METHODOLOGY
3. METHODOLOGY

Research is a process and a means to acquire knowledge about any natural or human phenomena. Rapid changes of modern times are causes as well as effects of new discoveries, investigations and findings in various walks of life.

The definition of scientific research has been given by Fred, N. Kerlinger in his book, “Foundation of Behavioural Research”, says "Scientific research is systematic, controlled, empirical, and critical investigation of hypothetical propositions about the presumed relation among natural phenomena." The primary objective of this chapter is to explain the method which the researcher has chosen to select the sample and the tools and techniques used in the present research. Keeping the objectives of the research in view the methodology is discussed as follows:

3.1 Research Design
3.2 Locale of the study
3.3 Sampling Procedure
3.4 Variables used in the Study
3.5 Study Tools
3.6 Reliability and Validity
3.7 Procedure of Data Collection
3.8 Tabulation of the data
3.9 Methods used for Statistical Analysis
3.1 RESEARCH DESIGN

The research design is the most important and the crucial one in research methodology. In a broad sense research design is the entire process of planning and carrying out the research. To put it in Kerlinger’s (1990) words, “Research design is the plan, structure and strategy of investigation so as to obtain answer to research questions and to control variance”. The Plan is the overall scheme or programme of research. The study entitled “The effect of mental health, lifestyle and nutrient intake on the health of adolescents” is exploratory in nature. Exploratory design is planned to gain more knowledge and familiarity with a phenomenon or the subject concerned to achieve new insight into the problem.

Accordingly, after a thorough and meaningful formulation of the problem, specific objectives were set to provide the basis for the inquiry. In the light of these objectives, the scope of the study was delimited and techniques of investigation to be adopted, tools to be used and the pattern of statistical analysis to be followed were decided.

Further the theme of present study was developed and given a definite shape in the form of outline of the study to evaluate their relevance in the light of the set objectives. An effort was made to obtain a thorough review of the relevant literature relating to the previous research in the field. To provide the scientific basis of the study a proper hypothetical framework was also developed, to provide a definite direction and specific scope to the whole investigation. The needed conceptual clarification about the terms and items used within the framework of the study was also considered necessary for which the help of both literal and operational definitions was taken.
The finding of the study have been discussed in the light of the available research material on the subject and were subsequently summarized throwing light on all the major aspects covered by the study. The result has been pointedly presented under the title of finding and discussion and the possible suggestions on the basis of conclusion of the study are given in later chapter.

**LOCALE OF THE STUDY**

**UTTAR PRADESH:** Uttar Pradesh is the fourth largest state of the country in terms of geographical area. The total area of the state is 2,38,566 sq. km. Uttar Pradesh is between north latitude 24 degrees and 31.5 degrees and east latitude 77 degrees and 84.5 degrees. It is border state along the foot hills of Himalayas and Nepal and Tibet (China) to its north, bordered by Bihar in the East. Himachal pradesh, Haryana and Rajasthan are in the West and Madhya Pradesh in the South.

The British combined Agra in one province called the United Provinces of Agra. The name was shortened to United Provinces in 1935. After independence, the United Province was renamed Uttar Pradesh. The state of Uttaranchal came into being in November 2000.

Uttar Pradesh is the largest state of India in term of population also. According to 2001 census it has a population of 16,60,52,859 representing 16.17% of the country’s population. One out of every six persons in India belongs to Uttar Pradesh. Every 31st person of the world is a resident of U.P. It would have ranked as the seventh largest country in the world, if it was an independent nation according to population. Its population density of 689 persons per square kilometer is far above that
of 324 for entire country. The state of Uttar Pradesh is divided into 14 divisions and 70 districts.

**ETAH:** Etah district is a part of Agra division and situated in south-west part of Uttar Pradesh. Etah is a town which is also the district headquarter. Etah is located on the Grand Trunk Road which connects Delhi to Kanpur. This is the old Sher Shah Suri Marg built during the Mughal period. Awagarh is one of the town areas of Etah, used to be an important princely state before independence.

