## Page Numbers

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Introduction to the Study</td>
<td>2-7</td>
</tr>
<tr>
<td>1.2</td>
<td>Introduction to Fluoride</td>
<td>8-13</td>
</tr>
<tr>
<td>1.3</td>
<td>Introduction to Jowar</td>
<td>14-17</td>
</tr>
<tr>
<td>1.4</td>
<td>Introduction to Case control design</td>
<td>18-32</td>
</tr>
<tr>
<td>1.5</td>
<td>Introduction to Nutritional Epidemiology</td>
<td>33-43</td>
</tr>
<tr>
<td>1.6</td>
<td>Study Area Introduction</td>
<td>44-45</td>
</tr>
</tbody>
</table>
1.1 INTRODUCTION

Fluorine, the most electronegative element, has notable chemical and physiological properties that are of great importance for human health and well being. (1)

The ability of fluoride to reduce dental caries at optimum levels (0.8-1.2ppm) is well documented. Additionally, its beneficial influence on the tooth form and anti plaque action has been reported (2, 3). While all these are the positive aspects of fluoride, it also has a darker side. Excessive fluoride ingestion during the period of tooth formation leads to the mineralization disorder, dental fluorosis (3). Fluoride inhibits enzymes such as metallo and serine proteinases (3) that are required for calcification of enamel matrix, leading to lowering the calcium content, resulting in hypo calcified enamel.

Dental fluorosis has significant implications because of the cosmetic nature (4) of the condition (associated with impaired self-image or loss of self-esteem) and its role as a biomarker for skeletal fluorosis (4). Dental fluorosis is irreversible, but preventable by appropriate and timely intervention.

Mottled teeth were recognized as a phenomenon more than a century ago. Earlier observations of this relatively common geographically defined and distinctive pattern of abnormality of enamel initiated clinical, epidemiological and laboratory research which identified an association between mottled enamel [dental fluorosis] and fluoride.

Investigations ranging from fundamental research to cellular physiology, to the epidemiology of fluorosis and dental caries, resulted in successful application of fluorides, as a major preventive agent against dental caries (1). The effect of
fluoride on biological system, particularly its mechanism of action in preventing caries has been recommended by prodigious investigative resources. For more than 60 years, dental fluorosis has remained a major issue in the mainstream of basic, applied and health services research because (1):

- It has led to the discovery of a trace element with the inhibitory property of dental caries.
- It is a permanent record of abnormal tooth development.
- Mechanism of fluorosis remains unexplained.
- Fluorosis has important implications in the community because of its cosmetic, environmental and preventive role (5, 6).

Endemic fluorosis remains a challenging and extensively studied national health problem in India, which has affected 17 out of the 32 states (5). In the affected states, 10% to 25% of the rural population has been estimated to be at a higher risk (7, 8) of developing fluorosis. Children up to the age of eight years are most prone to fluorosis, as their dental tissues are in the formative stage during this period (9, 10).

According to the World Health Organization (WHO), 1.5 ppm is the safe limit of fluoride in drinking water (11-14). The fluoride concentration in drinking water ranges from 0.02 ppm to 15 ppm in different endemic areas of India (5). The major source of fluoride intake in endemic areas is drinking water, although some food materials contribute to a considerable amount of total fluoride intake (15).

The increased prevalence of dental fluorosis in India may be due to

- Population growth resulting in increased need of potable water.
- Indiscriminate digging of bore wells and extensive usage of ground water.
• Total unawareness of the importance of water quality and drinking water from all the sources (5).

Dose and duration of fluoride intake, age of the individual, renal efficiency, poor nutritional conditions are some of the major factors that influence fluoride toxicity and bring about variations in the clinical presentation (10).

The Rajiv Gandhi National Drinking Water Mission (RGNDWM; a Govt. of India initiative) has reported that most of the areas in North Karnataka, India fall under the endemic fluoride belts (16). Preliminary studies in these areas have reported a 58% prevalence of dental fluorosis at 0.74 ppm fluoride concentration in drinking water (17) and occurrence of 13% objectionable fluorosis (severe dental fluorosis). This is an unusual finding and cannot be attributed to water borne fluorides alone (18, 19).

Although other factors like tropical climate have been attributed, the severity is not comparable with other endemic regions in India, which have similar climatic conditions (17, 18). High incidence of dental fluorosis even with low fluoride content in water is probably related to dietary factors (20, 21). The variation in the prevalence and severity of fluorosis within the same community or between communities with similar fluoride content in drinking water indicates that dental fluorosis is not due to waterborne fluorides alone (19). While fluoride content of drinking water is considered to be the most important factor responsible for endemic fluorosis, the prevalence and severity of the dental fluorosis do not always run parallel with levels of fluoride in drinking water.

Interaction between fluoride and nutrients is considered to be very important in understanding the nature of the disease (14). Dietary habits have a
key role in determining the severity of dental fluorosis. This is evident based on reports from Tanzania and Kenya where people who use “Trona (Magadi),” a type of salt to tenderize the vegetables, frequently suffer from fluorosis (22). This Trona contains large amounts of fluoride resulting in dental fluorosis (22).

A study in the state of Andhra Pradesh and northern part of Karnataka, India, reported four percent prevalence of Genu valgum (skeletal fluorosis) among jowar consumers compared to one percent in rice consumers (23). A cross sectional study in Tamil Nadu state, India, observed 50.8% prevalence of dental fluorosis among children in the age group of five to 14 years, when the fluoride level was 1.2 ppm; whereas in another village, with a fluoride level of 1.3 ppm, the prevalence was only 41.1%. The authors attributed this higher proportion of dental fluorosis to dietary habits such as consumption of jowar (24).

Jowar grown in endemic fluoride areas contained significantly higher amount of molybdenum than that was grown in non-fluoride areas in India (25, 26). The studies (27-29) have pointed out that molybdenum interacts with fluoride metabolism thus increasing fluoride retention during excretion, resulting in fluorosis. The jowar-based diet increases the fluoride retention and therefore the effect of fluoride is higher than wheat- or rice-based diets, when fluoride intakes are similar (animal models) (30, 31). These observations suggest that dietary patterns and intake of nutrients can modify the effect of fluoride.

The geographical variation in fluoride toxicity may be due to jowar, a staple food in North Karnataka, which tends to uptake higher amounts of molybdenum from the soil, when compared to other cereals (23).
In an experiment on rats, when fluoride intakes were similar, it was observed that jowar-based diet led to higher fluoride retention in urine and bones when compared to rice and wheat-based diets (32). A 32% increase in fluoride retention was observed in rats receiving 50 ppm molybdenum in their drinking water and low fluoride (0.5 ppm) corn diet, when compared to no molybdenum group (28). This study also demonstrated that molybdenum increases retention of fluoride, even at low fluoride concentration in diet and drinking water.

