STUDYING VARIATIONS IN CAVITY EXPERIENCE AMONG THE MAJOR ETHNIC GROUPS OF DOOARS REGION USING DATA MINING

Amrita Kundu¹, Dr. Rajashri Kundu²

¹Department of Computer Science, Banasthali Vidyapith (India)
²Assistant Professor, Department of Conservative Dentistry and Endodontics, MMDCH (India)

ABSTRACT
Dental caries is a common problem today though its existence can be traced even in ‘Sushruta Samhita’ where words like ‘mukha roga’ or ‘krimi dantaka’ have been mentioned. Till date almost everyone need to visit a dentist to get these caries treated and the rate is increasing day by day. So studying the reason behind this has proved to be the need of time. Studies are carried out on a large population of vivid type to find out the factors governing the caries prevalence like age, sex, ethnicity etc among people and suggest the possible solutions. Data mining has shown remarkable use in knowledge discovery from huge data sets. So in this paper we used data mining techniques to study the caries experience among the major ethnic groups of Dooars area of West Bengal, India to see how it varies among them and the possible reasons behind the variations.

Keywords- Caries, Caries Prevalence, Data Mining, DMF Index, Ethnicity.

I INTRODUCTION

1.1. Problem

According to Shafer (2006), Caries is an irreversible microbial disease of the calcified tissue of the teeth that can be characterized by demineralization of the inorganic portion and destruction of the organic substance of the tooth that leads to cavitations [1]. The history behind dental caries is very long. Dental caries problem was noticeably less through the Bronze and Iron ages mostly due to dependency on animal protein food. The increase of caries during the Neolithic period may be attributed to the increase of plant foods containing carbohydrates in them. It is believed that the beginning of rice cultivation lead to an increase in dental caries. 5000 BC old Sumerian text describes “toothworm” as the cause of caries [2]. Though Pierre Fauchard was the first one to reject the idea of worms as the cause of tooth decay [2], it was W.D. Miller who conducted a series of studies in the year 1890 and found that bacteria inhabited the mouth and they produced acids that dissolved tooth structures in the presence of fermentable carbohydrates [3]. Cavity is one of the most common disorders of childhood and has many undesirable effects on child population like-

a. Pain.
b. Improper sleep.
c. Problems in speech.
There are some factors affecting the prevalence of caries beside carbohydrate and microbes. According to Shafer they are ethnicity, age, gender and family history [1].

- **Ethnicity:** Some studies done in the past has shown remarkable difference between various ethnic groups. For example among the American blacks and whites living in the same geographical area under similar conditions [1].

- **Age:** Several studies have shown that about 20% of children have experienced dental caries in their dentition. By the age of 12, 90% of children experience dental caries [1].

- **Gender:** Studies indicate that the total caries experience in permanent teeth is greater in females than in males of the same age. This is attributed largely to the fact that the teeth of girls erupt at an early age than the teeth of boys of the same age [1].

- **Family history:** Siblings of individuals with high caries susceptibility are also generally caries active, whereas siblings of carries immune individuals generally exhibit low caries rate (Garn et al) [1].

### 1.2. Previous work

1.2.1. Creighton W.E carried out an Epidemiological study on dental caries experience of Negro and Caucasian children in Portland, Oregon in the year 1969. Baseline data was collected from 1374 children and 761 black children to make the cross racial comparison of prevalence of dental caries. He found that black children have fewer dental caries than the whites [5].

1.2.2. Downer M.C made a comparison of dental caries in European and Negro girls aged 13 and 14 years and of periodontal disease in girls aged 11 to 14 years attending a secondary school in London in the year 1970. The Negro girls had a significantly lower prevalence of caries experience and a significantly higher prevalence of destructive periodontal disease than the Europeans [6].

1.2.3. Bedi R, Uppal RD conducted an oral health survey on the oral health of minority ethnic communities in the United Kingdom in 1996 and found significant differences in caries experience among different ethnic groups [7].

1.2.4. Okunseri C, Badner V, Kumar J, Cruz, G D conducted a study to estimate dental caries prevalence and treatment need among racial/ethnic minority school children in the Bronx in 2002. Hispanic children (mean DMFS of 1.71) experienced more dental caries as compared to African-Americans (mean DFMS of 1.14) [8].

