CHAPTER VI

RESEARCH METHODOLOGY

6.1 Research Design

Research is an organized, systematic, data based, critical, objective, scientific inquiry or investigation into a specific problem, undertaken with the purpose of finding answers or solutions to it with view to increasing knowledge. It is undertaken to explore, test and establish relationships between variables of a selected and identified scope of study.

Research design involves and highlights the methodological rigor and appropriateness of the intellectual design for conducting research. It aids in drawing a careful, detailed and exacting approach to conducting a research. It throws light on the basic research questions and problems addressed in the scope of the study. A research design is a framework that guides how research should be conducted, based on certain philosophies, principles and assumptions.

The flowchart of research design which is followed for the current study is placed on the next page. In this study Branding and Positioning, the two major concepts are studies with respect to pharmaceutical industry. This study took design of Exploratory, descriptive and predictive research. After designing theoretical framework data is collected and analysis is performed.

Research process is explained in the form of a chart in the next section of the thesis, and detailed explanation follows subsequently.
6.2 Research Process

**FIGURE 9: RESEARCH PROCESS**

Source: Uma Sekaran and Roger Bougie - Research methods for Business
6.6.1 Exploratory study

An exploratory study is undertaken when not much is known about the situation at hand, or no information is available on how similar problems or research issues have been solved in the past. For the present study extensive preliminary work is done to gain familiarity with the phenomenon in the situation to understand what is occurring, before the development of model. The respondents for the study were asked about reasons for prescription, what do they prefer, what comes to their mind first, elementary statements were generated. The data was collected through observation and interviews. This qualitative data then is converted into the scale through which brand positioning is understood. Theories are developed and hypothesis is formulated for the further investigation.

6.2.2 Descriptive Study

A descriptive study is undertaken in order to ascertain and able to describe the characteristics of the variables of interest in a situation. In these variables like years of practice, age group, gender, education, and company name recall are studied. This study helped in describing characteristics of the group. The goal of this study is to understand profile and relevant aspects of individuals, organizations and industry.

6.2.3 Causal Study

In a Causal study variables are selected as dependent and independent, predictor and criterion, stimulus and response. The study of change in dependent variable is performed based on change in independent variable.
6.3 Sampling Methodology

Sample is drawn from population. For physicians, data is collected from major hospitals and clinics from Mumbai. Being capital city, Mumbai is a commercial place with advance medical facility. Major pharma companies are based in Mumbai. It is a good representative of population. Sample size for patients is 442 and for physicians 464. Eight major companies are studied in detail for understanding their positioning with respect to physicians.

**TABLE 2 : SAMPLE COMPOSITION FOR THE STUDY**

<table>
<thead>
<tr>
<th>Sample Subjects Under Study</th>
<th>Sample Size (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians</td>
<td>464</td>
</tr>
<tr>
<td>Patients</td>
<td>442</td>
</tr>
<tr>
<td>Companies</td>
<td>8</td>
</tr>
</tbody>
</table>

Based on Cohen (1988) provides extensive tables for calculating the number of participants for a given level of power. Some guidelines used for sample calculations are, for standard $\alpha$-level of 0.05 and recommended power of 0.08, 783 participants are required. For the current study 1000 questionnaires were sent for data collection. After data cleaning, the precise data used for data analysis is 906, which is a good representative for the study. The practicing physicians are selected as sample subjects for collection of data. Data for both patients and physicians is collected from hospitals and dispensaries. While for positioning map, from ims India data companies are selected. The stratum chosen is financial performance. The availability of the data was also one of the important factors.
6.4 Data collection Methodology

Types of data: Majorly for this study secondary and primary data is collected.

6.4.1 Secondary data

The data is collected from various internal and external sources.

1. Internal company Records
2. Company reports
3. Internal computer databases
4. Reports and publications of government agencies
5. Other publications.
6. Computerized databases.

The data for the research work is collected from source like EBSCO, PROQUEST, EUROMONITOR, ACEANALYSER, EMARALD AND IMS INDIA.

6.4.2 Primary data

Primary data is collected through various methods like Interview and questionnaire.

Initial phase of Understanding positioning with respect to pharmaceutical industry is carried out by conducting open ended discussion with physicians and practicing doctors. Reasons for prescriptions were asked to them.

6.5 Questionnaire design

Questionnaire is the preferred method for data collection for its advantages of administration and convenience. The speed of data collection is high with questionnaire as compared with
other methods. Suitability of the method is checked and questionnaire is administered to the respondents.

Following steps are undertaken for designing questionnaire.

Specification of the requirement is stated, target audience is decided. Attempts are made to reduce both surrogate information error and respondent’s error. Initial cover letter is introduced for respondents understanding of objective of the study and their importance in the study. The privacy disclaimer is inserted at the end note. Initial demographic details are asked to patients like includes age, education and occupation. More details about diseases and companies awareness are also asked. An interval scale that is Likert scale is preferred for understanding important factors for them when it comes to medicines and drugs. The Five point scale for agreement and seven point scale of relevancy is preferred because of the perception related constructs.

