CHAPTER 4

RESEARCH METHODOLOGY

4.1: Introduction:

Literature review of Corporate Governance, Corporate Social Responsibility, Clinical Governance and Organizational Climate are discussed in the previous chapter. The need and importance of good corporate governance and CSR practices are highlighted along major issues in health care organization affecting them. Organizational climate of health care organization is slightly different because of the complex nature of health care sector. Clinical governance is most discussed aspects in health care organizations. Hence this chapter of the thesis summarizes the research methodology which reflects about research setting, the universe of the study, sampling method, and determination of sample size, data sources and questionnaire design. It is followed by the essential research tools and techniques that have been applied to successfully validate the research objectives and hypotheses.

4.2: Research Settings:

As research is an art of scientific investigation, it deals with the systematized effort to gain new knowledge and information. It comprises of numerous systematic activities such as; defining and redefining problems, formulation of hypotheses, collecting, organizing data, evaluating and interpreting data, making deductions and reaching to a conclusion. The research design and methodologies for the research process needs to be pre-set before the real execution. For systematic implementation of research, this section has been subdivided into the following five sub-sections.

4.3: Research Design:

For the purpose of thorough investigation of the research problem, descriptive research design has been used in the present research. Auberbach & Silverstein (2003), has described “The descriptive research attempts to describe, explain and interpret conditions of the present i.e. ‘What is’. The purpose of a descriptive research is to examine a phenomenon that is occurring at a specific place and time. A descriptive research is concerned with conditions, practices, structures, differences or relationships that exist, opinions held processes that are going on or trends that are evident”. Descriptive work consists of three main categories such as;
observational research, case methods and survey research. These works have been used in previous research work to analyze the factors of the research problem effectively. The present work adopts exploratory research methods, as the present research is seeking to generate meaningful hypothesis by examining collected dataset and looking for building relations between the variables. The exploratory research method is adopted, in order to formulate a more precise problem or develop hypotheses. If the theory happens to be too general or too specific, a hypothesis cannot be formulated. Exploratory research is also helpful in formulating relevant hypotheses for more definite investigation (Shields & Rangarjan, 2013).

4.4: Research universe:
The prime objective of this research is to study the nature of Corporate Governance and CSR practices of health care organizations, specifically private health care organizations are taken into consideration for the study. So the private health care organizations operating in Bangalore were identified for the study. Both primary and secondary data sources along with various multivariate techniques were applied together to analyze and validate the hypotheses proposed in the study.

4.5: Development of the Questionnaire:
The process of designing the questionnaire is as follows. Whenever previously used scales were available, they were first screened by the researcher for their easy interpretability in the Indian private healthcare context. This was followed by a formal pilot study in one hospital in the same geographical area. During pilot study, in addition to an intensive questionnaire survey supplemented with occasional interviews, several respondents sat with the researcher and answered each question. During the pilot study majority of respondents were Doctors and Nurses. This process helped the researcher not only to know how respondents understood the questions, but also to appreciate how this understanding differed across organization types and between Clinicians and Non- Clinicians. The analysis of the pilot study results pointed to the need for further changes. A changed format was then prepared and tried out on a different set of healthcare organization acquaintances.

4.6: Sampling Plan:
As it is not feasible to contain all populace in the study, selection of samples and data collection is the most important part of research study. Thus, it is significant to draw out data samples that can take the satisfactory size of population. The simple random sampling method is used as probability sampling for availing respondent's opinion on the subject, whereas convenience and
cluster sampling is used as non-probability sampling for selection for hospitals. The simple random sampling method is fitting for this research as it undertakes to limit the possible data to those which are “less extreme” by insuring that all sections of the populace are corresponding to the data in order to increase the effectiveness, by minimizing the error of probability (Agresti & Finlay, 2008). The respondents of the structured questionnaire were mainly the internal stakeholders (employees) of the Indian private healthcare setting. As a group of sample data was brought into reflection, therefore, it was important that the data collected for the purpose were enough for interpreting the research results. Data size refers to the number of elements (the respondents) to be included in the study. Determining the sample size is difficult and it involves numerous qualitative as well as quantitative reflections to the research study. Some important qualitative dimensions that should be measured in formulating the data size includes the significance of the judgment, environment of the investigation, a number of items, nature of the analysis, data size used in familiar works and resource.

