

[Egyptian Informatics Journal (2016) 17, 251–263](http://dx.doi.org/10.1016/j.eij.2015.09.005)

Cairo University

Egyptian Informatics Journal

[www.elsevier.com/locate/eij](http://www.elsevier.com/locate/eij) [www.sciencedirect.com](http://www.sciencedirect.com/science/journal/11108665)

FULL-LENGTH ARTICLE

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



Satarupa Banerjee [a](#_bookmark0),[\*](#_bookmark5), Aishwaryaprajna [b](#_bookmark1), Debjani Chakraborty [c](#_bookmark2), Amita Giri [d](#_bookmark3),

Ranjan Ghosh [e](#_bookmark4), Badal Chandra Sarkar [e](#_bookmark4), Jyotirmoy Chatterjee [a](#_bookmark0)

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

Abstract Health questionnaire data assessment conventionally relies upon statistical analysis in understanding disease susceptibility using discrete numbers and fails to reflect physician’s perspec- tives 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 mem- bership 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 con- sidering 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.

© 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information,

KEYWORDS

Fuzzy rulebase; Oral health; Oral habit;

Susceptibility assessment; Oral pre-cancer and cancer

Cairo University. This is an open access article under the CC BY-NC-ND license [(http://creativecommons.](http://creativecommons/)

org/licenses/by-nc-nd/4.0/).

\* Corresponding author. Tel.: +91 3222 281225; fax: +91 3222

282221.

E-mail addresses: [satarupa@smst.iitkgp.ernet.in](mailto:satarupa@smst.iitkgp.ernet.in), satarupabando@ gmail.com (S. Banerjee).

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

**Production and hosting by Elsevier**

1. Introduction

India experiences one of the highest incidence rates of oral can- cer globally [[1]](#_bookmark18). Oral cancer is the leading cancer type in men and the third most common cancer in women [[2]](#_bookmark20). In India, oral cancer is usually detected at advanced stages and the five year survival rate for advanced oral cancer is very low [[3]](#_bookmark22), posing an important public health challenge. Hence the early detection of

<http://dx.doi.org/10.1016/j.eij.2015.09.005>

1110-8665 © 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University. This is an open access article under the CC BY-NC-ND license [(http://creativecommons.org/licenses/by-nc-nd/4.0/).](http://creativecommons.org/licenses/by-nc-nd/4.0/))

252 S. Banerjee et al.

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]](#_bookmark24), where betel quid/areca nut is commonly chewed with smokeless tobacco (SLT) [[6]](#_bookmark26). 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 hav- ing excluded (other) known diseases or disorders that carry no increased risk of cancer [[7]](#_bookmark28). Annually, approximately 2–3% of oral lesions show malignant transformation. Long term follow-up studies clinically suggest that OLK with severe dys- plasia is more susceptible to transformation into OSCC [[8]](#_bookmark29). OSF is a chronic, premalignant condition, characterized by progressive submucosal fibrosis inside oral cavity [[9]](#_bookmark14). 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]](#_bookmark15), while for OSF, the major etiology of the disease considered was chewing areca nut [[9]](#_bookmark14).

* 1. *Literature review for rulebase construction*

Consumption of tobacco is the major cause of death and dis- ability 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 com- bustion” [[11]](#_bookmark16). 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]](#_bookmark16). 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]](#_bookmark17). SLT use in the United States has been associated with an increased risk for oral cancer in a dose–response fashion [[13]](#_bookmark19).

In Indian population role of Bidi and SLT is well known [[14]](#_bookmark21). When smoking tobacco is to be considered, in a meta analysis, Bidi smoking had higher odds ratio (OR) than cigar- ette smoking [[15]](#_bookmark23). So Bidi smoking was considered more harm- ful than cigarette smoking. Cigarette smoke condensate was found to enhance matrix metalloproteinases (MMPs), MMP- 2 and MMP-9 expression and thus increase collagen degrada- tion which ultimately increases chance of metastasis in cancer patients [[16]](#_bookmark25). 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 con- sumption for oropharynx and esophagus cancer. As OSCC

for OSF [[18]](#_bookmark30). Even chewing betel quid with or without tobacco itself is considered as an independent cancer causing factor [[19]](#_bookmark31).

