953	0.974	0.973
48–53 (F)	0.986	0.944	0.951	0.903	0.962	0.884	0.949	0.98	0.971
54–59 (F)	0.952	0.94	0.945	0.832	0.937	0.84	0.957	0.949	0.974
60 above (F)	0.985	0.947	0.928	0.851	0.952	0.849	0.94	0.943	0.966
20–29 (Disease)	0.985	0.966	0.939	0.914	0.966	0.832	0.937	0.969	0.928
30–39	0.983	0.985	0.987	0.874	0.947	0.879	0.985	0.969	0.965
40–49	0.985	0.978	0.984	0.919	0.984	0.869	0.944	0.966	0.959
50–59	0.967	0.969	0.964	0.913	0.982	0.848	0.95	0.98	0.951
60 above	0.985	0.979	0.975	0.926	0.981	0.881	0.953	0.974	0.978

M = Male, F = Female.

Acknowledgments

This study is partially supported by MHRD, Government of India, New Delhi (Institute Approval Number. IIT/SRIC/S MST/IAN/2013–14/222).

References

[1] Coelho KR. Challenges of the oral cancer burden in India. J Cancer Epidemiol 2012:17.

[2] Sankaranarayanan R. Oral cancer in India: an epidemiologic and clinical review. Oral Surg Oral Med Oral Pathol 1990;69(3): 325–30.

[3] Messadi DV, Wilder-Smith P, Wolinsky L. Improving oral cancer survival: the role of dental providers. J California Dental Assoc 2009;37(11):789–98.

[4] Reichart PA. Identification of risk groups for oral precancer and cancer and preventive measures. Clin Oral Invest 2001;5(4):207–13.

[5] Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol 2009;45(4):309–16.

[6] Guha N, Warnakulasuriya S, Vlaanderen J, Straif K. Betel quid chewing and the risk of oral and oropharyngeal cancers: a meta-analysis with implications for cancer control. Int J Cancer 2014;135(6):1433–43.

[7] Warnakulasuriya S, Johnson N, Van der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. J Oral Pathol Med 2007;36(10):575–80.

[8] Kurokawa H, Matsumoto S, Murata T, et al. Immunohistochemical study of syndecan-1 down-regulation and the expression of p53 protein or Ki-67 antigen in oral leukoplakia with or without epithelial dysplasia. J Oral Pathol Med 2003;32(9):513–21.

[9] Moutasim KA, Jenei V, Sapienza K, et al. Betel-derived alkaloid up-regulates keratinocyte alpha-beta6 integrin expression and promotes oral submucous fibrosis. J Pathol 2011;223(3):366–77.

[10] Bokor-Bratić M, Vučković N. Cigarette smoking as a risk factor associated with oral leukoplakia. Arch Oncol 2002;10(2):67–70.

[11] Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless tobacco and cancer. Lancet Oncol 2008;9(7):667–75.

[12] Carr AB, Ebbert JO. Interventions for tobacco cessation in the dental setting. A systematic review. Community Dent Health 2007;24(2):70–4.

[13] Rodu B, Cole P. Smokeless tobacco use and cancer of the upper respiratory tract. Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol 2002;93(5):511–5.

[14] Jayalekshmi PA, Gangadharan P, Akiba S, Koriyama C, Nair RR. Oral cavity cancer risk in relation to tobacco chewing and bidi smoking among men in Karunagappally, Kerala, India: Karunagappally cohort study. Cancer Sci 2011;102(2):460–7.

[15] Rahman M, Sakamoto J, Fukui T. Bidi smoking and oral cancer: a meta-analysis. Int J Cancer 2003;106:600–4.

[16] Allam E, Zhang W, Al-Shibani N, et al. Effects of cigarette smoke condensate on oral squamous cell carcinoma cells. Arch Oral Biol 2011;56(10):1154–61.

[17] Yang YH, Ho PS, Lu HM, Huang IY, Chen CH. Comparing dose–response measurements of oral habits on oral leukoplakia and oral submucous fibrosis from a community screening program. J Oral Pathol Med 2010;39(4):306–12.

[18] Jacob BJ, Straif K, Thomas G, et al. Betel quid without tobacco as a risk factor for oral precancers. Oral Oncol 2004;40(7): 697–704.