1.2.5. Xiaojuan Zeng, Yan Luo, Minquan Du, Raman Bedi measured the dental caries experience of preschool children from different ethnic groups in Guangxi Province in China in 2005. Zhuang children had a significantly higher prevalence of rampant caries (13% vs. 9%), mean DMFT (3.36 vs. 2.66) and mean DMFS (5.10 vs. 3.76) than the Han children. An analysis of the factors contributing to the racial variation showed that dietary influence was not the only factor responsible. The variation in caries susceptibility based on racial differences has been postulated [9].

1.2.6. Jalili et al. (1993) designed a study amongst the tribal children (number=1016; aged 6-13 years) in Mandu, Madhya Pradesh. The tribal children exhibited a low prevalence of caries in teeth, in permanent as well as permanent dentition, compared to rural as well as urban Indian children of the same age group [10].
1.2.7. Retnakumari studied the prevalence of dental caries in school children (aged 6-12 years) in Kerala in the year 1999. The findings indicated that highest prevalence (75.9%) was recorded in the subjects of age group of 10 years and the lowest in the age group of 8 years. Statistically significant association was found with dental caries and oral hygiene status [11].

1.2.8. Singh et al conducted a study on the prevalence of dental caries in school children (aged 12-16 years) of rural settings of Haryana in 1999, was found to be 39.4%. It was 37.9% in boys and 42.6% in girls. A significant increase in prevalence of dental caries with age was observed (33% in 12 year old children to 45.8% in 24 year old children) [12].

1.2.9. Chawla et al measured the trend of dental caries in 2000 in children of Chandigarh and noticed a continuous decline among the children (12-15 years) but a slight increase was found in the dental caries status of 5-6 year children [13].

1.2.10. David et al. (2005) conducted a study on dental caries and its associated factors in 12 year old school children in Thiruvananthapuram (Kerala). The authors observed the prevalence of dental caries in permanent dentition as 27%. The study indicated that urban living conditions were associated with more dental caries. The pattern of prevalence of dental caries in the primary dentition among 5 year old children (number=1009) in urban Pondicherry was assessed by Saravan et al. (2006) [14].

1.2.11. Dutta A (1965) conducted a study in 1424 children in Calcutta to study the prevalence of caries. It was found that the DMF increases with age. The observed DMF was 0.18 for 6 years and for 12 years it was 2.40. Caries prevalence was high among the low socioeconomic group as compared to high socioeconomic group [15].

1.2.12. Saha et al (1996) studied the prevalence of dental caries and oral hygiene status in rural and urban areas of Calcutta. 9600 children of age group 6-14 years were chosen for the examination. Dental caries prevalence in urban area was 68.0% in boys and 69.9% in girls respectively in the age group of 6 years and the combined percentage was 69.0%. In 12 years old it was 71.1% and 72.2% respectively in boys and girls and the combined percentage was 71.4%. In 6 years old group of the rural area the prevalence was 54.0% and 54.6% for boys and girls respectively and the combined percentage was 54.3%. In the age group of 12 years the prevalence was 55.3% for boys and 57.35% for girls with a combined value of 66.0% [16].

1.2.13. Yoh Tamaki et al (2008) proposed a model for the prediction of dental caries using data mining. About 500 students (age 5-8 years) were examined and data was collected mostly using questionnaires. A decision tree was designed to assess the data. MS levels in saliva were tested. The mean number of decayed and filled teeth (DF) was about 0.054 at the baseline [17].

1.3. Purpose

In this paper we have used data mining techniques on the data set of four major ethnic groups of Dooars region of the state of West Bengal (India) [Annexure-1(Fig. a.)]. Dooars or Duars are the footplains of the eastern Himalayas in north-East India including parts of Bhutan. The Sankosh river divides this area into the Eastern Dooars or the Assam Dooars and the Western Dooars or the Bengal Dooars consisting of an area of about 8800 square kilometers or 3400 square miles [18]. The various major towns in this region are Goalpara, Kokrajhar, Barpeta, Bongaigaon and Dhubri in Assam and Siliguri, Coochbehar, Alipurduar, Jalpaiguri, Malbazar, Mainaguri and Birpara in West Bengal.
Bengal. Also Phuentsholing, the commercial capital of Bhutan can be considered a part of this region [19]. The native people of this region generally have East Asian features. They include numerous tribes like Boro, Lepcha, Mech, Limbu, Orawn, Rajbanshi etc. Also a large area of Dooars is occupied by Bengali and Nepali population [20].