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 under- lying 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]](#_bookmark32). In India, there are many things to do for upgrading oral health awareness [[21]](#_bookmark32). Again a direct relation was noted between the favorable dental health awareness, attitude, oral hygiene behavior, and socioeconomic status in the Indian population [[22]](#_bookmark32). Even poor oral health was considered as an independent causative condition of OSCC [[23]](#_bookmark32). The study sug- gested that educational level influences the oral conditions and should be considered in assessing risk, and in planning appro- priate preventive measures as health literacy also has impact on oral health [[24,25]](#_bookmark32). 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. *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 com- puting methods. It can also be considered as an approach of computing with words as linguistic language is always pre- ferred for expressing opinions. The beauty of real-life applica- tion of fuzzy logic lies in the precision in meaning of an outcome and getting an idea of a complex system with toler- ance of imprecision [[26]](#_bookmark32).

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 quan- tify the linguistic expressions here fuzzy numbers must have to be defined. Defining fuzzy numbers allows modeling of com- plex systems using a higher level of abstraction originating from defined knowledge and experience. A fuzzy number *A*~

[[27]](#_bookmark33) is a convex normalized fuzzy set defined on the universe of discourse ffi [the set of all real numbers) with a piecewise continuous membership function and bounded support.

Here, a fuzzy number *A*~ was used as a triangular fuzzy

number (TFN) denoted by *A*~ = (*m*, b, c) if its membership function is of the following form

was considered, we took account of chewing habit as poten- tially carcinogenic agent. In a study it was considered that, irri- tation caused due to high frequency of chewing may cause oral precancer (OSF) even without tobacco [[17]](#_bookmark27). Even dose

l*A*~(*x*) =

*m*—*x*

b

( 1 —

1 — *x*—*m*

c

if *m* — b 6 *x* 6 *m*

if *m* 6 *x* 6 *m* + c

[28]

response relationship in frequency and duration of betel quid chewing without tobacco was found to have an elevated risk

*A*~ may also be represented by (*a*; *a*; *a*¯) where *a* = *m* — b, *a* = *m*

and *a*¯ = *m* + c denote the left point, center and right point of *A*~.

Oral pre-cancer and cancer susceptibility assessment 253

In epidemiological data analysis, there often remains an uncertainty in the form of individual exposure estimates [[29]](#_bookmark34) and increasing knowledge gaps [[30]](#_bookmark38). 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 illit- erate 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]](#_bookmark35). This impreciseness can be well interpreted applying the consensus of fuzzy logic in the epidemiological data analysis. Similarity measures [[32]](#_bookmark35) 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 clin- icoepidemiological data does not consider such human percep- tion for information extraction from massive data to get a crisp solution as does the fuzzy logic [[33]](#_bookmark35). 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 con- sidering the complex vagueness of situations. Therefore intro- duction of the notion of logical fuzzy *If-Then* rule to understand such a complex process can provide a better infor- mation in the decision making process for a certain range of uncertainty [[34]](#_bookmark35).

Thus, this study, intends to apply a fuzzy rule – base for better prediction of malignancy or pre-malignancy susceptibil- ity 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 stud- ies 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.

1. Methods
   1. *Subject selection*

Study population of Terrain and Duars region, the northern region of West Bengal, India, was chosen in this study. How- ever 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.

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

254 S. Banerjee et al.

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.

* 1. *Data analysis strategy*
     1. *Conventional statistical analysis*

Clinicoepidemiological data of the studied subjects were ana- lyzed 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 v2 test [[35]](#_bookmark35) and the cutoff signif- icance 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 indi- cates their protective nature.