[19] Muwonge R, Ramadas K, Sankila R, et al. Role of tobacco smoking, chewing and alcohol drinking in the risk of oral cancer in Trivandrum, India: a nested case-control design using incident cancer cases. Oral Oncol 2008;44(5):446–54.

[20] Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. Bull World Health Organ 2005;83(9):661–9.
[21] Paul B, Basu M, Dutta S, Chattopadhyay S, Sinha D, Misra R. Awareness and practices of oral hygiene and its relation to sociodemographic factors among patients attending the general outpatient department in a tertiary care Hospital of Kolkata, India. J Family Med Prim Care 2014;3(2):107–11.
[22] Chandra Shekar B, Reddy CVK, Manjunath B, Suma S. Dental health awareness, attitude, oral health-related habits, and behaviors in relation to socio-economic factors among the municipal employees of Mysore city. Ann Trop Med Public Health 2011;4(2):99–106.
[23] Guha N, Boffetta P, Wunsch Filho V, et al. Oral health and risk of squamous cell carcinoma of the head and neck and esophagus: results of two multicentric case-control studies. Am J Epidemiol 2007;166(10):1159–73.
[24] Paulander J, Axelsson P, Lindhe J. Association between level of education and oral health status in 35-, 50-, 65- and 75-year-olds. J Clin Periodontol 2003;30(8):697–704.
[25] Guo Y, Logan HL, Dodd VJ, Muller KE, Marks 3rd JG, Riley JL. Health literacy: a pathway to better oral health. Am J Public Health 2014;104:e85–91.
[26] Singh H, Gupta MM, Meitzler T, et al. Real-life applications of fuzzy logic. Adv Fuzzy Syst 2013.
[27] Dubois D, Prade HM. Fuzzy sets and systems: theory and applications. Academic Press; 1980.
[28] Chakraborty D. Structural quantization of vagueness in linguistic expert opinions in an evaluation programme. Fuzzy Sets Syst 2001;119(1):171–86.
[29] Spiegelman D. Approaches to uncertainty in exposure assessment in environmental epidemiology. Annu Rev Public Health 2010;31:149.
[30] Viswanath K, Breen N, Meissner H, et al. Cancer knowledge and disparities in the information age. J Health Commun 2006;11(S1): 1–17.
[31] Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. Proc Natl Acad Sci 2006;103(42):15693–7.
[32] Hwang CM, Yang MS. New similarity measures between generalized trapezoidal fuzzy numbers using the Jaccard index. Int J Uncertainty Fuzziness Knowledge-Based Syst 2014;22(6):831–44.
[33] Abbod MF, von Keyserlingk DG, Linkens DA, Mahfouf M. Survey of utilisation of fuzzy technology in medicine and healthcare. Fuzzy Sets Syst 2001;120(2):331–49.
[34] Papageorgiou EI. A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. Appl Soft Comput 2011;11(1):500–13.
[35] Lacagnina V, Leto-Barone MS, La Piana S, La Porta G, Pingitore G, Di Lorenzo G. Comparison between statistical and fuzzy approaches for improving diagnostic decision making in patients with chronic nasal symptoms. Fuzzy Sets Syst 2014;237:136–50.
[36] Hong T-P, Lee C-Y. Induction of fuzzy rules and membership functions from training examples. Fuzzy Sets Syst 1996;84(1): 33–47.
[37] Bouquot JE, Speight PM, Farthing PM. Epithelial dysplasia of the oral mucosa—diagnostic problems and prognostic features. Curr Diagnostic Pathol 2006;12(1):11–21.
[38] Patil PB, Bathi R, Chaudhari S. Prevalence of oral mucosal lesions in dental patients with tobacco smoking, chewing, and mixed habits: a cross-sectional study in South India. J Family Community Med 2013;20(2):130.
[39] Sujatha D, Hebbar PB, Pai A. Prevalence and correlation of oral lesions among tobacco smokers, tobacco chewers, areca nut and alcohol users. Asian Pac J Cancer Prev 2012;13(4):1633–7.
[40] Kingdon GG. Does the labour market explain lower female schooling in India? J Develop Stud 1998;35(1):39–65.
[41] Rani M, Bonu S, Jha P, Nguyen S, Jamjoum L. Tobacco use in India: prevalence and predictors of smoking and chewing in a national cross sectional household survey. Tobacco Control 2003;12(4), e4 (1–8).