Ethnic group constitutes of an aggregation of biological and socio-cultural characteristics. Various Castes, Scheduled Castes, Scheduled Tribes and Communities are included in ethnic group (the names of Scheduled Castes and Scheduled Tribes after Manual of Election Law 1982, Government of India, New Delhi) [21]. In this paper the major ethnic groups selected were:

a) Bengali-The Bengali people are an ethnic community native to the historic region of Bengal (now divided into India and Bangladesh) in the southern region of Asia. They speak Bengali which is an Indo-Aryan language [22].

b) Nepalese-Nepalese are descendants of migrants from parts of earlier Greater Nepal, Tibet, India and parts of Burma and Yunnan along with the native tribal population [23].

c) Rajbanshi-The Rajbanshi were primarily Hindus (both Shauva and Vaishnabhite) [24].

d) Adivasi-The Adivasi people of North Bengal largely residing in different parts of Dooars, specially the tea garden areas make a large portion of the population. The term Adivasi refers to a set of ethnic and tribal groups claimed to be the aboriginal population of India [20].

The cavity experience in these four ethnic groups has been studied in this paper and prevalence of decayed, missing and filled tooth has been compared using data mining techniques. Prevalence refers specifically to all current cases (old and new) existing at a given point of time or over a period of time in a given population [25]. Before prevalence can be recorded, a quantitative measurement is required that will reflect accurately the extent of the disease in a population. In this case of dental caries, the measurement used is called the DMF index [26] where:

- **D**-The number of decayed teeth with untreated carious lesions.
- **M**-The number of teeth which have been extracted due to caries and are therefore missing.
- **F**-The number of filled teeth.

![Decayed tooth](a) ![Missing tooth](b) ![Filled tooth](c)

**Fig. 1.** (a) Decayed tooth, (b) Missing tooth, (c) Filled tooth. [Source- Dr. Rajashri Kundu].

DMF index is the arithmetic index of the cumulative caries attack in a population [25]. The advantages of this index are:

- It’s simplicity.
- It’s versatile.
- It’s rapidity.
- It’s universal acceptance [25].

### 1.4. Contribution
Dooars region is spread across an area of 8800 square kilometer [18] with about 13 major towns and is a major tourist attraction. The tea gardens in the area contribute to the supply of high quality tea [27] for which India is famous. However irrespective of its importance no study report have been found by us about the dental caries, oral hygiene or comparative levels of caries status among the ethnic groups residing there (Rajbanshi, Nepali, Bengali and Adivasi etc [20]). This study conducted on the major ethnic groups (Bengali, Nepali, Rajbanshi and Adivasi) can throw light on the factors leading to the variations in the cavity experience among them and suggest ways to modify the present scenario mostly by means of preventing dental caries as restoration techniques can be out of reach of many people residing in the area with a low socioeconomic status. Also we can compare the caries index of this area with the rest of the world to better understand its variation and prevalence [Annexure-1(Fig. c.)].

II MATERIALS AND TECHNIQUES USED

2.1. Essential instruments and materials used by the doctor:

- Dental mirror,
- Tweezers
- Gloves
- Mask
- Disinfectants

2.2. Tools used for data collection and mining:

- Questionnaire
- Interview
- Observations
- Microsoft Excel 2007
- SPSS 16.0

2.3. Information about study area and the studied population

Some necessary information was collected before the study like:

- Map of the area from census office.
- Surnames of the different ethnic groups from the Block Development Office.
- List of secondary schools in the area from the office of Secondary School Inspector.

