CHAPTER II
REVIEW OF LITERATURE

2.0. Introduction:

This chapter gives a brief overview of selected studies on efficiency of health services and health sector financing with a view to understand the issues involved and identify the appropriate methodology for this study. The review is made non-technical, keeping the technical details to the next chapter. The studies relating to health sector efficiency are reviewed separately for developed and developing countries. This is followed by studies on Andhra Pradesh. In the penultimate section, we give a brief review of the health financing issues and studies that stress on user charges. The last section contains conclusions of the chapter.

2.1. Studies on hospital efficiency:

2.1.1. Studies using performance indicators:

Lasso (1986) analyzed the performance of 73 Colombian hospitals for the period 1977-80. Hospitals are categorized into different groups on the basis of their bed strength. Using combined utilization and productivity (CUP) analysis for assessing their performance, the author found that hospitals with larger bed strength have higher BOR and ALS followed by other groups with smaller bed size.

Sear (1991) conducted a study of efficiency of investor owned and non-profit hospitals by using a sample of 142 hospitals in the State of Florida, USA in 1988. Three indicators (ratios), namely total number of full time equivalent (FTE) personnel per active bed, number of man hours per adjusted patient day, and the total wages paid per adjusted patient day are used for the purpose. The author found that investor owned
hospitals used significantly fewer FTE staff per bed, fewer man hours per adjusted patient
day and paid significantly less in wages

2.1.2. Econometric studies:

There are a number of studies, which used econometric models for studying hospital efficiency. The problems addressed here include, quality and case mix heterogeneity, uncertainty, and input price variations. The nature of these problems and the way they can be tackled are briefly discussed below in the context of developed country. The review of studies for both developed and developing countries follows there after.

(0 Quality heterogeneity:

Most studies on health services ignore the problem of quality heterogeneity. Quality variations in health services hinder efficiency comparisons between providers. Failure to control quality differences may ascribe higher efficiency to lower quality producers and vice-versa. The next question is what is health care quality and how it can be measured and controlled for?

There are alternative definitions for health care quality. Wyszewianski et al (1987), for example, adopted a conventional, technical definition of health care quality. It is based on both the extents to which health care providers comply with technical norms of care and the service's ability to improve the patient's health status. According to them, "quality is higher for care A than for care B, if for the same patient, care A is likely to make a greater net contribution to the patient's health and well-being than is care B."

Barnum and Kutzin (1993) recognize that the quality of health care has two dimensions, the medical or technical and the consumer's dimension."Quality has both
supply and demand characteristics. The critical demand issue is 'perceived' quality, the consumer assessment of the relative quality of different health care providers. Adequate staff and supplies are obvious supply side factors affecting 'actual' quality of services that are important in affecting perceived quality" (Chapter 3, p 17).

The notion of quality that we adopt for our study is demand side one. This includes the demand side dimensions such as attitude of medical personnel, cleanliness of premises, and other measures of the degree to which patients' preferences are met. Because of the multi-dimensional nature of quality, it is not possible to infer the estimates of efficiency by combining quality information with cost of care. To do so, knowledge about the weights of different dimensions of quality would be required. That, however, is beyond the scope of this study.

It is sometimes argued that perceived and technical quality may or may not be in agreement with one another. Discrepancy in these viewpoints can be greater in developing countries where poor, uneducated populations may have a more limited ability to discern the technical merits of health services. Barnum and Kutzin (1993) suggest that quality perceptions by patients are in some cases related with technical quality measures, such as the availability of drugs and medical supplies. Thus, the policy measures aimed at improving technical quality through increase in supplies can also improve quality perceptions.

To sum up, the quality of care can be assessed and controlled for in various ways when measuring efficiency of health services. These include measuring quality according to (1) degree of compliance of medical norms of practice, (2) clinical outcomes and (3) various measures of patient's satisfaction.
(ii) **Case mix:**

Comparative studies of health facility cost and efficiency have to overcome the difficult methodological problems arising from the wide diversity of health care activities produced by alternative providers and the effects that such heterogeneity has on resource use, cost, and efficiency.