It is bounded on the west by district Aligarh, Hathras, Mathura and Agra, on the south by Mainpuri and Firozabad and north by Farrukhabad. It lies between the parallels of 27018’ and 28020’ North latitude and 79017’ East longitude. Its area is 4446 square kms. The district comprises of six Tehsils, 15 blocks, 149 Nyay Panchayats and 905 Gram Sabha with 1610 villages.

**Population:** According to 2001 census it has a population of 107098. Males constitute 53% of the population and females 47%. ‘Etah’ has average literacy rate 68% higher than the national average of 59.7% and female literacy is 63%. In ‘Etah’, 14% of the population is under 6 years of age. The primary occupation of the people of the district is Agriculture.

**Climatic Conditions:** The climate is drier and hotter than that of the neighbouring districts. The hot weather lasts longer than in the north, usually from March to June. The rainy season is short (July-September); the late start and early withdrawal of the monsoon and long dry spells in between depending on the formation of depression are quite common features. The cold weather lasts from November to January. Rainfall is mostly confined to July, August and September, but in the first half of
<table>
<thead>
<tr>
<th>st Stage</th>
<th>IInd Stage</th>
<th>IIIrd Stage</th>
<th>IVth Stage</th>
<th>Vth Stage</th>
<th>VIth Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>From each house one girl (11-19) and one boy (13-21) were selected</td>
</tr>
<tr>
<td>Urban</td>
<td>Total 25 wards; out of them ward no. 7 and 23 selected</td>
<td>Each ward has 6-8 Mohallas; out of these 2 Mohallas for each ward selected</td>
<td>Each Mohalla has 160-220 houses; out of these houses 50% were selected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETAH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
October rains are not unusual. About 90 percent of the rainfall is received during this period, and the winter rains are uncertain, very little and undependable.

**Rivers:** The two major rivers Yamuna and Ganga passes through the Etah district.

**Crops:** The three crops of Etah district are Rabi, Kharif and Jayad. Rabi crops are sown in the month of October and November and are harvested in the month of March or April. This crop includes wheat, barley, chana and lentil. The Kharif crops are sown in July and are harvested in the month of December. The main products of this crop are millet, maize, bazra, urd, mung and arhar. Where water is available, Jayad crops are grown near rivers, ponds and canals. Main products of this crop are tobacco, pumpkin, watermelon, melon, cucumber and other vegetables.

### 3.3 SAMPLING PROCEDURE

Data may be collected by complete enumeration called census inquiry or by partial enumeration called sample inquiry. In the former case information is collected about each and every item comprising the whole while in the latter case information is collected about a smaller number of items which are representative of the whole so as to form an estimate of the characteristics of the whole. If such a sample is an adequate representative of the whole, is properly drawn and interpreted, then it is most likely to represent the conditions of the whole and can be fairly relied upon as if the observation has been based on complete enumeration. The primary object of sampling is to obtain maximum
information about population with minimum effort, and also to set out the limits of accuracy on estimates based on sampling.

A sample from a statistical population is the set of measurement/data that are actually selected in the course of an investigation/inquiry. The respondents selected should be representative which constitutes what is technically called a “Sample” and the selection process is called “Sampling technique”.

The listing of all the sampling units in the universe is defined as sampling frame. When each unit in the population is numbered for identification it would be called a sampling frame. Such a frame helps in identifying any particular item in the population such as electoral list of households, the layout map of a town etc.

According to Gupta (2000) “A sample is that part of universe which we select for the purpose of investigation. It should exhibit the characteristics of universe, that is, it should be a micros (small universe)”.

Sampling helps to collected vital information more quickly. Even small samples, when properly selected, help to make estimates of the characteristics of the total population in a shorter time. Any research process includes selection of elements or objects for study, collecting information from them, organizing the data and drawing conclusion from them. Much of time and cost is saved at each stage of research of sampling technique adopted is scientic as per the study.