**The actual respondents to the questionnaire are 452**, which include 140 from multi specialty hospitals and 166 from teaching hospitals the remaining 146 from super specialty hospitals in Bangalore. 5 point Likert scale is used in the questionnaire.

### 4.7: Data Collection Procedure:

The field research methods used here for data collection are questionnaire method and interview method and systematic observation methods. In the first round researcher visited two health care organizations in each category (multispecialty, teaching and super specialty). These Health Care organizations are situated in the capital city of Karnataka i.e., Bangalore. The measuring instrument was used to capture the responses of employees of the organizations. The questionnaires were prepared only in English language as respondents were familiar with the same. During the pre-test, the participants were told that the study was a part of a doctoral level research. The participants were told that their participation in this study was voluntary and confidentiality was assured. The complete filling of questionnaire takes about 10 to 15 minutes.

The research sample was selected from employees of each health care organization (HOD, Doctors, Nurses, Deans, Principals, and Directors). Respondents’ inclusion criteria comprised of employees working in the organizations at least since last one year and willing to participate in the survey. The sample respondents were chosen randomly in each private hospital and the questionnaires were filled with face to face fashion.
4.8: Research Methods Applicable:

4.8(a): Exploratory Factor Analysis

Since its inception a century ago (Spearman, 1904, 1927), factor analysis has become one of the most widely used multivariate statistical procedures in applied research endeavors across a multitude of domains (e.g., Psychology, education, sociology, management, healthcare etc). The fundamental intent of factor analysis is to determine the number and nature of latent variables or factors that account for the variation and Co variation among a set of observed measures, commonly referred to as indicators. Specifically, a factor is an unobservable variable that influences more than one observed measure and that accounts for the correlations among these observed measures. In other words, the observed measures are inter-correlated because they share a common cause (they are influenced by the same underlying construct); if the latent construct was partially out, the inter correlations among the observed measures would be zero to one. Thus, factor analysis attempts a more stringent understanding of the co variation among a set of indicators because the number of factors is less than the number of measured variables.

In social science and management research, factor analysis is most commonly used in psychometric evaluations of multiple-item testing instruments. It is a data reduction technique. In the early stages of scale development, researcher might use factor analysis to examine the plausibility or credibility of this assumption i.e., the ability of a single factor to account for the inter correlations among the different variables and to determine if all variables are reasonable indicators of the underlying construct of particular dimension or factor i.e., how strongly is each item related to the factor?.

In the present research factor analysis was used to identify the issues affecting Corporate Governance and CSR practices of health care organizations. The main objective of factor analysis is to reduce the number of items and to notice the construction of the relationships between various items; i.e., to segmentation of variables. Hence, the factor analysis is applied as an item reduction or structure detection method (the term factor analysis was first introduced by Thurstone, 1931). Therefore, factor analysis has been classified into two categories. Those are Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Both EFA and CFA aim to reproduce the observed relationships among a group of variables with a smaller set of latent variables, but they differ fundamentally by the number and nature of a priori specification and restrictions made to the factor model. EFA is a data-driven approach such that
no specifications are made in regard to the number of latent factors (initially) or to the pattern of relationships between the common factors and the indicators.

4.8(b): Confirmatory Factor Analysis (CFA):