A patient-day care may differ between hospitals, between different departments of the same hospital and over time. According to Tatchell (1983), these differences may arise due to differences in (a) technology, (b) quality of care, (c) case-mix, (d) case complexity and severity, and (e) institutional characteristics (size, teaching status, location, composition, ownership etc.) So the output of the hospitals need to be standardized in order to bring it to a comparable form for measuring efficiency. Among the different methods of standardising the output, service mix is the foremost. Service mix may be defined as the type of services available or the services and the procedures actually performed. Cohen (1967), Anderson (1976), Kouner (1969) have used the method of service-mix, taking the actual procedures performed in the hospital.

Another method of standardizing hospital output is to derive a measure of hospital case-mix using the case types treated in the hospital. One of the early attempts to standardize hospital output for case mix was based on specialty mix. Feldstein (1967) made an attempt to standardize the hospital output by dividing the various specialties into 8 mutually exclusive groups. Another approach to deal with case mix is diagnostic mix. The studies related to this latter method have used either international classification of disease (ICD) groupings or diagnostic related groupings (DRG) for standardizing the output. Evans (1971) and Evans and Walker (1972) have used ICD groupings for...
dividing the diagnosis into different homogenous categories Fetter et al (1980) and Barer (1982) have used DRG for standardising hospital output Other methods of standardising hospital output are found in Lave and Lave (1970), Lave, Lave and Silverman (1972) However, Tatchell (1983) gave a detailed review of the studies related to the issue of case mix and service mix He identifies two methods of standardizing hospital output, namely service mix and case mix

The service mix approach uses information about the type of services offered by the hospital (pediatrics, gynecology, etc) as a basis for standardising hospital output Tatchell (1983) argues that the existence of common services among providers does not imply that the mix and complexity of cases, as well as quality of care are similar among providers He also points out that whether or not certain services are offered by a provider is not sufficient to characterise its output, due to lack of information about the extent to which these services are actually used

The case mix approach attempts to standardise hospital output according to the mix of cases actually treated in the hospital Vitaliano (1987) calls this method as demand-side one In fact, both methods reflect demand and supply forces since the services offered by hospital generally respond to a demand The case mix approach, however, captures supply as well demand relations better because it is based on utilisation information Under the case mix approach, Tatchell distinguishes several methods They are, specialty mix, diagnostic related groupings, information theory, case severity, and others The reminder of this section focuses on the use of the case mix approach to standardize output when measuring health facility efficiency

The case mix denotes composition as well as complexity of cases While
measuring hospital cost, these two aspects are of utmost importance and need to be considered. For example, if only composition is taken into account, then wards in two hospitals would appear identical if both of them treat same number of diabetic cases over a given period of time. If information on case complexity is also taken into consideration, it could be that the two facilities differ in terms of the proportion of complicated cases. Also, it is to be noted that the facilities, which treat more number of complex cases, consume more resources resulting in higher cost per treatment. Thus, study of health facility efficiency, technical or economic, requires knowledge about the case mix. This information can be used to control for case mix variability while estimating cost.

However, while it is possible to obtain data about composition of output, it is usually difficult to get information on output complexity. Complementary techniques must then be used. Some of these are given below.