Multi stage stratified random sampling techniques was used for selecting the unit of information in present study. Etah district was selected purposively as it was convenient to the researcher and the need of study.
Etah district is divided into rural and urban area and urban area was selected purposively in the second stage. Etah urban is divided into 25 wards and out of these wards, two wards namely ward 7 and ward 23 were selected randomly in the third stage. Each selected ward consist of 6 to 8 mohallas, out these mohallas two mohallas namely from each selected ward were selected randomly in the fourth stage. Each selected mohallas consist of 160 to 220 houses. Out of these houses, fifty percent of the houses were selected randomly in the fifth stage. From each selected house, one adolescent girl aged 11 to 19 years and one adolescent boy aged 13 to 21 were selected randomly in the sixth stage. Thus, about three hundred adolescent boys and girls formed the unit of information for the present study.

3.4 VARIABLES USED IN THE STUDY

A variable means some characteristic of each member of the unit that is to be studied such as income, age, land holding etc. which can be expressed numerically. It is a significant of any research characteristic. A variable can be regarded as some kind of yardstick that gives us a basis for the evaluation of the single unit of analysis. A variable is a set of values from a class fixation. A value is anything that can be predicted of a unit. A variable can be defined as a thing that is observed and that is of such a nature that each single observation can be classified into only one of a number of mutually exclusive classes. Hence, it can be said that variables are the conditions or characteristics that the researcher manipulates, controls or evaluates. It is any characteristic which will be useful in deducing the truth of hypothesis. Variables can be of two types.
3.5 INDEPENDENT AND DEPENDENT VARIABLES

**Independent variable:** According to Kerlinger (1990), “Independent variable is the presumed cause of the dependent variable”. The independent variable is the antecedent; the dependent variable is the consequent. The variable from which the predictions are made is independent variable. For example, age, education, religion, caste, mental health, life style and nutrient intake etc. are the independent variables in the present study. Independent variables are the conditions or characteristics that the experimenter manipulates in his/her attempt to ascertain their relationship to observed phenomenon. The following independent variables have been selected for this study i.e. age, sex, education, family type, family size, residence, religion, caste, occupation, income, mental health, life style and nutrient intake etc.

**Dependent variable:** According to Kerlinger (1990) “Dependent variable is the presumed effect of an Independent variable”. The variable that is predicted is called the dependent variable. The dependent variable is the condition which the researcher tries to explain. The dependent variables may be the health of the adolescents in the present study. Dependent variables are the conditions or characteristics that appear or change as the experimenter introduces, removes or changes the independent variables. Health of the adolescents was used as dependent variable.

3.6 STUDY TOOLS

An interview schedule specially prepared with critical discussion has always proved useful. In the present study the interview schedule was
developed in consultation with subject matter specialists and pediatricians. Subsequently necessary changes were made in the light of objectives. The questions which were important were added while irrelevant questions were deleted. Final schedule was used for data collection and proper measures were taken to avoid vague and ambiguous answers. Schedule was pretested in the field on a sample of 50 respondents to check its reliability and validity, required modification were made in the language and the format, wherever needed as per the suggestions of the experts.

Interview schedule was divided into two main parts on the basis of objectives framed:

(a) General information
(b) Specific information

(a) General Information:

This information was further divided into 3 sub groups.

Information regarding respondents: The main variables in this category were age, sex, residence, religion, caste, education and occupation.

Information regarding family and socio-economic status of respondents: Information in this section includes the age, education, occupation, total family income, members in the family and type of the family.

Information regarding the adolescents: To know the health status of the adolescents questions were asked about adolescent’s name, age, sex and physical activities. Anthropometric measurement was taken to assess the body mass index of the adolescents.
**Height:** Height is a linear measurement made up of four centre apartments legs, pelvic spine and skull. While measuring the height first the subject was aspect to remove her/his shoes /sandle stand on even flour over with the centre of their back touching the scale with their feet parallel and heels buttocks shoulders and back of their head touching the wall. The head was to be held comfortable erects. The arms were hanged loosely by the side. the measuring tape was held parallel to the body and smooth this scale was held on the top of the head in a centre crushing the hair at right off the nearest 0.5 cm (Gopal Das 1987).

