3.1 Research Design

Research design decides the fate of a proposal and its outcome. If the design is defective, the whole outcome and report will be faulty and undependable. It is upon the design that nature of data to be calculated will very much depend. It is therefore desirable that research design should be methodologically prepared. A good research design, however, depends on research purpose and is bound to be different in the case of exploratory, formulative and diagnostic studies. It is therefore desirable that research design should be methodologically prepared and it depends on research purpose. This chapter explains the various tools of research followed to achieve the specific objectives of this study.

In the present study, the research design is cross-sectional study in nature. It is a multi-stage cluster randomized sampling using the probabilities proportional size (PPS) method.

3.2 Sample and Sampling

Sample size was calculated taking into consideration the total urban Sikh population of Amritsar as reported by the Census report of 2011 and the assumed prevalence of T2DM, hypertension and MS in the subject population. The sample size of 1089 subjects was calculated. The standard WHO 30 cluster sampling approach was used to select the subjects from the sample population. Randomization of the subjects was done using the probability proportional size (PPS) method. The selection of outcome measures was based on the diagnostic criteria for Type 2 Diabetes Mellitus, Hypertension and Metabolic syndrome. Questionnaire was developed to elicit information on these lifestyle disorders divided into 11 components. The multivariable logistic regression model was used to adjust for age and sex as confounders.
The fasting blood sample was taken for the analysis of biochemical variables. The blood was analysed for serum cholesterol, (CHO), serum triglycerides (TG), (High density lipoprotein (HDL), (Low density lipoprotein (LDL), very low density lipoprotein (VLDL). The anthropometric readings were taken namely height, weight, waist (WC) and hip circumference (HC).

3.3 **Inclusion Criteria**
- Males or females aged ≥20 years,
- No limitations of physical activity and
- Those people who had shifted from the rural agricultural background to the urban setup over the last one decade.

3.4 **Exclusion Criteria**
- Severe orthopaedic/cardiovascular/respiratory conditions restricting physical activity,
- Women who were pregnant and lactating at the time of study and
- Known case of HIV infection.

3.5 **Instrumentation and Outcome Measurement**

To carry out the study following tools were used.

Subjects were assessed for the following ABC (Anthropometry, Biochemical, and Clinical)

1. **Anthropometric Assessment**: All hip and waist circumferences were taken with the subject in standing upright position. All the measurements were repeated three times at same position and conditions. Average of the three readings taken by the same observer was used for the final analysis.

1.1 **Body mass Index**: BMI was calculated by using formula weight (kg)/height (m²).
**Design and Methodology**

a) **Height (cm):** Height was measured without shoes by using a stadiometer. Standing height was assessed through maximum vertical stature for persons who could stand unassisted. A fixed stadiometer with vertical backboard, fixed floorboard and movable headboard was used. Subjects stand with the heels of their feet against the vertical backboard a wall, with feet pointing outward at approx a 60 degree angle. Body weight was distributed evenly with both feet flat on the floor. The back, buttocks, claves and heels was touching the backboard, feet together. The head was in the Frankfort plane (an imaginary line from the ear canal to just below the lower orbit of the eye should be parallel to the floor). Once the subject was positioned the headboard or a flat ruler was placed on the top of the head, with sufficient pressure to compress the hair.

b) **Weight (kg):** Weighing machine (Acto inc) was used to measure body weight to the nearest 0.1kg. Subject stand still on the platform, with the body weight evenly distributed between both the feet. After removing all heavy clothes (e.g. jackets, coats, shoes) with empty pockets weight was recorded.