Register and Bruning (1987) in their comparative study of efficiency between profit and non-profit oriented U.S. hospitals have used an alternative method for dealing with case mix, since the information on case mix of individual hospitals was not available. Case mix proxies are used to limit the confounding effects of case heterogeneity on efficiency measures. They assumed that rural and urban hospitals differ in their case mix and therefore confined to only urban hospitals in their study. All long-term (older) federal hospitals are also eliminated from the sample to reduce heterogeneity. Hospitals that provide a specified set of services are retained, thus eliminating all hospitals, which are too 'basic' or 'high tech'. It is posited that hospital size, measured by the number of beds, is also associated with the case mix. Therefore, they limited their sample to hospitals within the range of 100 to 200 beds.
To verify their hypothesis that hospital size is associated with case-mix, the authors group the sample hospitals by size and compare the groups on the basis of bed-to-doctor and bed-to-nurse ratios. The differences in ratios are interpreted as differences in case mix. However, it should be noted that the above procedure might introduce a bias in their analysis because bed-to-labor ratios may reflect not only case mix but also efficiency differences. The authors have also estimated technical efficiency using multiple-output production function approach. This estimation method allowed them to control case mix composition, while the previous procedure of restricting the sample could control the case complexity as well.

Eakin and Knieser (1988) in their study of economic efficiency of US hospitals also confronted the problem of case mix differences among facilities. To quote them:

"The output data indicate that hospitals are a heterogeneous group with respect to mix of cases treated. Consequently, we adopt a multiple output specification in this study. In particular, we use four categories of outputs - general medicine, obstetrics, gynecology, weighted surgeries, and outpatient visits" (p 587)

The above procedure controls for case mix composition. In order to control for case severity and factors other than case mix, the authors proceeded as below:

"Three additional variables - TEACH, a dummy variable indicating medical school affiliation, ALOH, the average over all length of hospitalisation and the average overall occupancy rate - are included in the list of repressers to control for case severity and capacity utilisation" (p 587)

The variable TEACH is used to capture the (positive) effect on hospital costs of teaching activities. The length of stay variable is used to control for severity. However, this variable may represent differences in medical practice, not attributable to severity but to efficiency. The occupancy rate variable is introduced to discriminate between technical and economic efficiency, a useful distinction. A hospital may be technically efficient (it
can be on its production possibility frontier) but, due to low occupancy it may be economically inefficient.

Vitaliano (1987) adopted a service mix approach to control for case mix and its effect on hospital efficiency. For each hospital in the sample, an un-weighted index of 11 available hospital facilities or services is constructed and used as an independent variable to estimate a total cost function. He finds that the service-mix variable is statistically significant and has a positive effect on total cost.

To sum up, there are several methods available to control case mix when evaluating health services efficiency. These include:

(a) limiting the sample of providers to those that provide a similar case mix,
(b) in econometric studies of efficiency, include indicators of case mix such as type and volume of services produced, measures of output complexity etc as explanatory variables and
(c) limiting the comparison of efficiency to a few, well defined medical services (e.g., appendectomies, hernia repair, etc)

The author has also mentioned several other methods under case mix approach, such as specialty mix, DRGs, information theory, case severity and others.

(iii) Uncertainty:

The uncertainty in the hospital care arises from different dimensions. These are uncertainty with respect to (a) the timing of an illness and hence the demand for hospital services, (b) physician's diagnosis, and (c) the final outcome of the treatment. Since demand uncertainty in particular is likely to have important consequences on hospital costs, it is necessary to understand the cost and uncertain demand.
If one views the hospital, in addition to providing direct or anticipated patient care, as also providing sufficient capacity to assure that hospital services are available at the time of unexpected demand, this insurance or option demand of the community at large should be treated as yet another service provided by the hospital. Failure to take this standby service into account in econometric cost models may result in specification error.

Cowing et al (1983) suggest the following formula for measuring the level of standby services

\[ k = \frac{(\text{Beds} - \text{Average Daily Census})}{(\text{Average Daily Census})^{1/2}} \]

Here, larger the 'k', there is less chance that services will not be available for yet another patient (Jaskow, 1980) Although this issue has been recognised by various authors, it could not be incorporated in hospital costs in a more formal manner A detailed review of the studies dealing with this issue is given in Cowing et al (1983) In brief, it suggests that demand or output uncertainty accounts for some portion of the observed excess capacity of hospitals and it had lead to decreasing cost with increase in hospital utilisation

(iv) Input price variation:

Input price variation across providers complicates the interpretation of differentials in efficiency For example, let there be two providers who operate with equal technical efficiency, but facing different input prices An analysis of efficiency that focussed on technical efficiency alone would conclude that both the providers are equally efficient. This would contrast with the outcome of a study on economic efficiency for the following reasons If the technology and input mix were similar between providers, the provider facing higher input prices would exhibit the lower economic efficiency
Furthermore, if factor prices vary among providers, no inferences can be drawn about economic efficiency, unless a more careful analysis is conducted. Social cost of resources, instead of factor prices, may have to be considered in such cases.


2.1.3. Studies for developed countries:

During the past two and half decades, estimation of cost function for hospitals has become a common place for industrialised countries. Reviews by Cowing, Holtman and Powers (1983), Wagstaff (1989), and Ricardo (1992) have documented this literature. The initial attempts to estimate cost functions using data from hospitals in industrialised countries employed composite measures of hospital output (eg. Cohen 1967). They used average or unit cost of inpatient day or admission as the dependent variable, and a variety of interrelated explanatory variables such as occupancy rates, patient flow, length of stay, and capacity as explanatory variables (Mann and Yett 1968).

Most of the early empirical studies [Feldstein (1967), Berry (1967), Carr and Feldstein (1967), Ingbar and Tayler (1968), Francisco (1970)] were concerned with the effect of size on hospital costs - that is, the existence of scale effects and the related questions of optimal size. Authors of these studies estimated rather restricted versions of the general cost function, usually without recognizing many of the implicit assumptions.
involved. For example, most of the models did not include input prices in the estimated cost function, an omission which is equivalent to assuming zero input substitutability, that is, a fixed proportions technology for producing hospital care.

In addition, many of these studies typically include two types of variables in the cost function. The first one is a flow type variable, number of patient days of service provided, to account for variations across hospitals in actual or current output. The second variable is a stock type variable, the total number of beds, to account for variations in size or capacity. Some models have used other measures of output and the ratio of output to capacity. The basic purpose of these latter models appeared to be that of relating cost changes caused by output changes, holding capacity constant, to the short-run cost structure, and cost changes caused by changes in capacity, holding output constant, to the long-run cost structure. This in turn is related to the existence of economies of scale. The general conclusion of these studies was that there was evidence of significant economies of scale, at least up to moderate sized hospitals of around 500 beds.

Later studies have specified cost functional forms and included variables that are consistent with a theoretical production structure (e.g., Cowing and Holtman 1983, Grannemann et al. 1986, and Vita 1990). They have used total cost (rather than average cost) functions with multiple outputs and have employed flexible functional forms. The data used were mostly from USA or UK. We discuss briefly some of these studies.

Grannemann et al. (1986) departs from previous studies of hospital cost by introducing a more flexible functional form for total cost. They combined a translog form with a conventional cubic cost function where outputs enter directly, not in the
logarithmic form, to permit zero output levels. Capital is included as an independent variable, reflecting their assumption that capital is not a fully fixed cost to the hospital. This allows exogenous variations in the hospital's capital cost even in the short-run. They also included various hospital outputs, like emergency care, home visits, and two measures of inpatient care, to account for differences among hospitals in average length of stay. Due to data limitations, they have not included the interaction terms between input prices and output levels in their model. The authors estimated various measures of hospital performance such as marginal cost, product specific economies of scale and scope. Data on hospital costs, inputs and outputs are taken from American Hospital Association's annual survey for 1981. Ordinary least squares estimation method was used for estimating the cost function.