* + 1. *Defining fuzzy numbers and rulebase generation*

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

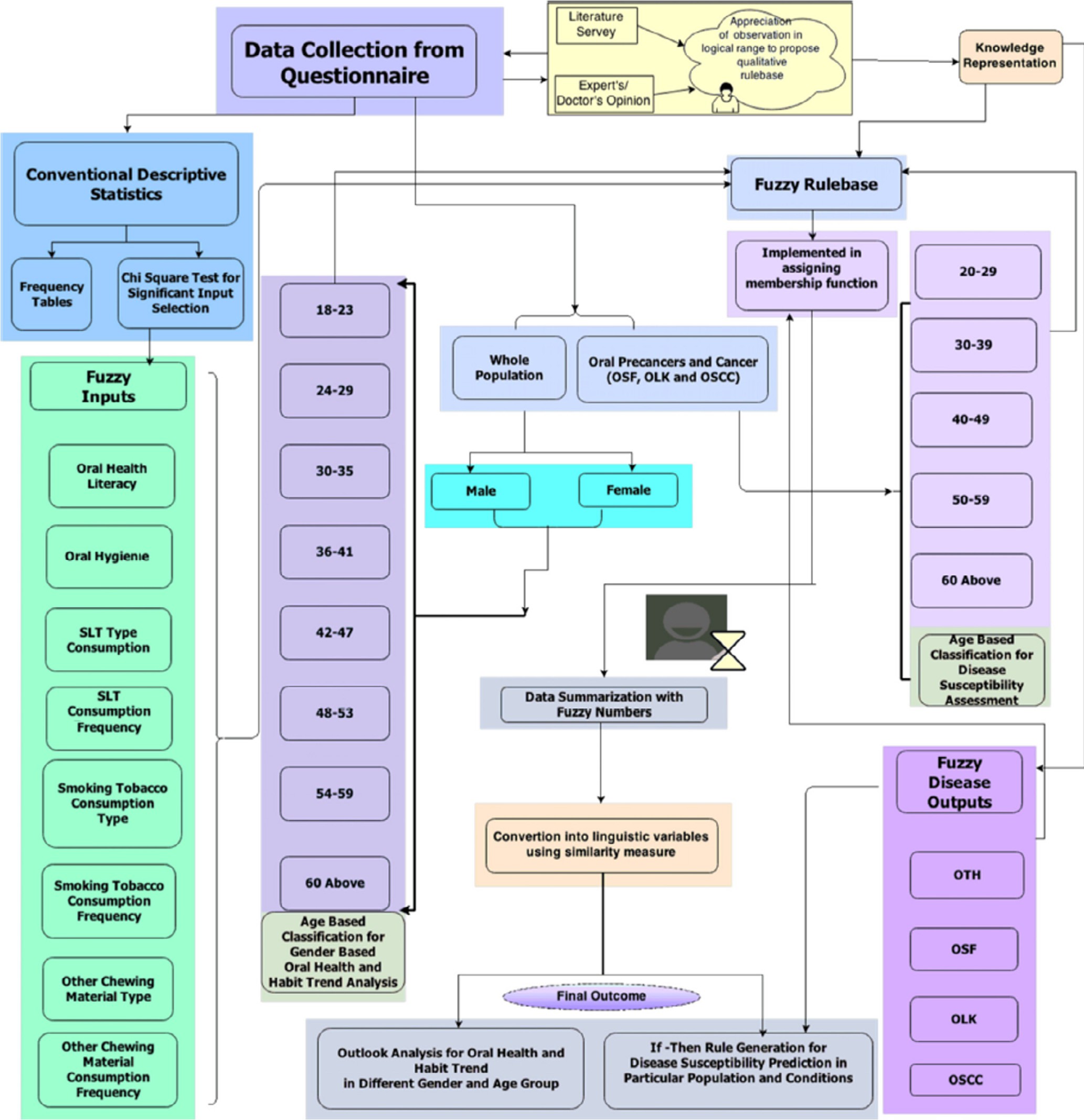


Figure 1 Proposed methodology on application of fuzzy consensus in assessing oral precancer and cancer susceptibility.

numeric data for fuzzy reasoning is important [[36]](#_bookmark35). 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 pre- pared using physician’s intuition and information extracted from the literature review [[11–25]](#_bookmark16). 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 differ- ent 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 dis- eases (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 poten- tiality [[37]](#_bookmark35). Input variables were quantitatively described through intuition with their assigned membership functions in a fuzzy scale of 0–1 in [Table 1](#_bookmark6).

* + 1. *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 cor- responding fuzzy scale. Each fuzzy number, thus obtained was then added for each group to get the final assigned fuzzy num- ber which summarizes the condition of the group in a quanti- tative manner.

For the patients with oral precancer and cancer, the data were summarized in a slightly different way. 41 patients diag- nosed 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](#_bookmark13).