**Measuring Tape (Height scale):** A good flexible non-stretchable take graduated in millimeters and centimeters was used to measure the height. Subject was straight stand without foot wear.

**Weight:** “Weight is usually regarded as the most meaningful method because weight deficit is directing related to nutritional deficiency. A portable weighing deficit machine was used for this purpose. The scale was adjusted to 200 before minimum clothing was asked to stand on the platform of the scale without shoes/ sandle and without touching anything, weight was recorded to the nearest of 100 gm (Gopal Das 1987).

**Weighing Machine:** The subjects were weighed on the portable weighing machine with light clothing and without shoes and weight was recorded to the nearest 100gm. This machine was used on plain surface.

**Health Status:** Health status measured by using the Body mess index formula given by Gomez (Gopal Das, 1987). Body mass index was calculated from the weight in kg of each individual was divided by the height in meter square (m²) on the basis of BMI, subject were classified into following group given in Table.
Table 1  
*Body Mass Index*

<table>
<thead>
<tr>
<th>Body Mass Index</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underweight</td>
<td>&lt;18.5</td>
</tr>
<tr>
<td>Normal</td>
<td>18.5 - 24.9</td>
</tr>
<tr>
<td>Overweight I</td>
<td>25 - 29.9</td>
</tr>
<tr>
<td>Overweight II</td>
<td>30 - 34.9</td>
</tr>
<tr>
<td>Overweight III</td>
<td>35 - 39.9</td>
</tr>
</tbody>
</table>

*Source: Gopal Das, 2004*

Specific Information:

This part consisted of specific information related to adolescent mental health, life style, food pattern and dietary survey. The main variables in this category were mental health, life style, food pattern and nutrient intake of the selected adolescents. Thus, specific information was collected from the respondents under following sub-divisions:

* Information regarding mental health of the selected adolescents.
* Information regarding life style of the selected adolescents.
* Information regarding food pattern of the selected adolescents
* Information regarding clinical examination of the selected adolescents
* Information regarding dietary survey of the selected adolescents

In all five categories of questions were asked about mental health, life style, food pattern and dietary survey related to following fields:
Section A - Mental health of the selected adolescents were assessed through a standardized schedule developed by Dr. A.K. Singh and Alpna Sen Gupta (2000).

Section B - Information regarding life style of the selected adolescents were also collected by asking questions related to type of work, exercise and regularity in exercise etc.

Section C - Information regarding food pattern of the selected adolescents were also collected by asking questions related to dietary habits, regularity of meal intake, number of meals taken in a day, appetite and food liked etc.

Section D - Clinical examination is an essential feature of all nutritional surveys since their ultimate objective is to access levels of health of individuals of population group in relation to the food they consume (Park, 2004). The following organs normally hair, eyes, lips, tongue, teeth, gums, skin, nails, muscular system and skeletal system were clinically examined by the researcher herself.

Section E - Information regarding dietary survey of the selected adolescents were obtained by 24 hours recall method.

3.7 VALIDITY AND RELIABILITY OF DATA

The importance of the inferences of any investigation depends upon the validity and reliability of data. Various measures were adopted to ensure validity and reliability in the context of the need of the study.

Questions were framed in simple and easy language to enable the urban women to understand them quickly and all the interviews were conducted...
separately in order to control group influence and to protect privacy of the individual respondent.

**Pilot Study:** A pilot study was conducted in a selected area on 30 respondents. This was done to assess the reliability and validity of the schedule which was developed. Questions were asked from each respondent and their responses were discussed with the experts in the light of the objectives. Subsequently necessary changes were made. The questions which were important were added while irrelevant questions were deleted. Wherever modifications were required in the language of questions, such changes were included.