*Pretesting of Scale for Brand positioning measurement and measures:*

The scale is pretested over few respondents for clarity and lucidity. It was observed that physicians could take 7-8 minutes to fulfill the survey instrument i.e. questionnaire.

*Pretesting of Scale for Brand Personality:*

The scale is adapted for measuring Brand personality from Aaker (1997). Some important traits related to pharmaceutical concepts are added. The reliability and validity of the same is checked.

*Pretesting of Scale for Brand Trust:*

The scale of Hess (1995) is adapted for understanding the concept of trust whose reliability and validity is checked with reference to objective of the study.
As suggested by Churchill (1979), Cronbach Alpha and exploratory factor analysis was undertaken to check the reliability and validity of the data. The reliability and validity tests were confirmed and were similar to those in literature.

Unit of analysis:

The unit of analysis is Patients, Pharmaceutical companies and Physicians. The data primarily collected from the above-mentioned unit of analysis. Secondary data is collected about the industry, number of players and top 300 brands contributing highest to sales.

6.5.1 Validity and Reliability of questionnaires

Validity and reliability in the data collection relates not only to accuracy of the data items that are measured but also its accuracy with respect to the purpose for which it was collected. The reliability of a measure is established by testing for both consistency and stability. Consistency indicates how the items measuring a concept hang well together as a set. Cronbach’s alpha is a reliability coefficient that reflects how well the items in a set are positively correlated to one another. Cronbach’s alpha is computed in terms of the average inter-correlations among the items measuring the concept. The closer Cronbach’s alpha in to 1, the higher the internal consistency reliability (Kerlinger, 1986).

Another measure of consistency (reliability) used in specific situations is the split half reliability coefficient. Since this reflects the correlations between two halves of a set of items, the coefficient obtained will vary depending on how the scale is split-up. Sometimes split-half reliability is obtained to test for consistency when more than one scale, dimension, or factor is assessed, and the items across each of the dimensions, or factors are split based on
some predetermined logic (Campbell, 1976). In almost every case, Cronbach’s alpha is an adequate test of internal consistency / reliability.

6.6 Statistical Analysis Techniques

Suitable statistical analysis tools used to analyze the data. Appropriate univariate, bivariate and multivariate analyses were used depending on the nature of variables and objectives of the study.

Univariate Analysis:

Univariate analyses refers to the analysis in which there is single variable. In this study univariate analysis was used for identifying the descriptive characteristics of the data. Frequency tallies, Histogram, and descriptive are obtained by this method.

Bivariate Analysis:

Bivariate analysis includes the simultaneous analysis of two variables. It is undertaken to establish relationship between two variables. Correlation and ANOVA (Analysis of Variance) are the bivariate techniques used in this study. Correlation aims at ascertaining whether or not tow variables are varying together.

Multivariate analysis

These techniques are used when more than two variables are studied at a time. Meaningful conclusion can be found out from the combination of data by using these techniques.

6.6.1 Analysis of Variance (ANOVA)

ANOVA is used to uncover the main and interaction effects of categorical independent variables (called “factors”) on an interval dependent variable. A “main effect” is the direct
effect of an independent variable on the dependent variable. An “interaction effect” is the joint effect of two or more independent variables on the dependent variable. The key statistic in ANOVA is the F-test of difference of group means, testing, if the means of the groups formed by values of the independent variable (or combinations of values for multiple independent variables) are different enough not have occurred by chance. If the group means do not differ significantly then it is inferred that the independent variable(s) did not have an effect on the dependent variable, then multiple comparison test of significance are used to explore just which values of the independent(s) have the most to do with the relationship.

If the data involve repeated measures of the same variable, as in before-after matched pair’s tests, the F-test is computed differently for the usual between groups design, but the inference logic is the same. There are also a large variety of other ANOVA designs for special purposes, all with the same general logic.

It is noted that analysis of variance tests the null hypotheses that group means do not differ. It is not a test of differences in variances, but rather assumes relative homogeneity of variances. Thus a key ANOVA assumption is that the groups formed by the independent variable(s) have similar variances on the dependent variable (“homogeneity of variances”). Levene’s test is standard for testing homogeneity of variances. Like regression, ANOVA is a parametric procedure which assumes multivariate normality (the dependent has normal distribution for each value category of the independent(s)).