Retention of fluoride due to jowar and rice based diet was studied in a clinical trial in India. A significantly higher retention of fluoride was found in jowar diet group compared to rice group, when the fluoride intake (10ppm) was similar in both groups for ten days (31).

Jowar (sorghum) millets are a grass of East African origin; a drought-resistant, heat-tolerant member of the grass family. Jowar is commonly used for food, fodder, and for the production of alcoholic beverages. “Bhakri,” a variety of unleavened bread made from sorghum, is the staple diet of 30 to 40% farming communities of Northern Karnataka and Maharashtra in India. Jowar becomes an economically viable crop to grow in these communities (North Karnataka, India) due to extreme drought conditions.

It has also been observed that fluoride promotes accumulation of molybdenum in jowar (33). Molybdenum improves growth of some crops. Use of commercial fertilizers having molybdenum can extend the fluorosis problems to the areas that were unaffected. It also acts as an additional source of molybdenum to diet (34). It is observed that fluoride content of food items like jowar grown in endemic fluorosis areas is high (35-36). However, no clear-cut evidence is available
to explain the mechanism of the interrelationship between jowar consumption and dental fluorosis. The current evidence has postulated that metabolism of fluoride differs among jowar consumers. It may be due to reduction in serum fluoride available for renal filtration or due to enhanced tubular fluoride reabsorption.

The problem of dental fluorosis shares certain features with several other important public health problems. The poor population in the rural areas is widely affected. The reasons for its development are unknown. There is no known treatment once the disease sets in. Hence, it is imperative to prevent fluorosis.
1.2 **INTRODUCTION TO FLUORIDE**

**History**

(L. and Fr. fluere: flow or flux)

In 1529, Georigius Agricola described the use of fluorspar as a flux, and in 1670, Schwandhard found that glass was etched when exposed to fluorspar with acid. Scheele and many later investigators, including Davy, Gay-Lussac, Lavoisier, and Thenard, experimented with hydrofluoric acid with some of the experiments ended tragically. Fluoride was finally isolated in 1866 by Moissan after nearly 74 years of continuous effort and was awarded Noble Prize in 1931 (37).

**Properties**

Fluorine is the most electronegative and reactive of all non metal elements. It is a pale yellow, corrosive gas, which reacts with most organic and inorganic substances. Finely divided metals, glass, ceramics and even water, reacts with fluorine in bright flame. Until World War II, there was no commercial production of elemental fluorine. The nuclear bomb project and nuclear energy applications, however, made it necessary to produce large quantities of fluoride (38).

**Uses**

Fluorine and its compounds are used in producing uranium (from the hexafluoride) and more than 100 commercial fluorochemicals. Fluorochlorohydrocarbons are extensively used in air conditioning and refrigeration. Elemental fluorine has been studied as a rocket propellant as it has an exceptionally high specific impulse value.
Table 1.1  Physical and Chemical Properties of Fluorine.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Fluorine</td>
</tr>
<tr>
<td>Discoverer</td>
<td>Henri Moissan</td>
</tr>
<tr>
<td>Group</td>
<td>17</td>
</tr>
<tr>
<td>Series</td>
<td>Halogens</td>
</tr>
<tr>
<td>Discovery Location</td>
<td>Paris, France</td>
</tr>
<tr>
<td>Discovery Year</td>
<td>1886</td>
</tr>
<tr>
<td>Atomic Volume</td>
<td>17.1 cm³/mol</td>
</tr>
<tr>
<td>Relative Gas Density (Air=1)</td>
<td>1.31</td>
</tr>
<tr>
<td>Number of Electrons (with no charge)</td>
<td>9</td>
</tr>
<tr>
<td>Oxidation States</td>
<td>-1</td>
</tr>
<tr>
<td>Number of Neutrons (most common/stable nuclide)</td>
<td>10</td>
</tr>
<tr>
<td>Number of Protons</td>
<td>9</td>
</tr>
<tr>
<td>Flammability Class</td>
<td>Non-flammable gas (extreme oxidizer)</td>
</tr>
<tr>
<td>Atomic Mass Average</td>
<td>18.99</td>
</tr>
<tr>
<td>Melting Point</td>
<td>53.63K -219.52°C -363.14°F</td>
</tr>
<tr>
<td>Greenish-yellow gas of the Halogen family</td>
<td></td>
</tr>
<tr>
<td>Optical Refractive Index</td>
<td>1.000195</td>
</tr>
<tr>
<td>Physical State (at 20°C &amp; 1 atm)</td>
<td>Gas</td>
</tr>
</tbody>
</table>

Metabolism of fluoride

Biological effects of fluoride intoxication are related to the total amount of fluoride ingested whatever be the source, food, water or air.

Sources of fluoride

Food

Nearly all food contains small quantities of fluoride. The total daily intake of fluoride in any average human diet is small, except in endemic regions. In certain endemic regions of India, the fluoride content of vegetables and food may be very high. The contribution of food to the total daily intake of fluoride varies from region to region. Staple diets rich in jowar, Ragi or Bajra containing high fluoride aggravate fluoride toxicity in some endemic areas of India (37, 38).
Water and Beverages

The variation in the fluoride content in natural water depends on factors such as the source of water, type of geological formation and the amount of rainfall. Surface water generally has low fluoride, while ground water has high concentrations of fluoride. The highest fluoride concentration of 28.9 ppm was reported from India. The fluoride content of seawater varies from 0.8 to 1.4 ppm, which explains why the fluoride content of diet rises, on seafood consumption. Among the beverages, tea has an exceptionally high fluoride content which varies in different brands from 122-260 ppm (37-38). The fluoride intake depends upon consumption of drinking water and beverages, which is further determined by factors such as body size, physical activity, food habits and variations in atmospheric temperature and humidity (40-42). In tropical countries like India, the daily fluoride intake is very high due to ingestion of ground water by farm laborers and therefore is at risk of developing fluorosis.

Air

The atmosphere has very low fluoride content and in 97% of non-urban areas fluoride is hardly detectable. The fluoride content of atmosphere is more in areas where there are volcanic eruptions or industrial activity or aluminum mining.

Total daily fluoride intake

In majority of endemic areas around the world, the main contribution is from water and only in few areas of India and China significant amounts come from food. The estimated range of safe and adequate intake of fluorides for adults is 1.5
to 4.0 mg per day and it is less for children and those with renal disease. The daily
intake of fluoride in endemic regions varies from 10 to 35 mg and can be even
higher in summer months (37, 39).