Reliability: The most important characteristic of a measuring tool is reliability. A test is said to be reliable when it functions consistently. Reliability means consistency or accuracy, i.e. the degree to which the test agrees with itself. A test cannot be valid unless it is reliable, but it may be reliable without being valid.

Though there are many methods of certaining reliability of a tool such as test-retest method, split-half method, parallel test method etc. Spearman’s Split-half Correlation Method was used on 30 respondents which were not included in the actual study. The reliability coefficient was worked out as 0.8732 which indicates that the schedule was reliable.

Validity: Validity is the extent to which a test measures what it purposed to measure. No measurement procedure is valid if it does not use the valid tool for data collection because tool is the information capturing device and the tool should capture the information which fulfills it’s purpose. Therefore, validity is the most important characteristic of a tool.
The experts gave their positive opinion in relation to the constructed questionnaire. Thus it was assumed that the present questionnaire has logical and face validity. It is quite apparent that if a questionnaire is valid in relation to the objective which was to be achieved than the same will be reliable also.

3.8 PROCEDURE OF DATA COLLECTION

The present schedule is basically prepared to administer individually. As there is no specified time limit to complete the schedule but this can be completed easily in 20-25 minutes. At the time of administration, the researcher is supposed to explain the purpose of administering the schedule and to ensure the subjects the information supplied by the subjects will be kept confidential and will be used for research purpose only.

The data collection work was carried out during May, 2008 to September, 2008. The respondents were individually contacted at their residences. Personal interviews were held with the respondents. Questions in the interview schedule were asked in Hindi, preferably in the local dialect and the responses were recorded. Sometimes the questions were repeated and read out to them to obtain an accurate response.

3.9. DESCRIPTION OF STATISTICAL ANALYSIS

The collected data were coded, tabulated and analysed using various statistical techniques. The statistical tests were used to know the relationship between dependent and independent variables among the various groups of study. The significance levels were used as 5 percent for testing the hypothesis.
3.10. METHODS USED FOR STATISTICAL ANALYSIS

The following statistical test used and their purpose is given in the table 1.

Table 2
Statistical Tests Used And Their Purposes

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Statistical tests</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Percentage</td>
<td>To study the distribution of both dependent and independent variables.</td>
</tr>
<tr>
<td>2.</td>
<td>Arithmetic Mean</td>
<td>To study the central value</td>
</tr>
<tr>
<td>3.</td>
<td>Standard deviation</td>
<td>To know the variability among the observations.</td>
</tr>
<tr>
<td>4.</td>
<td>t-test</td>
<td>To test the difference between means of two groups.</td>
</tr>
<tr>
<td>5.</td>
<td>Coefficient of correlation</td>
<td>To determine the relationship between two variables.</td>
</tr>
<tr>
<td>6.</td>
<td>t-test for correlation</td>
<td>To test the level of significance of correlation coefficient</td>
</tr>
</tbody>
</table>

The tests used are described below –

(1) Percentage:

Single comparisons were made on the basis of the percentage. For drawing percentages the frequency of a particular cell was multiplied by 100 and divided by total number of respondents in that particular category to which they belonged. The percentages were calculated till one place after decimal point, similar to the practice followed in National Family Health Survey (NFHS).
(2) Arithmetic Mean:

Arithmetic mean is the average used in the present study. "Arithmetic mean of a series is the figure obtained by dividing the total values of various items by their number" (Elhance, 2008).

\[ \bar{X} = A + \frac{\sum fu}{\sum f} \times i \]

where, \( \bar{X} \) = Arithmetic mean
\( A \) = Assumed mean
\( \sum fu \) = Product of frequency and deviation from the assumed mean
\( i \) = Class interval
\( \sum f \) = Total frequency.