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 ambigu- ous 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](#_bookmark9), [4b](#_bookmark11) and [5](#_bookmark12) 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](#_bookmark7) depicts the workflow of the proposed methodology for application of fuzzy consen- sus for oral precancer and cancer susceptibility assessment.

1. Result and discussion
   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](#_bookmark8). Variables having statistical signifi- cance with the disease outcome are depicted in [Table 3](#_bookmark10) 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 liter- acy rates (49.68% people were with education up to 8th stan- dard 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 Dar- jeeling district, in the Bengali Hindu community in this study.

* 1. *Interpretation of the summarized data*

Most of the current literatures focus on institution based stud- ies 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]](#_bookmark35). 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 esti- mation tool was used for primary feature extraction only. The methodical fuzzy mapping of features and linguistic con- version of assigned membership functions, using Jaccard index presented in tables ([Tables 4a](#_bookmark9), [4b](#_bookmark11) and [5](#_bookmark12)), was adopted for understanding such associations by trend analysis of the data- set. 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 pre- cancers and cancers (OTH) too.

In the context of appreciating feature summarization against the disease conditions, [Tables 4a](#_bookmark9), [4b](#_bookmark11) and [5](#_bookmark12) 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 4a Oral health and habit trend obtained through linguistic conversion of fuzzy numbers using Jaccard index in male.

Age Oral health Oral SLT types SLT Smoking

tobacco type

Low

Smoking tobacco

consumption frequency

Very Low

Other chewing

material type

Other chewing

Disease

group literacy

hygiene consumption consumption

frequency

material frequency chances

18–23 Satisfactory Very Very Low

Good

24–29 Satisfactory Good Very Low

Very Low

Low

Very Low

OTH

Very Low

Low

Very Low

Low

Very Low

OTH

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 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 | 8.417, 1, 0.004 | 0.419 | 0.229, 0.776 |
|  |  | standard 8th |  |  |  |
| 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 |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 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 | Bad | Good | Moderate | Moderate | Medium | Moderate | Low | Low | OSF |
| Above |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 4b Oral health and habit trend obtained through linguistic conversion of fuzzy numbers using Jaccard index in female. | | | | | | |
| Age Oral health Oral SLT types group literacy hygiene 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 Very Low | Very Low | Very Low | Very Low | Low | Very Low | OTH |
| Good |  |  |  |  |  |  |
| 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 Bad Poor Low | Low | Very Low | Very Low | Medium | Medium | OTH |
| above |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

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 Oral

literacy hygiene

SLT types SLT Smoking

consumption consumption tobacco

frequency type

Medium

Smoking tobacco

consumption frequency

Low

Other chewing Other chewing Disease

material type chances

20–29 (M

& F) 30–39 (M

& F) 40–49 (M

& F) 50–59 (M

& F)

60 Above (M & F)

60 above (Male only)

Satisfactory Very Good Low

Moderate

Low

material

frequency

Very Low

OSF

Bad

Satisfactory Medium

Moderate

Low

Very Low

Medium

Low

OLK

Bad

Good

High

High

Medium

Moderate

Medium

Moderate

OLK

Bad

Satisfactory Medium

Moderate

Medium

Moderate

Medium

Low

OLK

Bad

Satisfactory High

High

High

High

Medium

Moderate

OSCC

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 mate- rial 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](#_bookmark9), [4b](#_bookmark11) and [5](#_bookmark12) similarly por- trayed the oral health and habit scenario of the population in a compartmentalized manner, along with the disease predic- tion rules for males, females and diseased patients for different age groups separately.

When the trend between male and female population of dif- ferent age groups ([Tables 4a and 4b](#_bookmark9)) 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]](#_bookmark36). In this population, SLT type and consumption frequency were comparatively lower in females than males which also support the previous findings [[41]](#_bookmark37). 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](#_bookmark9)). The Areca nut and leaves chewing habit type and frequency when considered, the addiction was found to be more in females [[6]](#_bookmark26), which was further elaborated in [Table 4a](#_bookmark9) 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](#_bookmark11)).

Though OSF is associated with areca nut intake in dose dependent manner [[17]](#_bookmark27), 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]](#_bookmark15). 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](#_bookmark9) and [5](#_bookmark12)). Increase in deleteri- ous oral habits such as SLT consumption frequency, was also found in males from age group of 36 years ([Table 5](#_bookmark12)). In case of OSCC prevalence, the present study demonstrated an associa- tion of high smoking tobacco with SLT consumption, poor oral hygiene, bad oral health literacy and high age ([Table 5](#_bookmark12)) and can be corroborated with the findings of the other studies [[13,23]](#_bookmark19).