Absorption of fluorides

Soluble inorganic fluorides ingested through water and food is almost completely absorbed. But absorption of less soluble inorganic and organic fluorides varies from 60-80%. Fluorides are absorbed from the gastro-intestinal tract by a process of simple diffusion without any mechanism of active transport (37, 39). It has been noticed that when salts of calcium, magnesium and aluminum are added to diet, reduces the quantum of fluoride absorption on by the formation of their less soluble compounds (43-44). The increased fluoride toxicity has been observed on addition of substances like phosphates, sulphates and molybdenum to the diet.

Distribution of fluorides

About 96-99% of the fluoride is retained in the mineralized bones, since fluoride is the most exclusive bone seeking element because of its affinity for calcium phosphate. But it has been noticed that there is no significant retention of fluoride in the body, if very small quantities of fluorides are ingested (45). In fact, there was no discernible retention of fluoride when upto 4-5 mg was ingested daily. But when more than five mg was ingested about half of it appeared to have been retained by the skeleton and rest excreted through urine. This observations show that after absorption from the gut fluoride enters the circulation, where the plasma
fluoride accounts for three-fourths of the total amount of fluoride found in the red blood cells. Fluoride in plasma exists in free ionic and bound forms, the latter, bound to the serum albumin forming about 85% of the total amount of fluoride in plasma. Plasma fluoride in healthy individuals in non-fluoridated areas ranges from 0.14 - 0.19 ppm and is higher in fluorosis affected individuals. The sequestration of fluoride from the skeleton, urinary excretion and the loss sustained through sweat help in regulation of plasma fluoride.

Figure 1.2 Pharmaco-dynamics of fluoride (39)

The fluoride uptake by the skeleton is very rapid, which depends on the vascularity and rate of growth. The uptake of young bones is faster than that of mature bones. The fluoride is incorporated more readily in the active growing and cancellous areas than in the compact regions of bones. It has been observed that skeletal fluoride concentration increases almost proportionately to the amount and the duration of ingestion. It is this increase in the fluoride content of skeleton that provides the most reliable clue to excessive fluoride intake. The other indicators
such as urine and soft tissue levels, which manifest wide fluctuations, cannot be relied upon (39).

**Excretion of fluorides**

**Feces**

Fluoride present in feces results from two sources, the ingested fluoride that is not absorbed and the absorbed fluoride that is excreted into the gastrointestinal tract. About 10-25% of daily intake of fluoride is excreted in the feces.

**Urinary route**

The elimination of absorbed fluoride occurs almost exclusively via the kidneys. Urinary fluoride in healthy individuals fluctuates widely between zero and 1.2 ppm (average 0.4 ppm), when fluoride content of drinking water is 0.3 ppm. Urinary levels of fluoride are higher in individuals exposed to higher intake of fluoride. The renal clearance of fluoride is directly related to urinary pH. In alkaline urine the fluoride is present in ionic form and hence its renal clearance is rapid. In the acidic urine on the other hand, fluoride is present in nonionic form (Hydrofluoric acid-HF) and hence it is rapidly reabsorbed in renal tubules.
1.3 INTRODUCTION TO JOWAR

Sorghum popularly known as jowar is the most important food and fodder crop of dry land agriculture. In India, sorghum grain is consumed by human either by breaking the grain and cooking it in the same way as rice or by grinding it into flour and preparing ‘chapatis’ (Indian Bread). To some extent it is also consumed as parched and popped grain. This grain is also fed to cattle, poultry and swine. Sorghum grain contains about 10-12% protein, 3% fat and 70% carbohydrates; therefore, it can satisfactorily replace other grains in the feeding programme for dairy cattle, poultry and swine (46).

<table>
<thead>
<tr>
<th>Scientific Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plantae</td>
</tr>
<tr>
<td>Angiosperms</td>
</tr>
<tr>
<td>Monocots</td>
</tr>
<tr>
<td>Commelinids</td>
</tr>
<tr>
<td>Order: Poales</td>
</tr>
<tr>
<td>Family: Poaceae</td>
</tr>
<tr>
<td>Genus: Sorghum L.</td>
</tr>
<tr>
<td>Species: About 30 species,</td>
</tr>
</tbody>
</table>

Table 1.2 Scientific classification of sorghum (46)  Figure 1.3 Jowar plant

Sorghum belongs to the tribe Andropogonae of the grass family Poaceae. Sorghum is a genus of numerous species of grasses, some of which are raised for grain and many of which are used as fodder plants either cultivated, or as part of pasture. Sorghum is known under a variety of names: Great millet and Guinea corn in West Africa, Kafir Corn in South Africa, Dura in Sudan, Mtama in eastern Africa, Jowar in India and Kaoliang in China. In the United States it is usually referred to as Milo or Milo-maize (46).
Chapter 1

Introduction

History

Sorghum, being a tropical crop, has its history related to the hot and humid areas of the world. The cereal grain is said to have originated from Ethiopia as a wild grass as early as 8000 years ago. The cereal crop, once adopted and cultivated, spread across the African continent, especially the regions of Egypt and Sudan (47).

Figure 1.4 Sorghum Growing Areas in World (48)

Sorghum marked its entry into the Asian continent in the first millennium when it was brought to India and then it became popular among the other countries of Asia as well. The weather conditions in the continent suited well for its cultivation. The sorghum was distributed in the rest of the world through of slave trade in America. At that time, it was regarded as the food of the poor and the slaves were fed on sorghum (46).

Sorghum and millets have been important staple diet in the semi-arid tropics of Asia and Africa for centuries. These crops are still the principal sources of energy, protein, vitamins and minerals for millions of the poor people in these regions.

Sorghum and millets are grown in harsh environments where other crops yield poorly. They are grown with limited water resources and usually without
application of any fertilizers by a multitude of small-holder farmers in many countries. Jowar are mostly consumed by disadvantaged groups, they are often referred to as "coarse grain" or "poor people's crops". They are not usually traded in the international markets or even in local markets in many countries.

**Production of sorghum in India**

India has been able to maintain its position among the top three producers of the crop. The sorghum is produced both as a summer and a winter crop i.e khariff and rabi crops in the country. Indian production hovers around an average of nine million metric tons (49). India accounts for around 20% of the world total area used for the crop production. The major states where this cereal grain is produced are Maharashtra, Karnataka, Andhra Pradesh, Madhya Pradesh, Gujarat, Rajasthan, Uttar Pradesh and Tamil Nadu (49).

**Plant Characteristics**

It erects with much variability in growth characteristics. Depending on variety and growing conditions, the plant grows 0.6-0.5 m tall, five to over 30 mm in diameter. Its leaves are broad and coarse, similar in shape to those of corn but shorter and wider.

Sorghum plants can tolerate high temperatures throughout their life cycle better than any other cereal crop. It can tolerate drought conditions very well. It remains dormant under moisture but
resumes growth when favorable conditions reappear. The optimum temperature for growth is 30°C and it needs about 250-400 mm rainfall. It is rich in carbohydrate, and B complex vitamins. It is poor in vitamin A and rich in dietary fiber. Jowar is richer in protein compared to rice, but the quality is not as good as rice protein.