Vitaliano (1987) argued that econometric studies of efficiency, which use multiple outputs in a cost function are plagued by the problem of multicollinearity among outputs. He, therefore, used a total cost function with a single output to conduct a comparative study of economic efficiency using a sample of 166 New York State hospitals. Instead of using output measure as an explanatory variable, the author used the number of beds. Other explanatory variables used in the model include, condition of hospital, index of case mix of services provided in the hospital, regional dummy, and a measure of hospital's monopoly and monopsony power. The latter variable is measured by hospitals' share of the total supply of beds in the country. Vitaliano argued that monopsony or oligopsony hospital might exert downward pressure on factor prices. The estimation technique was weighted least squares. A quadratic U-shaped average cost function was obtained from the estimated total cost function. As expected from a U-
shaped average cost curve, the hospital costs have exhibited economies of scale in certain range of output. He attributed this to presence of high fixed cost such as specialised personnel and equipment.

Eakin and Kniesner (1988) proposed a general methodology for estimating hospital cost functions, which allows for the possibility that hospitals may not be cost minimisers. They point out that "erroneously assuming cost minimisation leads to inaccurate estimates of factor substitution possibilities" (p 584). The authors estimated a system of equations which include observed cost function and three observed factor share equations. They specify a 'hybrid' translog total cost function to accommodate zero output levels instead of their logarithm.

The model included four outputs and four inputs and was estimated using the method of seemingly unrelated non-linear regressions. The sample consisted of 331 US short-term hospitals, both profit and non-profit oriented. Allen elasticities of factor substitution and economies of scale are estimated. The empirical results show that hospitals undervalue the cost of capital and overvalue the price of physicians. Hospitals thus tend to overemploy capital and underemploy physicians. This resulted in allocative inefficiency, which equals to 5 percent of the total observed cost. They also found that elasticities of substitution and factor demands are sensitive to model specification.

Frank and Taube (1987) studied technical and allocative efficiency in the production of outpatient mental health clinics using data from 755 clinics obtained through a survey by the US National Institute of Mental Health in 1982. The study focussed on estimation of production functions for mental health visits. Like Eakin and Kniesner (1988), this study was also interested in exploring provider departures from cost.
minimisation behavior. The method consists of verifying marginal productivity theory, i.e., equality between the factor price ratio and the ratio of marginal products for each pair of inputs. Differences between the above two ratios are attributed to either over or under utilization of inputs. The authors also estimated the economies of scale and scope in the production of mental health visits. Finally, the study looked into the determinants of efficiency by focusing on differences in productivity between government-run and private clinics.

In the empirical analysis, Cobb-Douglas and Transcendental functional forms are used for specifying a production function. Ordinary least squares estimation was used. Both the forms indicated decreasing returns to scale and greater productivity of private clinics, as evidenced by an ownership dummy included in the production functions. The study also has shown violation of marginal productivity theory, i.e., factor price ratio between physicians and other clinical staff differed from the corresponding ratio of marginal products, with physicians being over-employed. This also signaled a departure from cost minimization in the production of outpatient mental health services in the USA.

Eakin (1991) used the results from Eakin and Kniesner (1988) to study the determinants of economic efficiency in hospitals. For this purpose, the estimated values of allocative inefficiency are regressed on several hospital and market-related characteristics considered to be the determinants of efficiency based on economic theory. The independent variables included ownership, regulatory factors, competitive factors, factors characterizing the sources of hospital revenue, measures of hospital size, and regional dummy variables. Estimation was done using generalized least squares. It was found that the regulatory environment is a critical determinant of economic efficiency.
Economic inefficiency was greater in larger hospitals with a larger market share. He does not find statistically significant differences in hospital efficiency between non-profit and profit-oriented hospitals and likewise between church operated and other hospitals.

2.1.4. Studies for developing countries:


Ramlingas-wami (1984) estimated the cost of medical education in 17 medical colleges in India and found that the recurrent cost of educating a medical graduate varied between Rs 55,000 to 1,34,000 in the year 1981 among different medical colleges. Thus, most of the earlier studies have concentrated in developing a suitable methodology for costing rather than evaluating hospital performance.