(3) Standard Deviation:

It is usually denoted by letter \( \sigma \) (small sigma) of the Greek alphabet and is a measure of dispersion. Standard deviation is the square root of the arithmetic mean (average) of the squares of the deviation measured from the mean or assumed mean. (Elhance, 2008).

\[ \sigma = i \sqrt{\frac{\sum fu^2 - (\sum fu)^2/n}{n - 1}} \]

where, \( \sigma \) = Standard Deviation.
\[ \Sigma f u^2 = \text{Sum of the product of frequency and square of deviation from assumed mean.} \]

\[ \Sigma f u = \text{Sum of the product of frequency and deviation from assumed mean} \]

\[ n = \text{Number of observations} \]

\[ i = \text{Class interval} \]

(4) Test of Significance:

t-test can be expressed as -

\[ t = \frac{|\bar{X}_1 - \bar{X}_2|}{\text{C.S.E.}} \]

where \( \bar{X}_1 \) and \( \bar{X}_2 \) are the means of the first and second group.

C.S.E. means combined standard error. It is calculated by the following formula -

\[ \text{C.S.E.} = \text{C.S.D.} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \]

Where \( n_1 \) and \( n_2 \) are the number of observations in the first and second group respectively. C.S.D. is the combined standard deviation. It is calculated by the following formula –

\[ \text{C.S.D.} = \sqrt{\frac{(n_1 - 1) \sigma_1^2 + (n_2 - 1) \sigma_2^2}{n_1 + n_2 - 2}} \]

If \( n_1 \leq 30 \) and \( n_2 \geq 30 \)
\[
\text{C.S.D.} = \sqrt{\frac{n_1 \sigma_1^2 + (n_2 - 1) \sigma_2^2}{n_1 + n_2 - 1}} \quad \text{If} \quad n_1 > 30 \quad n_2 \leq 30
\]

\[
\text{C.S.D.} = \sqrt{\frac{(n_1 - 1) \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2 - 2}} \quad \text{If} \quad n_1 \leq 30 \quad n_2 > 30
\]

\[
\text{C.S.D.} = \sqrt{\frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2}} \quad \text{If} \quad n_1 > 30 \quad n_2 > 30
\]

\[
\text{C.S.E.} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}
\]

\(\sigma_1\) and \(\sigma_2\) are standard deviations of first and second group respectively.

(5) Correlation Coefficient (r):

A mathematical method of measuring the intensity and magnitude of linear relationship between two variables of the series, was suggested by Karl Pearson, 1936 (Elhance, 2008).

The coefficient of correlation or the product moment correlation is calculated by the following formula –

\[
r = \frac{n \Sigma \Sigma f u v - \Sigma f u \Sigma f v}{\sqrt{[n \Sigma f u^2 - (\Sigma f u)^2][n \Sigma f v^2 - (\Sigma f v)^2]}}
\]

Where, \(r\) = stands for correlation coefficient,
\[ n = \text{number of observations.} \]

\[ \Sigma \Sigma f_{uv} = \text{sum of the product of deviation of } x \text{ and } y \text{ variables with their frequencies} \]

\[ \Sigma f_u = \text{sum of the product of deviation of } x \text{ variable with the frequency.} \]

\[ \Sigma f_v = \text{sum of the product of deviation of } y \text{ variable with the frequency.} \]

\[ \Sigma f_{u^2} = \text{sum of the product of squares of deviation of } x \text{ variable with the frequency.} \]

\[ \Sigma f_{v^2} = \text{sum of the product of squares of deviation of } y \text{ variable with the frequency.} \]

(6) **Formula for testing correlation coefficient by t-test:**

\[ t = r \sqrt{\frac{n-2}{1-r^2}} \]

Where, \( n \) is the number of observations.

\( r \) is the correlation coefficient.
3.11. LIMITATIONS

1. The study was conducted on 300 adolescents in Etah city so the findings of the present study cannot be generalized to other areas and other districts of the state.

2. Information related to anthropometric measurements in most of the respondents was obtained after visiting them two times and more.

3. Information related to nutrient intake required frequent visits.

4. Data for the present study has been collected only through interview schedule. Other methods like longitudinal case studies, observations etc. can act as a good source of information collection.