**Plant Varieties**

Varieties of sorghum are classified into 4 groups: *grain sorghums, grass sorghums, sweet sorghums, and broom corn*. Broom corn is grown from the branches of the seed cluster. Sweet sorghums have sweet juicy stems and are grown to be made into sorghum syrup (50). The syrup is made by pressing the juice out of the stems and boiling it down to the proper thickness. Sweet sorghums can also be made into animal feed or silage. Grass sorghums are grown for green feed and hay but can also be weeds.
1.4 Case control design

Case control studies are the first approach to test causal hypothesis and now it has evolved as a permanent method of epidemiological investigation. The case control method is analytical epidemiological study in which individuals who have the disease called cases are compared to individuals who are free of disease called controls. Exposures differences between the cases and controls are used to find potential risk or protective factors (51). It has three distinct features

a. Both exposure and outcome have occurred before the study
b. The study proceeds backwards from effect to cause and
c. It uses control group to support or refute an inference.

History of case control studies

The case control study proceeds from effect to cause, attempting to identify antecedents that led to the condition or the disease. The explicit use of controls appears to be of recent origin and was particularly emphasized in the experimental method formalized in the 18th and 19th centuries (Sartwell 1974). An early report of a case control approach was from Broders in 1920 concerning squamous cell epithelioma of the lip. He compared 537 cases to 500 controls and found that, percentage of tobacco use was markedly similar among cases and controls (51). However he pointed out that there is a remarkable difference in the method of smoking and suggested that pipe smoking played a role in the development of cancer of the lip. The modern case control studies dates from Lane Claypon in 1926 where his paper on reproductive factors in relation to breast cancer has been reported (51).
A landmark in the analysis of case control studies is the paper published by Cornfield (1951) which demonstrated that the relative risk can be estimated from either a case control or a cohort design. Mantel and Haenszel (1959), later showed how to efficiently estimate the relative risk from stratified data, and gave chi-square tests for association (51).

**Basic Steps in the design**

- Selection of cases and controls
- Matching
- Measurement of Exposure
- Analysis and interpretation (52).
Method of selecting cases and controls

After defining the eligibility criteria for cases and controls and specifying the sources from which they will be drawn, the next step is the unbiased selection of eligible cases and controls. A procedure is followed for the selection of a sample in a manner that assures that each individual has an equal chance of appearing in the study (51). This identification of subjects does not depend on the individual’s exposure status in regard to the study factors. All eligible cases arise with a defined period of time. When the case arises from a hospital, it is obvious that sampling is from a subset of the total population of cases of interest. Selection of the control group is important as in rare diseases; the number of eligible controls available for study greatly exceeds the number of cases (51).

Sampling is done from a frame, which is a list of potentially eligible cases and controls in the target population. A case control is an investigation of disease–exposure associations with sample population, termed as “target population”. It refers to a subset of general population, that is both at risk of the study exposure(s) and the development of the study disease. Irrespective of exposure, each eligible case and control should ideally have an equal chance of appearing in the study. Thus it avoids over-representation of exposed cases and controls in the study at hand (51).

A variety of sampling procedures exists for selection of cases and controls. Random sampling is a method of selecting individuals from a frame such that each possible sample has a fixed and determinate probability of selection. So it is an “equal probability of selection method” which is called Epsom. In multistage sampling designs, it is Epsom through several stages of selection process (51).
Chapter 1

Introduction

**Matching**

Matching is defined as the pairing of one or more controls to each case on the basis of their “similarity” with respect to selected variables. Matching is carried out for the attribute(s) which is already recognized as a “Confounder”. Matched sampling involves the pairing of one or more controls to each case on the basis of specified variables, the effect of which, one wants to eliminate from the case control comparison (51).

**Confounder**

The problem of confounding exists when the effect of a risk factor \( B \) is studied on the disease \( A \) and another factor \( C \) comes into question, which has an effect on both the disease and the risk factor. A factor can act as a confounder only, if it satisfies the following two criteria.

1. It should be associated with the outcome.
2. It should be associated with the frequency of the exposure factor of interest.

**Alternatives to matching**

There are several alternatives to control for confounding. All these serve the same purpose of assessing the extent of bias reduction or increase in precision expected from a matched design (51).

**Stratified sampling**

This involves the formation of subgroups by partitioning the ranges of specified variables and sampling a predetermined number of cases and a controls within cells created by the multiple cross classification. The stratum specific odds ratio and a pooled odds ratio are calculated (51). A difference between the stratum specific odds ratio shows the presence of interaction, whereas a difference...
between the unadjusted odds ratio and the pooled odds ratio shows the presence of a confounder (51). Stratification is done to assess interaction between two or more potential risk factors. The cases and controls are stratified based on one of the risk factors. To study the interaction between the like risk factors a and b, cases and controls are stratified based on b into two stratum

**Stratum 1 – when risk factor b present**

<table>
<thead>
<tr>
<th>Risk factor a</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A1</td>
<td>B1</td>
</tr>
<tr>
<td>Absent</td>
<td>C1</td>
<td>D1</td>
</tr>
</tbody>
</table>

Stratum specific odds ratio \( S_1 = \frac{A1D1}{B1C1} \)

**Stratum 2 - when risk factor b absent**

<table>
<thead>
<tr>
<th>Risk factor a</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A2</td>
<td>B2</td>
</tr>
<tr>
<td>Absent</td>
<td>C2</td>
<td>D2</td>
</tr>
</tbody>
</table>

Stratum specific odds ratio \( S_2 = \frac{A2D2}{B2C2} \)

When there is a difference between \( S_1 \) and \( S_2 \), then it indicates interaction between the risk factors. Then the question arises whether persons exposed to both factors have a risk for the disease than that would be expected of their risks from exposure of each factor alone. There is no single characterization of what to expect from a combination of risk factors. There are two alternative models, additive and multiplicative, which are often used to express the concept of interaction. A variable’s effect may be measured in terms of risk difference, risk ratio, difference in log odds, incidence rates etc. here it is odds ratio;
Consider two dichotomous risk factors $x$ and $y$.

$P_\infty =$ incidence rates of in the absence of both factors $x$ and $y$

$P_{10} =$ when $x$ but not $y$ is present

$P_{01} =$ when $y$ but not $x$ is present

$P_{11} =$ when both $y$ and $x$ are present

In an additive model for disease risk

If the joint effect of $x$ and $y$ equals the sum of their individual effects,

$\left( P_{11} - P_\infty \right) = \left( P_{10} - P_\infty \right) + \left( P_{01} - P_\infty \right)$

then there is no interaction.