Rodriguez and Jimenez (1985) conducted a comparative study of productivity between private and (decentralized) public Chilean hospitals. They measure the productivity in terms of inverse average length of stay (ALS), namely shorter the ALS, higher the productivity. The authors divided ALS into three components, namely diagnosis (D), medical treatment (T) and recovery (R) respectively. It was argued that while T must be performed within the hospital, D and R can partially be accomplished on outpatient basis. For a given patient-case-mix, the authors hypothesized that a series of individual specific variables, such as age and income can influence D and R. It was assumed that a patient's income is a close proxy of the type of health insurance or
coverage he/she has. Thus, income was assumed to closely correlate with out-of-pocket price of the services and it influenced individual behavior, as measured by D and R. The type and severity of illness, and the amount of medical inputs provided to the patients, were also assumed to affect the D, T and R.

Using a sample of 5 hospitals (3 private and 2 public) and 369 patients, the authors used OLS to regress length of stay on patient age and income, medical inputs and facility specific dummy variables. Three separate equations were estimated for obstetrics and gynaecology, surgery and internal medicine. The authors conclude that, other things being equal, private hospitals had lower length of stay and were more productive than public institutions. They cautioned, however, that their analysis was subjective to the accuracy of information about diagnosis, case mix and severity of the illness etc.

Lewis, Suletta and Forgia (1990) used cost accounting method to measure efficiency and quality of care at Aybara hospital, which is a 271-bedded government-run facility in the Dominican Republic. To estimate costs, the authors monitored a selected sample of patients during their treatment in the hospital, recording the cost of services provided to them. The sample consisted of three sets of patients, (i) emergency patients, (ii) people consulting in an ambulatory basis, and (iii) inpatients admitted to five departments (3 surgical and 2 ophthalmology) of the hospital during a one-week reference period. The sample patients underwent 23 surgical interventions consisting of wounds, appendicitis, cataract, hysterectomy and hernia during that week. The prices of non-labor inputs were obtained directly from supplier wherever possible. Labor costs were computed by measuring actual staff time devoted to medical procedures and multiplying it by workers actual wage (converted to an hourly basis). The total cost of
procedures was the sum of the variable and allocated fixed cost. The variable cost included cost of labor, drugs, ancillary services and consumables. Cost of overheads, depreciation of buildings, equipment and other fixed assets constituted the fixed cost. The quality of care was measured in two ways, (a) assessing the appropriateness of the qualifications of the medical staff involved in care and (b) comparing the actual diagnosis and treatment practices and services delivered with medical norms of care.

By extrapolating the percentage of labor used from the sample to all patient services in the hospital, the authors concluded that only 12% of the medical labor contracted by the hospital could actually be counted for. Although no allowances were made for down time by the medical staff, this unexpected result signaled a major inefficiency in operations of Aybara hospital. With regard to quality of care, major departures were found between the expected cost of meeting norms of diagnosis and treatment and the cost of such services actually provided. The cost of drugs dispensed and the tests performed represented about 10% of the costs implied by the norms.

Anderson (1980) studied the behavior of hospital costs in Kenya using a sample of 51 hospitals during 1975-76. The dependent variable was average cost per patient day. The explanatory variables were, capacity as measured alternatively by available and used beds, occupancy rate, average length of stay, number of outpatient visits per inpatient day, number of satellite ambulatory facilities operating under the hospital and the nature of hospital, provincial or non-provincial.

Four alternative specifications of the model were estimated using alternative dependent and independent variables. It was found that the hospitals were operating with increasing returns to scale, as evidenced by negative and statistically significant
coefficients associated with scale variables. Further, higher occupancy levels resulted in lower average costs implying that greater demand should be accommodated within the existing hospitals rather than through new ones. Outpatient activity was found to increase average cost. In contrast, length of stay did not come out statistically significant. Because the cost of hospital and its satellites are intertwined, the average cost used as dependent variable also included the cost of satellite facilities. The regression results showed that a greater level of satellite activity had a positive impact on aggregate average cost. Finally, the provincial hospitals were found to have higher average cost than district and sub-district hospitals.