2. **Biochemical Test:** The following investigations were performed on fasting venous blood sample.

a) **Blood glucose:** Blood glucose monitoring system Elegance CTX10 was used to determine blood sugar levels.

b) Nycocard Reader II was used for Nycocard HbA1c test to measure glycosylated hemoglobin. Plasma glycosylated hemoglobin (HbA1c) concentration was the main outcome for glycemic control.

c) **Lipid profile:** A fasting venous blood sample was obtained after anthropometry and physical examination for lipid profile. Estimation of total cholesterol (TC), serum triglycerides (TG), and high-density lipoprotein cholesterol (HDL-C), was performed on the sample drawn after 12 hours overnight fast. Bayer’s semi automatic analyzer for measuring lipid profile (HDL, serum triglycerides and total cholesterol levels). “Value of low density lipoprotein cholesterol (LDL) was calculated according to Friedewald’s equation” (Friedewald et al., 1972).
3. **Clinical History and Examination:** Following general physical examination was performed

a) **Blood Pressure:** Blood pressure was measured by a standard mercury sphygmomanometer, Lifecare\textsuperscript{TM} (N & B Medical Products Co.) according to JNC guidelines.

b) Past/Family history of Type 2 Diabetes mellitus, Hypertension/Metabolic syndrome.

4. **Questionnaires**

The first step was to design the questionnaire for urban Sikh community of Amritsar including both male and female population to elicit data regarding the diagnosis and awareness of T2DM, hypertension and metabolic syndrome. Questionnaire was developed to elicit information on these lifestyle disorders divided into 11 components. Questionnaire developed included examination of content validity. The key points to design the questionnaire were determined and were shown to three professionals including: Biostatistician, Physician and Endocrinologist to identify significant omissions. Based on their suggestions and recommendations, questionnaire was structured. The questionnaire was then shown to the English Professor for proof reading and grammatical corrections. It was finally reviewed by six experts including a Biostatistician, Physician, Cardiologist, Endocrinologist, Psychologist and Community Medicine specialist. They were asked to analyse the contents of the questionnaire, wordings and order of items in questionnaire. Based on the feedback of the experts, final content of the questionnaire was determined. The consensus questionnaire was administrated to the subjects. During the interview questionnaire items that required explanation or clarification were noted. The explanation given was recorded and was used to revise the items. The questionnaire was administered on the subject repeatedly till the time the questionnaire were found to be simple, clear and free from any mistakes. (Floyd, 2002; Mildred, 2004: Lim and Chen, 2007). A pilot study on 100 subjects was done to pretest the questionnaire.
To enumerate the information pertaining to following points were taken into consideration in the questionnaire:

a. Demographic data  
b. Family history  
c. Medical history  
d. Educational qualification  
e. Employment status  
f. General health  
g. Symptoms of the lifestyle disorders  
h. Dietary habits  
i. Physical activity  
j. Compliance indicators  
k. Awareness of risk factors

It was essential to include all these variables in the questionnaire, so that we could get an insight picture of the prevalence of these lifestyle disorders and their associated risk factors in the urban Sikh population of Amritsar. The demographic data provided us with the details of the family type, their educational status, occupation, marital status and their duration of stay in the current place of residence. The family history included that if either of their parents or siblings had suffered from T2DM, hypertension or metabolic syndrome. The medical history was also important to be taken into consideration as we came to know about the medications they were taking for any systemic disorders. The dietary habits and physical activity patterns showed us the pattern of the diet the subject was taking, their medium of cooking and milk and milk products intake. The awareness of risk factors was important as we could gather information about the risk factors the subject population was aware off.

3.6 Methodology and Experimental Protocol

This study focused on Urban Sikhs living in the Amritsar, Punjab, India as they constitute 70% of the total population of Amritsar according to 2011 census report. The area was arbitrarily divided into five equal zones namely ($Z_1$, $Z_2$, $Z_3$, $Z_4$, $Z_5$). Each zone
Design and Methodology

constituted of 13 Censes wards. The list provided the name, age and address of those eligible for voting (>18 years). The fieldwork was completed in a period of 20 months, starting in Jan 2012. The totals of 215 subjects from each zone were randomly selected. In the selection of the family in a particular ward the WHO method of sampling was followed to prevent any non uniformity in selection of the subjects. A model consent form was constructed to ensure compliance with ICMR guidelines regarding the use of human subjects in research. All protocols and consent documents were reviewed and approved by the Institutional Ethics Committee of Faculty of Sports Medicine & Physiotherapy, Guru Nanak Dev University, Amritsar. The complete detailed performa of all the family members was filled by the investigator which included number of members in the family, age, educational status etc. From the list provided one male and one female member were randomly selected. These two family members were requested to fill in the self designed, professionally validated questionnaire.