If the joint effects exceed the sum of individual effects then it is called synergism. If the joint effect is less than the sum of individual effects it is called Antagonism. Synergism and antagonism are special cases of positive and negative interactions.

A lack of interaction can be expressed in terms of excess relative risk.

$\left( R_{xy} - 1 \right) = \left( R_x - 1 \right) + \left( R_y - 1 \right)$ where $R$ is the relative risk when

$R_{xy} = \frac{P_{11}}{P_{00}}$

$R_x = \frac{P_{10}}{P_{00}}$

$R_y = \frac{P_{01}}{P_{00}}$

In a multiplicative model for disease risk:

If $R_{xy} = R_x R_y$ then there is no interaction. If not there is either synergism or antagonism as the case may be.

If there is interaction then the interactive term or the total risk if the two factors are present is; $R = R_{xy} R_x R_y$
**Frequency matching**

This is a variation of stratified sampling. Here the selections of cases are done at random and controls will be taken from corresponding subgroups in proportion to the number of cases. The expected size of case series in each subgroup is proportional to its population size, out of which one may deliberately oversample certain subgroups to allow their better representation in their study. An estimate of overall risk is derived by appropriately weighting the stratum specific estimates by the sampling fractions (51).

**Post stratification**

The use of above techniques needs to identify variables to control before investigations. Post stratification involves the classification of unmatched samples of cases and controls on the basis of their values on one or more variables ascertained. Therefore it is similar to stratified analysis except that variables used for stratification need not be ascertained before start of the study (51).

However, there are two limitations. *First*, there is loss of information resulting from subgroups in which only cases and controls occur. So a subgroup comparison cannot be made. This post stratification relies on overlap in the distribution of cases and controls. *Second*, the number of variables on which one can simultaneously stratify and the fineness of the classification are restricted by the number of cases and controls.

**Regression analysis**

The goal of regression analysis is to describe the mean (or expected value) of Y as a logistics function of a set of predictor variables $x_1, x_2, x_3, \ldots, x_n$

This involves fitting the logistic regression model.
\[ Y = b_0 + b_1 E + b_2 X \]

E is the study exposure, X is the confounding variable to be controlled, and 
\[ y = \ln \frac{p}{q} \] is the logarithm of odds of disease (the log odds of a case) \((51)\).

Among exposed individuals \((E=1)\) the log odds of disease is \(Y_1 = b_0 + b_1 + b_2 X\)

Among unexposed individuals \((E=0)\) the log odds of disease is \(Y_0 = b_0 + b_2 X\)

Thus, the log odds of disease has same dependence on \(X\) \((b_2\) is constant) among exposed and unexposed. The difference in the log odds for individuals with the same value of \(X\) \((y_1 - y_0) = b_1\). Thus, the effect of exposure on the log risk of disease, adjusted for the linear effect of \(X\), is represented by \(b_1\). In fact \(\exp (y_1 - y_0) = \exp (b_1)\) is the disease exposure odds ratio adjusted for the long linear effect of \(X\).

**Conditional logistic regression** is to describe the association between a dichotomous or binary dependent variable and independent variables. The curve is “\(S\)” shaped and the values can never be less than 0 or more than 1. In a case control study, the dependent variable is binary and indicates whether it is a case or a control. The coefficient of independent variable representing the exposure can be interpreted as the odds ratio. The advantage of using conditional logistic regression rather than just (point estimation) estimating the odds ratio from the sample, is that it takes into consideration of all the other selected independent variables (confounders), which goes through various processes of iterations and gives unconfounded odds ratio for each variable. This fundamental advantage is the ability to control conveniently for other factors that are measured, but not for that is matched. If the matching has introduced a material correlation in the exposure histories between cases and controls, conditional logistic regression analysis allows both the removal of confounding introduced by matching and the
control of additional unmatched confounding factors. In this situation, analyses and interpretations based on regression model should be considered (51).

**Bias in Case control Studies**

A bias is a systematic error in the design of a study that results in erroneous conclusions. There are three broad categories of biases (51, 52).

**Selection bias**

Case control studies are vulnerable to misleading associations arising from circumstances in which cases and controls are selected and are ascertained.

**Surveillance bias**

If the condition under study is mild or asymptomatic, it is liable, that cases are more likely to be detected in persons under frequent medical surveillance. When cases are from hospitals or clinic and the controls from a community, surveillance bias can be encountered. Use of control from a hospital when the cases are from the hospital will minimize such a bias (53).

**Diagnostic bias**

This type of bias occurs when probability of a case being diagnosed is higher, if the individual is exposed. In a study designed to assess the correlation between oral contraceptive usage and deep vein thrombosis, a group of female patients, clinically diagnosed to have deep vein thrombosis were compared with a group of apparently healthy women. Later, it was noted that when a woman complains of pain in her legs, her physician is more likely to make a diagnosis of deep vein thrombosis if she is a pill user. Naturally, exposure factor is going to be more frequently encountered among the cases (53).
Incidence and prevalence bias

If the exposure factor affect the duration of the disease, choosing prevalent cases results in a biased study. Suppose smokers who develop myocardial infarction (MI) are more likely to die quickly, then a case control study using surviving cases of MI would result in lower estimation of relative risk (53).

Non response bias

Here the cases and controls refuse participation in the study. The assessment of the impact of non response is calculated. If the non responders do not differ considerably from the rest of the people in the study, then non response does not have a negative impact on the results (53).

Information bias

Recall Bias

In a study to determine the effect of certain drug and congenital abnormalities, the mothers of affected children may be more likely to remember the details of her ANC period than those of unaffected children (53).

Reporting bias

This occurs when one uses surrogate informant for dead cases (53).

Interviewer Bias

Here the interviewer is not masked of the disease status; he or she tends to be more persuasive when questioning cases as compared to questioning controls.

Analytical Issues

- Association vs. Causal relationship.
- Adjustment of confounders.
- Matching.
• A common limitation of the adjustment: cannot account for the effects of the unobserved confounders (51).

The most effective way to proceed in the analysis of case control study is from simple to the complex. The exposure of interest is dichotomized into present or absent among cases and controls. Then the crude disease exposure association is determined by estimating the odds ratio in 2 X 2 table. The odds ratio is the parameter in the analysis of a case control study, since it approximates relative risk and is invariant across cohort and case control studies. Once the study is a matched one, then a matched pair analysis is done. Usually, an unmatched analysis precedes a matched analysis. If controls are matched then doing an unmatched analysis will underestimate the association.