III RESEARCH METHODOLOGY

For this study we chose the school going children belonging to the four ethnic groups of Dooars of age 12-15 years. The reason behind selecting this age group was that at this age generally all permanent teeth are erupted. It is the age of early permanent dentition. So it is ideal for measuring the severity of dental caries and its variation with ethnicity too. For collecting the data dental checkups were conducted in some of the schools of Dooars region like Malbazar and more than 2000 students of this age group were examined among which 500 students of each ethnic group (both male and female) were considered for the study. The school authorities were informed the date of checkups from beforehand and the children along with their parents participated in it. Inter racial marriage history of last three generations were collected from the parents. Printed questionnaires were distributed among the students where they had to fill some basic queries like their name, age, gender, name of the institution, address and food habit. The oral examination was then done by Dr. Rajashri Kundu using the necessary instruments and the number of decayed, missing and filled teeth was filled by her in the questionnaire itself. The inclusion criteria were:
Only the permanent teeth, including the third molars were considered.

Teeth were considered “decayed” when there was presence of caries, teeth with restoration and recurrent caries, teeth with temporary restoration (WHO DMF index modification, 1986) or those with decay indicated for restoration. They were counted under ‘D’.

Teeth were considered missing only when extracted due to carries and not due to any orthodontic reason or accident or those which were congenitally missing. They were counted under ‘M’.

A tooth was considered to be filled only when it had some kind of filling (e.g. Amalgam filling). They were counted under ‘F’.

A tooth was considered to be present even when the crown was totally destroyed and only the root was present.

The students were also asked about how they take care of oral hygiene, the kind of toothpaste they used and number of times they brushed. Fluoride concentration in the water plays an important role in oral cavity. Due to universal presence of fluoride in the earth’s crust, all water contains fluoride in varying concentration. This variation is noticed not only from place to place but also within the same locality. But as West Bengal comes under very low fluoride concentration area, its affect was not considered [Annexure-1 (Fig. b.)].

3.1 Classification
Finally, the students were classified into four major ethnic groups (Bengali, Nepali, Rajbanshi and Adivasi) by comparing their surnames with the list provided by the Block Development Office. Data of 500 students belonging to each group and the student names with no ambiguity were selected and stored in Microsoft Excel 2007 along with the number of decayed ‘D’, missing ‘M’, filled ‘F’ and total number of decayed, missing and filled teeth ‘DMF’ of each student.

3.2. Data Mining
Data mining [28] was then applied on this summarized data using Microsoft Excel 2007 and SPSS 16.0 to find out:

- Mean and standard deviation of decayed teeth ‘D’ of all the groups.
- Mean and standard deviation of missing teeth ‘M’ of all the groups.
- Mean and standard deviation of filled teeth ‘F’ of all the groups.
- Mean and standard deviation of decayed, missing and filled teeth ‘DMF’ of all the groups.
- Prevalence of ‘D’, ‘M’, ‘F’ and ‘DMF’ among all the groups.
- Frequency of ‘D’, ‘M’, ‘F’ and ‘DMF’ in all the four groups using histograms.

IV RESULTS AND OBSERVATIONS
The results obtained are summarized in the table below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bengali</td>
<td>0.820</td>
<td>1.110886</td>
<td>0.106</td>
<td>0.3674</td>
<td>0.162</td>
<td>0.5062</td>
<td>1.088</td>
<td>1.6091</td>
</tr>
<tr>
<td>Nepali</td>
<td>1.036</td>
<td>1.168518</td>
<td>0.130</td>
<td>0.4577</td>
<td>0.116</td>
<td>0.4806</td>
<td>1.282</td>
<td>1.5792</td>
</tr>
<tr>
<td>Rajbanshi</td>
<td>0.814</td>
<td>1.261184</td>
<td>0.020</td>
<td>0.1401</td>
<td>0.030</td>
<td>0.2987</td>
<td>0.864</td>
<td>1.2823</td>
</tr>
<tr>
<td>Adivasi</td>
<td>0.502</td>
<td>0.869199</td>
<td>0.000</td>
<td>0.0000</td>
<td>0.024</td>
<td>0.2179</td>
<td>0.526</td>
<td>0.8825</td>
</tr>
</tbody>
</table>
The comparative analysis of mean of ‘D’ (Fig. 1.) showed significant difference between groups with highest value of ‘D’ among Nepali (1.036) and lowest among Adiwasi (0.502). Bengali and Rajbanshi showed intermediary values of ‘D’ (Bengali 0.82 and Rajbanshi 0.814).