The objective of CFA is to identify latent factors that account for the variation and co-variation among a set of indicators. EFA and CFA both models are based on the common factor model and thus many of the concepts and terms that were discussed in EFA like factor loadings, unique variances, commonalities and residuals. In CFA the researcher must pre-specify all aspects of the factor model: the number of factors, the pattern of indicator, factor loadings and so forth. CFA requires a strong empirical or conceptual foundation to guide the specification and evaluation of the factor model. Accordingly, CFA is typically used in later phases of scale development or construct validation after the underlying structure has been tentatively established by prior empirical analyses using EFA, as well as on theoretical grounds. EFA and CFA often rely on the same estimation methods like maximum likelihood approach. CFA is strongly driven by theory or prior research evidence. Thus, whereas in EFA the researchers can only pre-specify number of factors; the CFA researcher usually tests a much more parsimonious solution by indicating the number of factors, the pattern of factor loadings and an appropriate error theory. Thus, every aspect of the CFA model is specified in advance. The acceptability of the specified model is evaluated by goodness of fit, the interpretability and strength of the resulting parameter estimates. CFA is more appropriate than EFA in the later stages of construct validation and test construction, when prior evidence and theory support more risky a priori prediction regarding latent structure. In addition to this CFA offers a very strong analytic framework for evaluating the equivalence of measurement models across distinct groups such as demographic groups like sex, race, culture or economic conditions of the respondents. CFA is a very important component within a broader class of methods called Structural Equation Modeling (SEM) or covariance structure analysis. Generally, CFA is used as a precursor to SEM, which specifies structural relationships among the latent variables.

Most of the numerical methods require at least one statistical test to establish the significance of an analysis. However, in CFA, many statistical tests are used to determine, how well the model fits to the data (Suhr, 2006). While reporting the results of a confirmatory factor analysis; Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Normed Fit Index (NFI),
Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA) are narrated. These model fit indices are described as follows:

4.8(c): Goodness of Fit Index
The goal of goodness-of-fit approach is to identify the solution that reproduces the observed correlations considerably better than stringent models (i.e., Models involving fewer factors). The GFI is able to reproduce these observed relationships equally or nearly as well as more complex solutions (i.e., Models with more factors). Its measure indicates how well a specified model reproduces the covariance matrix among the indicator variables. The GFI was an early attempt to produce a fit statistic that was less sensitive to sample size. This statistic is still indirectly sensitive to sample size due to the effect of N on sampling distributions (Maiti & Mukherjee, 1991). The possible range of GFI values is zero to one with superior values indicating better fit, the GFI value of greater than .95 considered good.

4.8(d): Adjusted Goodness of Fit Index (AGFI)
The AGFI tries to take into account the different degrees of model complexity. It does so by adjusting GFI by a ratio of the degrees of freedom used in a model to the total degrees of freedom available. The AGFI penalizes more complex models and favors those with a minimum number of free paths. AGFI values are typically lower than GFI values in proportion to model complexity.

4.8(e): Normed Fit Index (NFI)
The NFI is one of the original incremental fit indices. It is a ratio of the difference in the $\chi^2$ value in the fitted model and a null model divided by the $\chi^2$ value for the null model. It ranges between zero and one and a model with perfect fit would produce an NFI of 1. The CFI was derived from this index in an effort to include model complexity in a fit measure.

Comparative Fit Index (CFI)
Comparative fit indices are also referred to as incremental fit indices (Bentler, 1990; Hu & Bentler, 1998) which evaluate the fit of a user-specified solution in relation to a more restricted, nested baseline model. Typically, this baseline model is null or independence model in which the co variances among all input indicators are fixed to zero, although no such constraints are placed on the indicator variances. As you might expect, given the relatively liberal criterion of evaluating model fit against a solution positing no relationships among the variables, comparative fit indices often look more favorable i.e., more suggestive of acceptable model fit.
than indices from the preceding categories. The value of CFI ranges from zero to one with values closer to one implying good model fit.

4.8(f): Root Mean Square Error of Approximation (RMSEA)

A widely used and recommended index for this category is the Root Mean Square Error of Approximation (Steiger & Lind, 1980). The RMSEA is a population base index that relies on the non-central $\chi^2$ distribution, which is the distribution of the fitting function when the fit of the model is not perfect. The non-central $\chi^2$ distribution includes a non-centrality parameter (NCP), which expresses the degree of model mis-specification. The NCP is estimated as $\chi^2 – df$ (if the result is a negative number, NCP = 0). When the fit of the model is not perfect, the NCP is greater than zero and shifts the expected value of the distribution to the right of that of the corresponding central $\chi^2$. The RMSEA is an error of approximation index because it assesses the extent to which a model fits reasonably well in the population (as opposed to testing whether the model holds exactly in the population of $\chi^2$. In the RMSEA it is rare to see the exceeding value greater than one. As with the SRMR, RMSEA values of zero indicate perfect fit and values close to zero suggest good model fit.