**Unmatched or crude odds ratio estimation**

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>E+</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>a+b = e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>c+d = f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>a+c = n1</td>
<td>b+d = n2</td>
</tr>
<tr>
<td>a+b+c+d = n</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Unmatched or crude odds ratio, OR, \( \psi = \frac{ad}{bc} \)**

The odds ratio is the maximum likelihood estimate, assuming that the sample sizes for cases and controls are fixed at \( n1 \) and \( n2 \) respectively, and that simple random samples have been taken from theoretically infinite populations of cases and controls or that random samples with replacements have been taken from finite population. The sample odds ratio is also a maximum likelihood estimate under the assumption that \( a, b, c \) and \( d \) follow multinomial distribution with total sample size fixed at \( n \).
Estimation of risk

In epidemiology, the association between a risk factor or protective factor (exposure) and a disease may be evaluated by the “risk ratio” (RR) or the “odds ratio” (OR). Both are measures of “relative risk”—the general concept of comparing disease risks in exposed versus unexposed individuals. An odd is always higher than its corresponding probability, unless the probability is 100% (51).

The odds ratio (OR) is a measure of the relative magnitude of the odds of exposure among individuals who have the disease (cases) and the odds of exposure among individuals who do not have the disease (controls).

Odds ratio is odds of exposure among cases divided by odds of exposure among controls. Odds ratio is calculated by dividing odds of exposure among cases (a/c) by odds of exposure among controls (b/d) numerically the same as dividing the products obtained when multiplying diagonally across the 2x2 table (ad/bc) known as “cross-products ratio”

Validity of using odds ratio as a measure of Association

The usual calculation of amount of risk is by means of relative risk. Relative risk is a risk ratio. It is ratio of two incidence rates.

Relative Risk, \( R = \frac{\text{Incidence of the disease among exposed}}{\text{Incidence of the disease among non exposed}} \)

A case control study is best understood by applying principles from follow up studies. Relative risk is calculated from a follow up or a cohort study where in incidence rates can be calculated.

From the above table relative risk is \( R = \frac{(a/a+b)}{(c/c+d)} \)
Where \( a \) is the number of people with the disease among the exposed population of \((a+b)\) and \( c \) is the number of people with the disease among the unexposed population of \((c+d)\).

When the disease under investigation is rare,

\[
\text{then } a \text{ and } b, \quad (a+b) \approx b \\
(c+d) \approx d
\]

Then relative risk, \( R = (a/b) / (c/d) = ad /bc \)

Odds is the ratio of two probabilities

\[
\text{Odds} = \frac{\text{Probability of an event to occur}}{\text{Probability of an event not to occur}} = \frac{P}{1-P}
\]

Odds ratio is the ratio of two odds. For rare diseases, the risk of disease, \( p \) and the odds ratio, \( p/(1-P) \) are virtually identical.

Disease odds ratio = \( \frac{\text{odds for disease when exposed}}{\text{odds for disease when non exposed}} = \frac{a/b}{c/d} = \frac{ad}{bc} \)

So, the disease OR \( \approx \) Relative risk (RR)

In a case control study, the groups studied are based on the presence or absence of disease rather than exposure. So in a case control study

\( \frac{ad}{bc} = \frac{a}{c} \div \frac{b}{d} \)

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Exposure</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Hence, if we can obtain odds for exposure among cases and controls, the calculated odds ratio is equal to disease odds ratio.

\[
\text{Exposure odds ratio} = \frac{\text{Odds for exposure among cases}}{\text{Odds for exposure among controls}} = \frac{a/c}{b/d} = \frac{ad}{cb}
\]
In the case control study, what we calculate is the exposure odds ratio. But the exposure odds ratio = disease odds ratio.

A relative risk RR is approximately equal to the disease odds ratio when the disease under study is rare. The ratio calculated is equal to the relative risk, as the concept of "exposure opportunity" is irrelevant, if the setting is translated to that of a follow up study. This can be explained by using the probability theory and Bayes theorem.

\[
\text{Relative Risk, } R = \frac{P(D/E+)}{P(D/E-)} \quad \text{P = probability}
\]
\[
D = \text{disease} \quad \text{E = exposure}
\]

Using Baye's Theorem

\[
P(D/E+) = \frac{P(E+/D) \cdot (D)}{P(E-)}
\]
\[
P(D/E-) = \frac{P(E/-D) \cdot (D)}{P(E-)}
\]

So Relative risk, \( R = \frac{P(E+/D)}{P(E-/D)} \cdot \frac{P(E+)}{P(E-)} \)

Relative risk, \( R = \frac{\text{Odds for exposure among the diseased}}{\text{Odds for exposure among the population}} \)

As disease OR is numerically equal to the Exposure OR, the Exposure OR estimated will be equal to the relative risk.

**Interval estimate for unmatched odds ratio**

Point estimate involves the calculation of the odds ratios. A point estimate is an indicator of the extent of association or the magnitude of effect in the data. To incorporate into the estimation process, an allowance for random variability in the observations, it is preferably to estimate the effect using a range of values for the parameter that is consistent with the observed data with specified limits. The range...
of confidence interval and the process of calculating is called interval estimate (51).

The width of confidence depends not only on the amount of variability in the data, but also on an arbitrary value that specifies the degree of consistency between the limits of the interval and the data. This arbitrary value is the level of confidence, which is usually expressed as a percentage.

\[ 95\% \text{ CI} = \text{point estimate} \pm 1.96 \times (\text{standard error}) \]

95% confidence limits for crude odds ratio can be calculated using “Woolf Method”

The standard error is calculated for the \( \ln(\psi) \).

\[ Z_{\alpha} = 1.96 \]

\[ \text{SE}(\ln(\psi)) = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{x_1} + \frac{1}{x_2}\right)} \]

\[ 95\% \text{ C.I}(\psi) = \exp\{\ln(\psi) \pm 1.96 \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{x_1} + \frac{1}{x_2}\right)}\} \]

After the calculation of unadjusted odds ratio, the stratification analysis is carried out to identify interaction or confounding effect or effect modification.

**Advantages of case control design**

1. Quick, easy and cost effective
2. Most efficient design for rare diseases
3. It usually requires a smaller study population than cohort study

**Disadvantage of case control design**

1. It is uncertain on exposure-disease time relationship.
2. Inability to provide the direct risk estimation.
3. Not efficient for studying rare exposures.
4. Subject to biases (recall & selection bias).
1.5 \textbf{NUTRITIONAL EPIDEMIOLOGY}

\textbf{Introduction}

Nutritional epidemiology is the study of the nutritional determinants of disease in human populations. This new sub-discipline comes from the direct relevance to crucial health problems such as coronary heart disease, cancer, stroke, osteoporosis, diabetes etc. All of these diseases have been the object of research in nutritional epidemiology and the findings have already been applied in ways that may improve public health.