The mean of ‘M’ when compared among different groups (Fig. 2.), showed significant difference in result, Nepali having the highest number of missing teeth due to caries (0.13) followed by Bengali (0.106) and Rajbanshi (0.02). Surprisingly Adivasi people had no missing tooth due to caries with a mean ‘M’ equal to zero.

When the mean of ‘F’ was compared among the groups, a remarkable decrease in the value was noticed from Bengali to Adivasi.

Comparing the mean DMF of all the groups the highest ‘DMF’ was found among the Nepali (1.282) followed by Bengali (1.088), Rajbanshi (0.864) and Adivasi (0.526).

**TABLE II. Prevalence chart of D, M, F and DMF among the groups**

<table>
<thead>
<tr>
<th>Race</th>
<th>Prevalence of ‘D’</th>
<th>Prevalence of ‘M’</th>
<th>Prevalence of ‘F’</th>
<th>Prevalence of ‘DMF’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bengali</td>
<td>44%</td>
<td>9%</td>
<td>11%</td>
<td>44%</td>
</tr>
<tr>
<td>Nepali</td>
<td>52%</td>
<td>8%</td>
<td>6%</td>
<td>52%</td>
</tr>
<tr>
<td>Rajbanshi</td>
<td>37%</td>
<td>2%</td>
<td>1%</td>
<td>37%</td>
</tr>
<tr>
<td>Adivasi</td>
<td>30%</td>
<td>0%</td>
<td>0%</td>
<td>30%</td>
</tr>
</tbody>
</table>
Fig. 5. Prevalence of Decayed teeth ‘D’ among the groups.

Fig. 6. Prevalence of Missing teeth ‘M’ among the groups.

Fig. 7. Prevalence of Filled teeth ‘F’ among the groups.

Fig. 8. Prevalence of ‘DMF’ among the groups.

Fig. 9. Histogram of Bengali ‘DMF’.

Fig. 10. Histogram of Nepali ‘DMF’.

The highest number of filled teeth was found among Bengali group (11%) followed by Nepali (6%) and Rajbanshi (1%). Adivasi group showed 0% prevalence of filled tooth (Fig. 7).

The result of prevalence of ‘DMF’ showed highest value among Nepali (52%). Bengali group had the next highest value (44%) followed by Rajbanshi (37%) and Adivasi (30%) (Fig. 8).

Histogram of ‘DMF’ of Bengali group showed that nearly 300 students out of 500 had sum of D, M, F equal to 0. But the highest value noticed in this group is 6 and that count is more than all other groups (Fig. 9).
Histogram of ‘DMF’ of Nepali showed that more than 200 students out of 500 had sum of D, M, F equal to 0. Around 100 students had the value 1 and the highest value is 6 but the count is less than that of the Bengalis (Fig. 10.).

**Fig. 11. Histogram of Rajbanshi ‘DMF’**

Histogram of Rajbanshi showed that more than 300 students out of 500 had the sum of D, M and F equal to 0. Less than 50 students had the value 4 and the rest 1 to 3 (Fig. 11.).

Histogram of sum of D, M, F of Adivasi showed that about 350 students out of 500 had the value equal to 0 which is greater than any other group. 3 is the highest value among them which is less than the rest of the groups (Fig. 12.).

**Fig. 12. Histogram of Adivasi ‘DMF’**

Eye balling the whole result gives an idea about the status of caries among the four major ethnic groups. Caries experience is highest among Nepali and lowest among Adivasi.

**V DISCUSSION**

Oral health status and related behavior among minority ethnic groups are critical issues in a multicultural society. Socio-cultural changes or lifestyle issues may help to obtain a better understanding of disease etiology, thereby contributing to the development of effective prevention strategies. Dooars area also comprise of a multicultural society including the culture of Bengali, Nepali, Rajbanshi and Adivasi where Adivasi people are considered to be the minority. The socioeconomic status, lifestyle etc are different for these groups of people. Thus the results shown in our study with different value of ‘DMF’ index or DMFT (decayed-missing-filled teeth) among the ethnic groups are well matched with the previous studies attempted on various ethnic groups across the world [Annexure-1 (Fig .c.)]. The major factors we found important are discussed under:
5.1. Dietary factors: Most dietary factors found to be significant are related to the consumption of sugar—its amount, frequency or time of consumption. In our study, the consumption of sugar and sticky carbohydrates in the form of refined food are a new trend among the advancing ethnic groups of Dooars, like Bengali and Nepali which can contribute to their larger value of DMF index than the Rajbanshi and Adivasi children who consume mainly fibrous food.