Dor (1987) has estimated average cost function for 19 urban hospitals in Peru. The hospitals included were some from the ministry of health and others from social security ministry/department. Using the average cost function, Dor obtained an analytical expression for the optimal patient flow, i.e., the flow at which hospital average cost was a minimum. Using OLS and WLS methods, he estimated separate average cost functions for each of the three cost categories (labor, goods and services) as well as combined. He found that hospital average cost decreased with service intensity. This means, the average cost of hospitalization decreased either with increase in number of hospitalizations or decrease in the number of beds or both. The calculated optimal flow was 3.2 admissions per bed per month. This value was above the sample mean but below the actual flow of several hospitals in the sample. Thus, the study concluded that the average cost curve was 'U'-shaped.

Bitran and Dunlop (1990) studied the determinants of hospital costs using a sample of 15 government hospitals for Ethiopia, with one to three annual observations for
each hospital and total of 38 observations. Flexible functional form as in Grannemann et al (1986) was estimated using the OLS method. Several parameters of interest such as the marginal cost of inpatient care and outpatient care, average incremental cost, product specific economies of scale, and economies of scope measures were calculated from the estimated cost function.

The authors found that the hospitals in the sample were operating under constant returns to scale for patient days, laboratory tests, and deliveries. Economies of scope were found between first outpatient visit and inpatient days, signaling an economic advantage (reduction in cost) in joint production of inpatient and outpatient care. The limitation of the study was to use 'planned' hospital expenditure by the Ethiopian government to represent cost. This may differ substantially from the actual expenditure.

Wouters (1990) studied the cost and efficiency of a sample of 42 private and public health facilities in Ogun State, Nigeria. She analysed efficiency and cost by estimating production and cost functions respectively. Technical efficiency was assessed using the estimated production function and the associated measure of marginal product of health workers. She found that the efficiency variable was insignificant and thus concluded that departures from cost minimisation have little effect on expenditures. She also found that the marginal costs are less than average costs and thus concluded that facilities in her sample exhibit increasing returns to scale both for admissions and outpatient visits. Based on the estimated economies of scope, it was concluded that there were no advantages in the joint production of inpatient and outpatient care.

Shephard et al (1991) conducted a cost effectiveness study of surgery in intermediate health units (IHU) in Cali, Columbia. In order to reduce the high cost of
treatment, IHUs were set up as an intermediate unit between PHCs and hospitals. The study showed that surgical anesthesia complications were higher in the hospital than in IHU, patient's satisfaction as measured by the time taken to return to their work after surgery, was higher at IHU. The average cost was much higher at the hospital than in IHU.

Barnum and Kutzin (1993) computed the same indicators as in Lasso (1986) for measuring the performance of public hospitals in some of the developing and industrialized countries. The data used by them was either average of all hospitals in the selected countries or average of sample hospitals in respective countries. They found that the industrialized countries perform better in terms of their utilisation as well as productivity compared to developing countries like Turkey, Ethiopia and Korea.

Barnum and Kutzin (1993) have also conducted an evaluation study of financial management of 8 Class I and Class II hospitals distributed throughout Columbia. The costs were evaluated at 1975 constant pesos, and service data correspond to each year during 1975-78, thus providing a pooled time series of cross sections. Using a total cost function, the authors found that at the sample average, the marginal costs for inpatients and outpatients were approximately equal to the respective average costs, implying short-run constant returns to scale with respect to both the variable factors.

Another study was conducted in 1993 by the same authors for China. A short-run variable cost function was used for the purpose. The sample included cost and service data for three years, 1984-86. The authors found diseconomies of scale and only mild economies of scope, implying thereby short-run inefficiency in the level of operation with respect to bed days and outpatient visits.
2.1.5. Studies for Andhra Pradesh:

Mahapatra and Herman (1991) used the method of ratio analysis for evaluating the performance of secondary level hospitals in Andhra Pradesh, India. Due to heterogeneity in the hospitals covered, the authors grouped the hospitals on the basis of available number of beds and the type of service provided. The hospitals providing similar type of service are grouped into one category. The authors found differences in the performance of different categories of hospitals.