Nutritional epidemiology has a direct relevance to human health, and it does not need to extrapolate from animal models or in vitro systems. The results are often used to estimate risks, which can be translated into specific recommendations for changes in nutrient intake or food recommendations and consumption patterns. The complex nature of diet and methodological bias in designs may sometimes limit the use of nutritional epidemiology.

\textbf{Characteristics of nutritional epidemiology}

1. Multiple causes - Dietary, genetic, occupational, psychosocial, infectious factors, levels of physical activity.

2. Long latent periods - cumulative exposure over many years or relatively short exposure occurring many years before diagnosis.

3. It occurs with relatively low frequency despite a substantial cumulative lifetime risk.

4. The complex nature of diet has posed an unusual challenge to nutritional epidemiology (54).
Diet represents a complex set of exposures that are strongly interrelated and cannot be characterized as present or absent.

Continuous variables often with a rather limited range of variation.

Individuals rarely make clear changes in their diet at identifiable points in time; typical consumption patterns evolve over a period of years.

Individuals are generally not aware of the content of the food that they consumed ⇒ consumption of nutrients is usually determined indirectly based on the reported use of food or on the level of biochemical measurements.

Small relative risk (1.5 ~ 3.0) and large population attributable risk.

**Epidemiologic approaches to diet and disease**

1. Sources of the concepts, hypotheses, and techniques of nutritional epidemiology (54).
   a. Biochemistry
   b. Cell culture methods
   c. Experiments in laboratory animals
   d. Metabolic and biochemical studies among human subjects.

2. The findings from in vitro studies and animal experiments sometimes cannot be extrapolated directly to human studies.

3. The basic science areas provide critical direction for information that can aid in the interpretation of the epidemiological findings.
Ecologic and Correlation studies

It is comparison of disease rates in populations, with the population per capita consumption of specific dietary factors. Usually, the dietary information is based on disappearance data (54).

Food produced and imported minus the food exported and fed to animals

Strengths of international correlation studies

➢ Contrasts in dietary intake are typically very large.

➢ The average diets for people residing in a country are likely to be more stable over a period of time, than are the diets of individuals within the country.

➢ The cancer rates on which international studies are based are usually derived from relatively large populations and are subject to only small random errors.

Problems of co-relational studies

➢ Many potential confounders: genetic predisposition, other dietary factors environmental or lifestyle practices cannot be controlled.

➢ Limited by the use of food “disappearance” data.

➢ It is cannot be independently reproduced.

Migrant studies and Secular Trends

Strengths

➢ Studies the genetic factors.

➢ It is useful to examine the latency or the relevant time of exposure
The changes in the rates of a disease within a population over time, provide evidence that non-genetic factors (environmental factors) play an important role in the cause of certain diseases (54).

**Case-Control study**

**Strengths**

- The confounding effects of other factors can be controlled in the design (matching or restriction) and in the analysis (multivariate methods).
- The dietary information can be obtained from the individuals more efficiently and rapidly than cohort studies because the number of subjects is typically smaller and no follow-up is necessary (54).

**Problems of case-control studies**

- Potential methodological bias: limited range of variation in diet, inevitable error in measuring intake ⇒ relative risks in most studies of diet and disease which are likely to be modest (0.5-3.0).
- The relative risks are usually based on small differences for cases and controls of only about 5% ⇒ a systematic error of even 3% or 4% can seriously distort such a relationship.

**Cohort studies**

**Strengths of prospective cohort studies**

- It avoids most of the potential sources of methodological bias associated with the case-control studies.
- The dietary information is collected before the diagnosis of disease; illness cannot affect the recall of diet.
The distribution of dietary factors in the study population may be affected by selective participation in the cohort.

It provides an opportunity to obtain repeated assessments of diet over time and to examine the effects of diet on a wide variety of diseases.

Limitations of prospective cohort studies

- The loss to follow-up varies by level of dietary factors that can result in distorted associations.
- The necessity to enroll large population of subjects for the diseases of relatively low frequency, even very large cohorts will not accumulate a sufficient number of cases within a reasonable amount of time.

Randomized controlled trails

Most rigorous evaluation of a dietary hypothesis, optimally conducted as a randomized, doubled-blind trial (54).

Strengths

- Minimizes the possibility of confounding by randomization.
- It is good for evaluating hypotheses, that has minor components of the diet (trace elements or vitamins) since these nutrients can be formulated into pills or capsule.

Limitations

- It produces an imprecise measure of the effect of exposure due to marginally adequate sample sizes.
- It is impossible to conduct due to practical or ethical reasons (e.g. smoking and lung cancer, alcohol use and breast cancer risk).
Causal criteria in nutritional epidemiology

Consistency

Evaluation of consistency in nutritional epidemiology is a challenge. Nutritional studies often have null findings for a variety of reasons including measurement error, lack of variation of intake in the population, or a distribution of intakes unrelated to the disease process. Careful evaluation of inconsistencies between positive and null studies can be informative (55).

An assessment of the level of intake needed for an effect to be observed across studies is difficult. Food-frequency questionnaires are adequate for comparisons within a study, but are not accurate in terms of absolute nutrient values. Thus, cutoff points in one study usually cannot be compared with cutoff points in another study. In some circumstances, a lack of consistency across vastly different study populations may provide insights rather than suggesting a lack of effect.

For example, cutoff points for fruit intake in a null study of invasive cervical cancer in the United States was 7.3 and ≥19 servings/wk for the lowest and highest quartiles, respectively (56), whereas in a Latin American study they were 4.3 and ≥30 times/wk, respectively, which showed protective effects (57). This example suggests that, although the findings were inconsistent between studies, very high fruit intakes may be necessary to observe an effect on cervical cancer. Extrapolation of absolute values for cutoff point intakes is not usually possible, unless a study has validation data that can be used to estimate what the true cutoff points might be. Stated in another way, the presence of consistency across studies
with different cut off points is hampered unless information on true (i.e., absolute) values of intake is available.

**Strength of association**

Nutritional epidemiology is fraught with evidence of weak associations. It is far more common to find risk estimates of 0.8–3.0 than to find a 2-fold (much less a 4-fold) estimate of risk. Indeed, strong risk estimates (≥4) arising in a single study is so uncommon that they may be viewed as the result of bias, if they are not reproduced in other, similarly designed studies. Generally, weak associations are also viewed with caution because they too can often be explained by bias (55).

Indeed, the criterion for strength of association is more likely to be problematic in nutritional studies, as a result of the frequent occurrence of measurement error. Although this fact could be used to claim that small risks are likely to be underestimated and therefore are stronger than observed. Weak associations in dietary studies may have large public health effects, if the dietary factor is common and the disease presents an important public health concern. It may be reasonable for this threshold to vary according to a prior hypotheses and that depends upon whether the exposure consists of a food group, a diet-derived nutrient, or a blood marker.