5.2. Oral hygiene maintenance: Oral hygiene maintenance has been found to be a significant risk factor for dental caries. In our study, Bengali and Nepali children were found more conscious about oral hygiene maintenance than the tribal groups including Rajbanshi and Adivasi children of Dooars. But still the Rajbanshi and Adivasi group had lower caries index which may be related to their tribal ethnicity.

5.3. Socioeconomic status: In our study population, Bengalis are socioeconomically sounder than the rest of the population and hence show more concern about dental care. This is supported by the evidence of early treated teeth like filled teeth. On the other hand the other less socioeconomically sound groups did not show much concern about the dental care.

5.4. Tea intake: Tea extracts has been reported to contain substances, notably polyphenols that has antibacterial properties against oral pathogens such as Streptococcus mutans which is the bacteria associated closely with the caries of teeth. Some studies suggest that a diet supplemented with green tea may be beneficial in dental caries management [29]. In the present study, Adivasi group of children are mainly living in areas of tea garden where they consume tea on a regular basis. This could result in the lower prevalence of dental caries among them.

5.5. Fluoride: Our study area is reported to have much less Fluoride level in drinking water than needed to reduce the caries index [Annexure-1 (Fig. b.)]. So its contribution in the variation of caries index among the groups cannot be considered.

VI CONCLUSION
In the current study it was observed that there is a definite difference in the caries level among the ethnic groups. A majority of carious lesions were unrestored in all the groups and were active by nature. If allowed to continue would certainly result in complications. The following measures are suggested from our side that can be taken to promote better dental care among them-

- Dental health services should focus primarily on the prevention of dental caries since the cost of management and restoration of teeth may not be within the reach of many residing in the Dooars area with a low socioeconomic status.
- The role of primary health care workers should not be underestimated as they can perform the primary prevention programs among the school children in the form of oral health education with emphasis on their food habits and regular checkups.
- Use of fluorides especially in toothpastes has been documented with much benefit in caries prevention, and therefore its use in caries prevention should be encouraged.
- At primary health centre a special manpower as “oral health educator” can be kept after training the existing health educators by conducting crash courses so that they can take care of oral health education of the people residing in Dooars.
Atraumatic Restorative Treatment (ART) is a well accepted treatment by rural school children. Specific manpower (A special dental auxiliary), named as ‘RURAL SCHOOL DENTAL NURSE’ can be trained to deliver ART to the rural school children.

Beside the chances of error in the reports so many subjects, there are also chances of observer biasness. In future more surveys can be conducted with larger data and vivid ethnic groups and different data mining techniques can be applied on those for extracting various kinds of knowledge. Training using Neural Networks can be introduced to manage such huge data in an intelligent manner and better information retrieval techniques can be applied for easier and faster retrieval.

VII ACKNOWLEDGEMENT

I would like to take this opportunity to express my gratitude to all those people who have in various ways helped us including the schools, students and their parents without whose cooperation our survey could not be successful. My sincere thanks goes to Dr. Saurabh Mukherjee (Associate Professor, Dept. of Computer Science, Banasthali Vidyapith) and Dr. Arnab Mukherjee (MDS, Oral and Maxillofacial Pathology) for sharing their wisdom, knowledge and expertise to better understand the issues related to this work.

REFERENCES


[27] http://www.westbengaltourism.gov.in/web/guest/dooars-main

[28] Jiawei Han, Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems, 2011, 3rd edition).


Biographical Notes

Ms. Amrita Kundu is presently pursuing M.Tech. first year in Computer Science Engineering Department from Banasthali University, India.

Dr. Rajashri Kundu is an Assistant professor at Conservative Dentistry and Endodontics Department MMDCH, India.
ANNEXURE

Fig. a. Map of Dooars area.

Fig. b. Fluoride level in India.

Fig. c. DMFT worldwide.