4.8(g): Structural Equation Modeling:

Structural equation modeling is a statistical methodology used by many different sector researchers like economists, educational researchers, marketing researchers, medical researchers, healthcare professional and a variety of social and behavioral scientists. Structural equation modeling (SEM) uses various types of models to depict relationships among observed variables, with the same basic goal of providing a quantitative test of a theoretical model hypothesized by the researcher. More specifically, various theoretical models can be tested in SEM that hypothesize how sets of variables define constructs and how these constructs are related to each other. One reason for its pervasive use in many scientific and social research fields of study is that structural equation modeling provides researchers with a comprehensive method for the quantification and “testing of theories”. Other major characteristics of structural equation models are that they explicitly take into account the measurement error that is ubiquitous in most disciplines and contain latent variables.

SEM models essentially combine path models and confirmatory factor models; i.e., SEM models incorporate both latent and observed variables. In the management and behavioral sciences, researchers are often interested in studying theoretical constructs that cannot be observed
directly. These abstract phenomena are termed as latent variables or factors. Because latent variables are not observed directly, it follows that they cannot be measured directly. Thus, the researcher must operationally define the latent variable of interest in terms of behavior believed to represent it. As such, the unobserved variable is linked to one that is observable, thereby making its measurement possible. Assessment of quality, then, constitutes the direct measurement of an observed variable, albeit the indirect measurement of an unobserved variable (i.e., The underlying dimension). It is important to note that the term quality used here in the broad sense to include scores on a particular measuring instrument.

Due to the mathematical complexities of estimating and testing the proposed assertions, computer programs are a must in applications of structural equation modeling methodology. The hypothesized model can then be tested statistically in a simultaneous analysis of the entire system of variables to determine the extent to which it is consistent with the data. If goodness-of-fit is adequate, the model argues for the plausibility of postulated relations among variables; if it is inadequate, the tenability of such relations is rejected. It takes a confirmatory rather than an exploratory approach to the data analysis (although aspects of the latter can be addressed). Furthermore, by demanding that the pattern of inter variable relations be specified a priori, SEM lends itself well to the analysis of data for inferential purposes. By contrast, most other multivariate procedures are essentially descriptive by nature (e.g., exploratory factor analysis), so that hypothesis testing is difficult, if not impossible. Secondly, whereas traditional multivariate procedures are incapable of either assessing or correcting for measurement error, SEM provides explicit estimates of these error variance parameters. The model-fitting process can therefore be summarized as follows:

\[
\text{Data} = \text{Model} + \text{Residual}
\]

Where data represent score measurements related to the observed variables as derived from persons comprising the sample. The model represents the hypothesized structure linking the observed variables to the latent variables and, in some models, linking particular latent variables to one another. Residual represents the discrepancy between the hypothesized model and the observed data.

SEM follows a logical sequence of five steps or processes: model specification, model identification, model estimation, model testing, and model modification. These basic building blocks are essential in conducting SEM models.
4.9: CONCLUSION
This particular research is based on survey of internal stakeholders’ perception about Corporate Governance and CSR practices in the private healthcare sector and was conducted according to the principles of the Declaration of Helsinki as revised in 2013. At the same time prior oral permission was taken from the respondents about their willingness to participate in the survey. This chapter outlines the different research methodologies such as EFA and CFA with SEM. The methodologies have been executed for meaningful and systematic outcome of the research carried out in this study. Data was collected through primary and secondary sources. The internal stakeholders who are working (employed for more than one year) with selected private healthcare organizations have been considered as a valid sample for the research. The selection of data for this research work was based on simple random sampling method and an attempt was made for equal representation of respondents from different types of hospitals. The succeeding chapter would examine different techniques and the motivation for the selection of research techniques in detail for data analysis.