Hence, the proposed oral pre-cancer, cancer and other oral diseases susceptibility assessment methodology with embedded fuzzy analytical dimensions depicted the association of multi- ple 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.

1. Conclusion

Fuzzy rule-base approach has been utilized for value addition to the findings from conventional statistical approach in defin- ing particular association between significant clinicoepidemio- logical 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 under- lying 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 uncer- tainty 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 dimen- sion of questionnaire data handling involved population of

specific demography, and same methodology can be imple- mented for other demographic conditions as well. It is also first of its kind and can help clinicians and policy makers in adopt- ing 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

If smokeless any 3 If smokeless any 2 If smokeless any 1

If smokeless occasional

If smokeless of 4 types or more If smokeless any 3, 2

If smokeless 1

If smokeless occasional If smokeless all 4

If smokeless any 3, 2

If smokeless 1

If smokeless occasional If smokeless all 4, 3

If smokeless 2

If smokeless 1

If smokeless occasional If smokeless 4, 3, 2

If smokeless 1

If smokeless occasional If smokeless 4, 3, 2

If smokeless 1

If smokeless occasional If smokeless 4, 3, 2

If smokeless 1

If smokeless occasional If smokeless 4, 3, 2

If smokeless 1

If smokeless occasional

Then susceptibility very high

Age group 18–23

Age group 18–23

Age group 18–23

Age group 18–23

Age group 24–29

Age group 24–29

Age group 24–29

Age group 24–29

Age group 30–35

Age group 30–35

Age group 30–35

Age group 30–35

Age group 36–41

Age group 36–41

Age group 36–41

Age group 36–41

Age group 42–47

Age group 42–47

Age group 42–47

Age group 48–53

Age group 48–53

Age group 48–53

Age group 54–59

Age group 54–59

Age group 54–59

Age group 60 and above Age group 60 and above Age group 60 and above

Then susceptibility high

Then susceptibility moderate Then susceptibility low

Then susceptibility very low Then susceptibility very high Then susceptibility high Then susceptibility moderate Then susceptibility low

Then susceptibility very high Then susceptibility high Then susceptibility moderate Then susceptibility low

Then susceptibility very high Then susceptibility high Then susceptibility moderate Then susceptibility low

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

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

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| Age group | Education | smokeless | smokeless | Smoking | Smoking | Chewing | Chewing | Oral | Y (disease |
|  |  | type | frequency | type | frequency | type | frequency | hygiene | 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 | 0.98 | 0.974 | 0.976 | 0.896 | 0.973 | 0.837 | 0.955 | 0.984 | 0.936 |
| (M) |  |  |  |  |  |  |  |  |  |
| 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 | 0.985 | 0.966 | 0.939 | 0.914 | 0.966 | 0.832 | 0.937 | 0.969 | 0.928 |
| (Disease) |  |  |  |  |  |  |  |  |  |
| 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](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0005) [Cancer Epidemiol 2012:17](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0005).
2. [Sankaranarayanan R. Oral cancer in India: an epidemiologic and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0010) [clinical review. Oral Surg Oral Med Oral Pathol 1990;69(3):](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0010) [325–30](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0010).
3. [Messadi DV, Wilder-Smith P, Wolinsky L. Improving oral cancer](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0015) [survival: the role of dental providers. J California Dental Assoc](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0015) [2009;37(11):789–98](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0015).
4. [Reichart PA. Identification of risk groups for oral precancer and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0020) [cancer and preventive measures. Clin Oral Invest 2001;5(4):207–13](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0020).
5. [Warnakulasuriya S. Global epidemiology of oral and oropharyn-](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0025) [geal cancer. Oral Oncol 2009;45(4):309–16](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0025).
6. [Guha N, Warnakulasuriya S, Vlaanderen J, Straif K. Betel quid](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0030) [chewing and the risk of oral and oropharyngeal cancers: a meta-](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0030) [analysis with implications for cancer control. Int J Cancer](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0030) [2014;135(6):1433–43](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0030).
7. [Warnakulasuriya S, Johnson N, Van der Waal I. Nomenclature](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0035) [and classification of potentially malignant disorders of the oral](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0035) [mucosa. J Oral Pathol Med 2007;36(10):575–80](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0035).
8. [Kurokawa H, Matsumoto S, Murata T, et al. Immunohistochem-](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0040) [ical study of syndecan-1 down-regulation and the expression of](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0040) [p53 protein or Ki-67 antigen in oral leukoplakia with or without](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0040) [epithelial dysplasia. J Oral Pathol Med 2003;32(9):513–21](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0040).
9. [Moutasim KA, Jenei V, Sapienza K, et al. Betel-derived alkaloid](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0045) [up-regulates keratinocyte alpha-beta6 integrin expression and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0045) [promotes oral submucous fibrosis. J Pathol 2011;223(3):366–77](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0045).
10. [Bokor-Bratic´ M, Vucˇkovic´ N. Cigarette smoking as a risk factor](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0050) [associated with oral leukoplakia. Arch Oncol 2002;10(2):67–70](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0050).
11. [Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0055) [tobacco and cancer. Lancet Oncol 2008;9(7):667–75](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0055).
12. [Carr AB, Ebbert JO. Interventions for tobacco cessation in the](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0060) [dental setting. A systematic review. Community Dent Health](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0060) [2007;24(2):70–4](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0060).
13. [Rodu B, Cole P. Smokeless tobacco use and cancer of the upper](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0065) [respiratory tract. Oral Surg Oral Med Oral Pathol Oral Radiol](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0065) [Endodontol 2002;93(5):511–5](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0065).
14. [Jayalekshmi PA, Gangadharan P, Akiba S, Koriyama C, Nair](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0070)