Mahapatra and Herman (1994) applied the combined utilization and productivity (CUP) analysis for evaluating the performance of secondary level hospitals in the state of Andhra Pradesh (AP). Based on a sample of 108-109 hospitals for the years 1989-90, the authors computed bed occupancy rate, bed turnover rate and average length of stay for each hospital in the sample. Using the graphical method suggested by Lasso (1986), the authors found that a sizable proportion of hospitals, namely, 407 percent in 1989 and 394 percent in 1990, were associated with low productivity and low utilisation. Similarly, an estimated 31.5 percent of hospitals in 1989 and 33.9 percent in 1990 were associated with high utilisation and high productivity.

After the study by Mahapatra and Herman (1994), there were no published works, to our knowledge, that measured hospital performance in AP. Since there has been a substantial change in policies relating to secondary level hospitals, there is a need to re-examine the performance indicators of the secondary level hospitals in AP. This is being attempted in this study. Thus, although there are a large number of studies relating to hospital cost in developed countries using the cost function approach, there are only a very few studies for developing countries. To the best of our knowledge, there are no
published studies in India using cost function approach for measuring the hospital performance and efficiency. The present study aims at filling this gap by using data on secondary level district hospitals in AP.

2.2. **Financing of health services:**

Government intervention in the provision of health care services is supported on the grounds of externality, equity and efficiency. Several economists like Arrow (1963), Newhouse (1970), and Cuyler (1971) gave sufficient justification for the government provision of health services. They argue that private providers cannot supply the health services due to market failure. Since health services are basic and they help in improving the quality of human capital and thereby economic development, there is a need for the government provision of these services. World development report (1993) reiterates this view. Further, the report states that the government provision of health services helps in reducing poverty, inequity, and inefficiency.

As already discussed in Chapter I, World Health Organizations' Alma-Ata declaration of 1978 outlined a global strategy for 'health for all by 2000' through the primary health care system. WHO estimated that the implementation of this strategy requires annual per capita health expenditure of US $15 in 1981 prices for most of the developing countries. Since per capita public spending on health is currently only US $2.3, there is an annual resource gap of US $50 billion for all the developing countries. Even if developing countries could fund as much as 50% of this amount, they would have to seek external funding of about seven times the present level of international transfers. The growth of per capita domestic public spending on health, therefore, is not anywhere near the levels required to meet the goals of the global strategy. [WHO (1981c), World...
Because of this reason and others (discussed in Chapter VT), various researchers and policy makers suggest alternative method of generating funds to finance health services (Jimanez 1986a, 1986b) Among the suggested measures are, user fees, development of suitable insurance system etc It is in this context that the present study aims at examining the issue of financing the health services in developing countries in general and AP in particular The study aims at reviewing the traditional arguments in favor of government intervention and the current arguments for introducing the user fees for health services by examining the resource allocation pattern in the state of AP In order to examine the feasibility and its impact of introducing the user fees, the study aims at conducting a field study on government-run hospitals in AP

2.3. Conclusion:

The purpose of this chapter was to review the past studies on hospital performance and efficiency with a view to identify the broad issues that were addressed in the literature From the above review, we notice that earlier researchers have dealt with the issues of case mix, quality heterogeneity, uncertainty, and input price variations while analyzing hospital performance and efficiency However, most of the above studies were for developed countries and we could find only a few studies for underdeveloped countries Further, even the few studies that were done for developing countries could not address the above issues One of the reasons could be non-availability of appropriate and detailed data The next chapter tries to look into the alternative approaches that were used for evaluating hospital performance and efficiency The objective is to identify an appropriate methodology, which can be used for evaluating hospital services in the state of Andhra Pradesh.