6.6.2 Factor analysis

Factor analysis is a specific computational technique used to examine patterns of relationships (correlations) amongst select variables (also called factors that are common among large number of variables). The objective of this technique is to reduce a large number of variables to more manageable number variables (data reduction) from a larger set of
variables based on nature and character of these relationships. The data reduction process is based on relationship or intercorrelations among the variables within the correlation matrix. The most frequently used approach is principle component analysis. This method transforms a correlation or covariance matrix into set orthogonal components that are equal to the original variables. The new set up of variable Pi called principle components are linear combination of composite variable with weights also called factor load. These linear combinations of variable called factors contribute maximum total variance account in the data as a whole.

The Kaiser-Meyer Olkin test of Sphericity was used for measuring sampling adequacy (KMO). After selecting “Analyse” from the, SPSS menu bar, “Dimension” and “Factor Reduction” and then Factor was carried out. After clicking on “Descriptive”, in the statistics box initial solution was checked. In the correlation matrix. KMO and Bartlett’s test of Sphericity was checked and reproduced. The KMO statistics varies between 0 and 1. A value close to 1 indicates that patterns of correlations are relatively compact and so factor analysis should yield distinct and reliable factor (Malhotra & Dash, 2009)

6.6.3 Discriminant Analysis

The ordering of things into classes is a basic procedure for empirical science. The analyst used multi-measurements for differentiating between two or more groups if individuals, things, or events. The traditional method is to compute the significance of the difference between the means of groups, taking each characteristic separately. However, the method is inefficient in that it does not make it possible to evaluate the relative amount of information for differentiation; provided by several measurements taken together. Neither does it combine the information taking into account the interrelations, if they exist, between the characters dealt with. Hence, an alternative is to construct a linear combination of the variables, a weighted sum, in such a way that it best discriminates among the groups in some sense.
Discrimination analysis is the method by which such linear combinations are determined. Discriminant analysis is similar to regression analysis, which involves the investigation of the criterion/prediction variable relationship; also a weighted sum of the measurements is needed as in multiple regressions. The difference lies in the nature of criterion, which is qualitative rather than quantitative as in the case of multiple regressions.

Multiple discriminant analysis is a generalization of the method of discriminant analysis appropriate for only two groups. It is mainly used as a method for studying relationship among several groups for populations, and provides a basis for the classification of individuals among several groups. Discriminant analysis may be interpreted as a special type of factor analysis, which extracts orthogonal factors of the measurements, taking into account the differences among the criterion groups. The model derives the components that best separate the cell or groups. In many instances, the problem is one of studying group differences for classifying items into groups based on certain criterion.

6.6.3.1 Assumptions of discriminant analysis

In discriminant analysis, sample of individuals are assumed to be drawn from several different population where different \( p \) quantitative scores are available for each individual. The \( p \) measurements are assumed to follow a multivariate normal distribution with equal variance-covariance matrices within the several populations.

6.6.3.2 Scope of multiple discriminant analysis

The primary importance of multiple discriminant analysis is the study of relationships among several groups in terms of multiple measurements. This provides a basis for classification of individuals among the several groups. The approach provides tests of significance for certain important hypotheses for relationships among several groups; for example a single composite
score accounts for all significant differences among the groups. Some of the tests are discussed in the following section.

6.6.3.3 Testing statistical significance of discriminant functions

1. Mahalanobis’s distance: The first step in attest of significance of a discriminant function is to measure the separation or distinctness of the two groups. This can be done by compounding Mahalanobis’ distance, which can be tested using multivariate equivalent of the $t$-test for the equality of two means. Mahalanobis’s $d^2$ statistics is a squared distance measure. It is equivalent to the standard Euclidian distance measure, and measures the distance from each case to the group mean.

2. Wilk’s lambda: This is a statistical significance test of the discriminant function. The ability of the variables to discriminate among the groups, (beyond those from which the information was extracted for the previously computed function) is measured by Wilk’s lambda. This is multivariate measure of group’s difference over discriminating variables and can be calculated in many ways. In general, it is calculated such that values of 1.0 there is no discrimination. For the present study Wilk’s lambda for test of statistical significance.

3. Canonical correlations: The canonical correlation coefficient is a measure of the association that summarizes how the discriminant function is related to groups.

6.6.3.4 Interpreting dimensions of discriminant functions

The problem of interpreting the nature of the dimensions of discriminant function is a difficult task. One easy way to characterize the dimensions is in terms of groups they separate most. In the past, the nature of the discriminant function was described by examining relative magnitudes of the weighting coefficients. This can be problematic, as the coefficients are dependent on the unit of measurement, which may be different for different original
measures. Therefore, the effects of differences are largely removed by multiplying each discriminate function coefficient of the standard deviation of the particular variable to which the weight is applied. This is equivalent to dividing the weights by the square root of the pooled variance.

This will render the within-group variance to unity. Following this, the relative magnitudes of the coefficients can be compared to determine which variables contribute most to the definition of the composite function. Morrison (1969) has provided details of the interpretation for discriminant analysis.