[RR. Oral cavity cancer risk in relation to tobacco chewing and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0070) [bidi smoking among men in Karunagappally, Kerala, India:](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0070) [Karunagappally cohort study. Cancer Sci 2011;102(2):460–7](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0070).

1. [Rahman M, Sakamoto J, Fukui T. Bidi smoking and oral cancer:](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0075) [a meta-analysis. Int J Cancer 2003;106:600–4](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0075).
2. [Allam E, Zhang W, Al-Shibani N, et al. Effects of cigarette smoke](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0080) [condensate on oral squamous cell carcinoma cells. Arch Oral Biol](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0080) [2011;56(10):1154–61](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0080).
3. [Yang YH, Ho PS, Lu HM, Huang IY, Chen CH. Comparing](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0085) [dose–response measurements of oral habits on oral leukoplakia](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0085) [and oral submucous fibrosis from a community screening](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0085) [program. J Oral Pathol Med 2010;39(4):306–12](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0085).
4. [Jacob BJ, Straif K, Thomas G, et al. Betel quid without tobacco](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0090) [as a risk factor for oral precancers. Oral Oncol 2004;40(7):](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0090) [697–704](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0090).
5. [Muwonge R, Ramadas K, Sankila R, et al. Role of tobacco](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0095) [smoking, chewing and alcohol drinking in the risk of oral cancer](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0095) [in Trivandrum, India: a nested case-control design using incident](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0095) [cancer cases. Oral Oncol 2008;44(5):446–54](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0095).
6. [Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C.](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0100) [The global burden of oral diseases and risks to oral health. Bull](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0100) [World Health Organ 2005;83(9):661–9](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0100).
7. [Paul B, Basu M, Dutta S, Chattopadhyay S, Sinha D, Misra R.](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0105) [Awareness and practices of oral hygiene and its relation to](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0105) [sociodemographic factors among patients attending the general](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0105) [outpatient department in a tertiary care Hospital of Kolkata,](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0105) [India. J Family Med Prim Care 2014;3(2):107–11](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0105).
8. [Chandra Shekar B, Reddy CVK, Manjunath B, Suma S. Dental](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0110) [health awareness, attitude, oral health-related habits, and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0110) [behaviors in relation to socio-economic factors among the](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0110) [municipal employees of Mysore city. Ann Trop Med Public](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0110) [Health 2011;4(2):99–106](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0110).
9. [Guha N, Boffetta P, Wunsch Filho V, et al. Oral health and risk](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0115) [of squamous cell carcinoma of the head and neck and esophagus:](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0115) [results of two multicentric case-control studies. Am J Epidemiol](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0115) [2007;166(10):1159–73](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0115).
10. [Paulander J, Axelsson P, Lindhe J. Association between level of](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0120) [education and oral health status in 35-, 50-, 65- and 75-year-olds.](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0120) [J Clin Periodontol 2003;30(8):697–704](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0120).
11. [Guo Y, Logan HL, Dodd VJ, Muller KE, Marks 3rd JG, Riley](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0125) [JL. Health literacy: a pathway to better oral health. Am J Public](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0125) [Health 2014;104:e85–91](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0125).
12. [Singh H, Gupta MM, Meitzler T, et al. Real-life applications of](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0130) [fuzzy logic. Adv Fuzzy Syst 2013](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0130).