A person on antihypertensive drug or controlled/uncontrolled blood pressure or person falling into the diagnostic criteria of hypertension as defined by (Chobanian et al., 2003) was labelled as hypertensive. Similarly for T2DM diagnostic criteria of IDF guidelines (IDF Clinical Guidelines Task Force, 2015) was followed. Chances of false hypertensive or false T2DM is very small, however, it cannot be ruled out in those who has controlled blood pressure or controlled sugar, but on medication.

3.7 Statistical Analysis

The various statistical formulae and equations, which were used for the analysis of the findings, are presented as follows:

Descriptive Statistics

The mean standard deviation and standard error were calculated to describe the data.

1. Arithmetic Mean (\( \bar{X} \)):

Arithmetic Mean gives the average of the whole range of the data given by adding together all the items and by dividing this total by number of items, and is given by the following formula:
Design and Methodology

\[ X = \frac{\sum X}{N} \]

Where,

\[ X \] = Arithmetic mean
\[ \sum X \] = sum of all variables
\[ N \] = Total number of all variables

2. **Standard Error (SE):**

   It enables the measurements of magnitude of the sampling error. It is calculated by the following formula:

   \[ \text{S.E.} = \frac{\sqrt{\text{SD}}}{N} \]

   Where
   
   \[ \text{SD} \] = Standard Deviation
   \[ N \] = Total number of variables

3. **Standard Deviation (SD):**

   It gives the degree of deviation or dispersion of the recorded data from the mean. It is given by the formula:

   \[ \text{S.D.} = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} \]

   Where
   
   \[ \text{SD} \] = Standard Deviation
   \[ X \] = Individual Variables
   \[ \bar{X} \] = Mean of Variables

4. **Students ‘t’-test:**

   It gives the difference between the two independent random samples of size \( N_1 \) and \( N_2 \) with mean \( \bar{X}_1 \) and \( \bar{X}_2 \) and S.E. of \( X_1 \) and S.E. of \( X_2 \). It is calculated by the following formula:
Design and Methodology

\[ t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(SE_{X_1})^2 + (SE_{X_2})^2}} \]

Where

\( t \) = 't' test  \\
\( \bar{X}_1 \) = Mean of 1\text{st} variable  \\
\( \bar{X}_2 \) = Mean of variable  \\
\( SE_{X_1} \) = SE of 1\text{st} variable  \\
\( SE_{X_2} \) = SE of 2\text{nd} variable

5. Pearson's chi-squared test:

“The Chi Square (X^2) test is used to see the association between two categorical variables. Chi Square is employed to test the difference between an actual sample and another hypothetical or previously established distribution such as that which may be expected due to chance or probability”.

\[ X^2 = \sum \frac{(Observed \ frequency - Expected \ frequency)^2}{Expected \ frequency} \]

Degrees of freedom (df) = n-1, where n is the number of classes

6. Binary logistic regression analysis:

“Logistic regression analysis describes how a binary response variable (Y = 1 or Y = 0) is associated a set of explanatory variables. This method may be thought of as an approach that is similar to the multiple linear regression, but takes into account the fact that the dependent variable is categorical”.

Binary logistic regression

\[ \log \left( \frac{P(Y = 1)}{1 - P(UY = 1)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \]

Binary logistic regression model fits the log-odds (logit) by a linear function of the explanatory variables. As P(Y) ranges from 0 to 1, the logit ranges from - \( \infty \) to + \( \infty \).