**Dose response**

General view of the criterion of dose response is that, the presence of a statistically significant linear or otherwise regularly increasing trend clearly reinforces the evidence in favor of causality. However, such an ideal situation may not be achieved easily, when dietary data are being evaluated. In nutritional epidemiology, it is often the case that only the extreme categories of exposure are
related to risk. Under these circumstances, a test for trend may be significant. Although such a finding may represent a statistical artifact, it also does not preclude a trend nor does it preclude the possibility of a threshold effect. Nevertheless, studies reveal no obvious linear trend (e.g., relative risk estimates from lowest to highest quintile of 1.0, 1.2, 0.8, 1.0, and 1.4) but also show that a statistically significant trend test should be viewed with caution (55).

**Biological plausibility**

Biological plausibility is one of the most challenging and promising of all the causal criteria. In nutritional sciences, the biological evidence is collected from animal models, in vitro cell systems, and human metabolic and clinical studies. The relevance of each type of evidence is controversial. Decisions on usefulness of the findings tend to be rather subjective, as is frequently the case in causal inference. The incorporation of genetic and other biological markers as exposures in epidemiological studies, suggests that biological plausibility will become more important to causal inference in the future (58).

In situations in which a prior hypothesis of nutrient-disease association is linked with a known (i.e., established) biological mechanism, evidence of the association in an epidemiological study can be called biologically plausible. However, evidence of associations that were not anticipated, often emerges from epidemiological studies. Given the number of nutrients and food groups evaluated in most nutritional studies, some new findings may be due to chance. Furthermore, it may not be difficult to contemplate biologically relevant functions of nutrients, but post hoc justifications should not hold the same evidential status as a prior hypothesis. Since there are strong precedents in epidemiology and nutrition for
making public health recommendations without evidence of biologically plausible mechanisms, it is suggested that it may be reasonable to continue with this practice (55).

We must be prudent to recommend an increase or decrease in the consumption of food (in contrast to single nutrients) without understanding the biological mechanism. In these situations, a balance must be sought between the known benefits and harms (at a population level) of the foods and the uncertainty added because of the unknown effects of other nutrients found in these same foods. When biological evidence is unavailable and yet recommendations based on epidemiologic evidence seem prudent, it is wise to make clear that some recommendations are more tentative than others.

**Temporality**

For the purpose of reaching causal conclusions, it is necessary for the exposure assessment to precede the onset of disease. For the purpose of making public health recommendations, it is desirable to determine the extent to which dietary factors may influence either, onset or progression of disease. When criteria are used in causal inference for the purpose of making nutritional recommendations, no single criterion (of these 5: consistency, strength of association, dose response, biological plausibility, and temporality) is absolutely critical or irrelevant. Sometimes, the evidence from epidemiological studies may appear to be in conflict with the set of causal criterion. For such a criterion, nutritional recommendations are not necessarily precluded (55).
Interpretation of data

The knowledge that an association exists is not sufficient to make public or personal decisions. Such actions require some knowledge of the shape and quantitative aspects of the dose-response relationship. The knowledge of the approximate latent period between alteration in diet and change in disease incidence would be important (54).

Interpretation of null association:

a. Variation in diet is insufficient.

b. Variation may exist for the study population, but only within a “flat” portion of the total dose-response relationship.

c. The method of measuring dietary intake is not sufficiently precise to measure differences that truly exist.

d. The low statistical power, due to inadequate number of diseased and non-diseased subjects.

e. The temporal relationship between the measure of exposure and the occurrence of disease do not encompass the true latent period.

f. The unmeasured third variable was related to exposure and disease in opposite directions (negative confounding variables).

Limitations of the null finding

- The confidence intervals provide a sense of the range of values that are still consistent with the data and should be adjusted for measurement error.

Multivariate relationships of diet and disease

➢ Types and amount of food consumed may be related to important non-dietary determinants of disease (e.g. age, smoking, exercise, and occupation), which may confound or modify relationships with diet.
Intakes of specific nutrients tend to be interrelated so that associations with one nutrient may be confounded by other aspects of the diet.

Intake of one nutrient may modify the absorption, metabolism, or requirement for another nutrient, thus creating a biological interaction (54).

The inherent limits of epidemiology

It has become increasingly evident, that the science of epidemiology has inherent limits. Although epidemiology is very effective in identifying strong links between an environmental factor and a disease (for example, the link between smoking and lung cancer), it is less effective in discerning weaker associations (59). As Dr. Michael Thun of the American Cancer Society stated in the journal "Science" "With epidemiology you can tell a little thing from a big thing. What's very hard to do is to tell a little thing from nothing at all." Many of the associations between diet and disease are "little things": if the effects are real, they are relatively subtle. It may be impossible to determine, from epidemiology alone, whether relatively weak associations between diet and disease are real or whether they reflect some type of subtle bias or measurement error that the researchers were unable to eliminate (59).
1.6 Study Area Introduction

Davangere is a city in the Indian state of Karnataka. The city of Davangere is located on NH4 (National Highway) at a distance of about 260 km from the state capital of Bangalore and is geographically located at 14.31 N, 75.58 E (Latitude & Longitude). The sex ratio in the district is 952 women to 1,000 men. The literacy rate in the district is 67.4%. Previously known for its cotton mills, it is a developing city of Karnataka.

Davangere District lies in the median region on the Deccan plateau. There are six taluks in this district viz. Harihara, Honnali, Harapanahalli, Jaguar Davangere and Channagere. The climate of Davangere is marked by hot summers, low rainfall and a pleasant monsoon season. The summer starts early in the month of March and lasts till the beginning of June when the district comes under the influence of southwest monsoon which lasts till end of October. The average temperature is around 37°C, varies from 35°C to 39°C.
Food habits

Main staple food is Jowar, Rice and Ragi (Millet). Davangere has a mixed style diet of Northern and Southern parts of state. This variation in dietary habit was suitable for conducting the study. Both Jowar Roti which is common in North Karnataka and Mudde (Ragi, Jowar balls) which is common in South Karnataka are consumed in Davangere (60).

Occupation

Seventy percent of people are involved in agriculture, mainly Paddy, Maize, Jowar and Sugarcane. General living and socioeconomic conditions are comparable in all selected villages (61).

Source of drinking water

All selected villages have drinking water supply from tap water or tube wells coming under schemes of Mini water supplies or Accelerated rural water supply. In this scheme, water from the primary source such as bore well is pumped into storage tanks (Reinforced Cement Concrete or Masonry tanks) erected at one or more centrally located places of the village, and the water is supplied through taps provided at tanks. But in some villages had NRWS (National Rural Water Supply Scheme) where water is distributed to the consumers through a network of pipes and taps (61). The water supply remained constant and these bore wells were around 8 to 12 years old.