13. [Dubois D, Prade HM. Fuzzy sets and systems: theory and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0135) [applications. Academic Press; 1980](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0135).
14. [Chakraborty D. Structural quantization of vagueness in linguistic](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0140) [expert opinions in an evaluation programme. Fuzzy Sets Syst](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0140) [2001;119(1):171–86](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0140).
15. [Spiegelman D. Approaches to uncertainty in exposure assessment](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0145) [in environmental epidemiology. Annu Rev Public Health](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0145) [2010;31:149](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0145).
16. [Viswanath K, Breen N, Meissner H, et al. Cancer knowledge and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0150) [disparities in the information age. J Health Commun 2006;11(S1):](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0150) [1–17](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0150).
17. [Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0155) [disease spread and public health responses to bioterrorism and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0155) [emerging diseases. Proc Natl Acad Sci 2006;103(42):15693–7](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0155).
18. [Hwang CM, Yang MS. New similarity measures between gener-](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0160) [alized trapezoidal fuzzy numbers using the Jaccard index. Int J](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0160) [Uncertainty Fuzziness Knowledge-Based Syst 2014;22(6):831–44](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0160).
19. [Abbod MF, von Keyserlingk DG, Linkens DA, Mahfouf M.](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0165) [Survey of utilisation of fuzzy technology in medicine and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0165) [healthcare. Fuzzy Sets Syst 2001;120(2):331–49](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0165).
20. [Papageorgiou EI. A new methodology for decisions in medical](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0170) [informatics using fuzzy cognitive maps based on fuzzy rule-](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0170) [extraction techniques. Appl Soft Comput 2011;11(1):500–13](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0170).
21. [Lacagnina V, Leto-Barone MS, La Piana S, La Porta G, Pingitore](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0175) [G, Di Lorenzo G. Comparison between statistical and fuzzy](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0175) [approaches for improving diagnostic decision making in patients](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0175) [with chronic nasal symptoms. Fuzzy Sets Syst 2014;237:136–50](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0175).
22. [Hong T-P, Lee C-Y. Induction of fuzzy rules and membership](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0180) [functions from training examples. Fuzzy Sets Syst 1996;84(1):](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0180) [33–47](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0180).
23. [Bouquot JE, Speight PM, Farthing PM. Epithelial dysplasia of](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0185) [the oral mucosa—diagnostic problems and prognostic features.](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0185) [Curr Diagnostic Pathol 2006;12(1):11–21](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0185).
24. [Patil PB, Bathi R, Chaudhari S. Prevalence of oral mucosal](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0190) [lesions in dental patients with tobacco smoking, chewing, and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0190) [mixed habits: a cross-sectional study in South India. J Family](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0190) [Community Med 2013;20(2):130](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0190).
25. [Sujatha D, Hebbar PB, Pai A. Prevalence and correlation of oral](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0195) [lesions among tobacco smokers, tobacco chewers, areca nut and](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0195) [alcohol users. Asian Pac J Cancer Prev 2012;13(4):1633–7](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0195).
26. [Kingdon GG. Does the labour market explain lower female](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0200) [schooling in India? J Develop Stud 1998;35(1):39–65](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0200).
27. [Rani M, Bonu S, Jha P, Nguyen S, Jamjoum L. Tobacco use in](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0205) [India: prevalence and predictors of smoking and chewing in a](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0205) [national cross sectional household survey. Tobacco Control](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0205) [2003;12(4), e4 (1–8)](http://refhub.elsevier.com/S1110-8665(16)00003-7/h0205).