

29972

WHO/SHS/NHP/92.2
Original: English
Distr.: General

HOSPITAL ECONOMICS AND FINANCING IN DEVELOPING COUNTRIES

WILLIAM NEWBRANDER

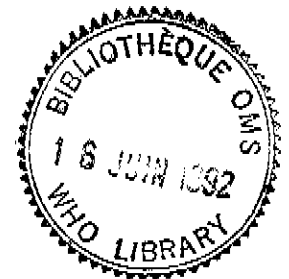
World Health Organization

HOWARD BARNUM

World Bank

JOSEPH KUTZIN

World Bank



WORLD HEALTH ORGANIZATION

Geneva

1992

TABLE OF CONTENTS

INTRODUCTION	1
The Relevance of Hospital Economics	
The Role of Hospitals in Health Systems	
Hospital Resource Issues	
RESOURCE ALLOCATION TO HOSPITALS	6
Economic Issues and Methodology	
Experiences of Countries	
Directions for the Future	
RESOURCE MANAGEMENT BY HOSPITALS	14
Economic Issues and Methodology	
Experiences of Countries	
Directions for the Future	
RESOURCE GENERATION FOR HOSPITALS	24
Economic Issues and Methodology	
Experiences of Countries	
Directions for the Future	
THE NEED FOR FURTHER HOSPITAL ECONOMICS WORK	34
References	37

INTRODUCTION

A common concern of most developing countries is health sector resources. Such concerns include: the sources of finance for health services, the ability of the public sector to maintain past funding levels, resource allocation patterns in the public sector and the efficiency of health services delivery (WHO, 1990). These concerns are legitimate due to the magnitude of expenditure on health services, which account for as much as 5 percent of gross domestic product and 5 to 10 percent of government expenditures in developing countries (Akin et al., 1987), but also because of pressures to increase health expenditures during the period of limited public sector resource availability currently being experienced in many countries.

The hospitals of these countries are an important element of the concern about health resources because they are the largest and most costly operational unit of these health systems and account for a large portion of the health sector's financial, human and capital resources. In aggregate terms,

- ◆ hospitals utilize nearly half of the total national expenditure for the health sector (Mills 1990a);
- ◆ hospitals commonly account for 50 to 80 percent of government recurrent health sector expenditure (Barnum and Kutzin, 1992);
- ◆ hospitals use a large proportion of the most highly trained health personnel: in Kenya 60 percent of the total number of physicians and 80 percent of the nursing officers are assigned to hospitals (Bloom et al., 1986).

Despite the level of resources devoted to hospitals and the health sector, there is a growing gap between available and required resources because the health needs of developing countries have not diminished, although they may have shifted with social, demographic and epidemiological changes. This gap has led many countries to explore ways to locate new resources for their hospitals and health systems. Insurance and user fees for hospital services are often seen as the primary means for obtaining additional resources. It is necessary, however, for countries to explore more seriously the possibilities for generating "additional" resources by using existing resources more effectively through improved allocation patterns and increased efficiency in the management of hospital operations. This

paper examines the economic and financial issues of hospitals in developing countries to identify the work which lays ahead to promote improved policies, planning and management of hospital resources.

The Relevance of Hospital Economics

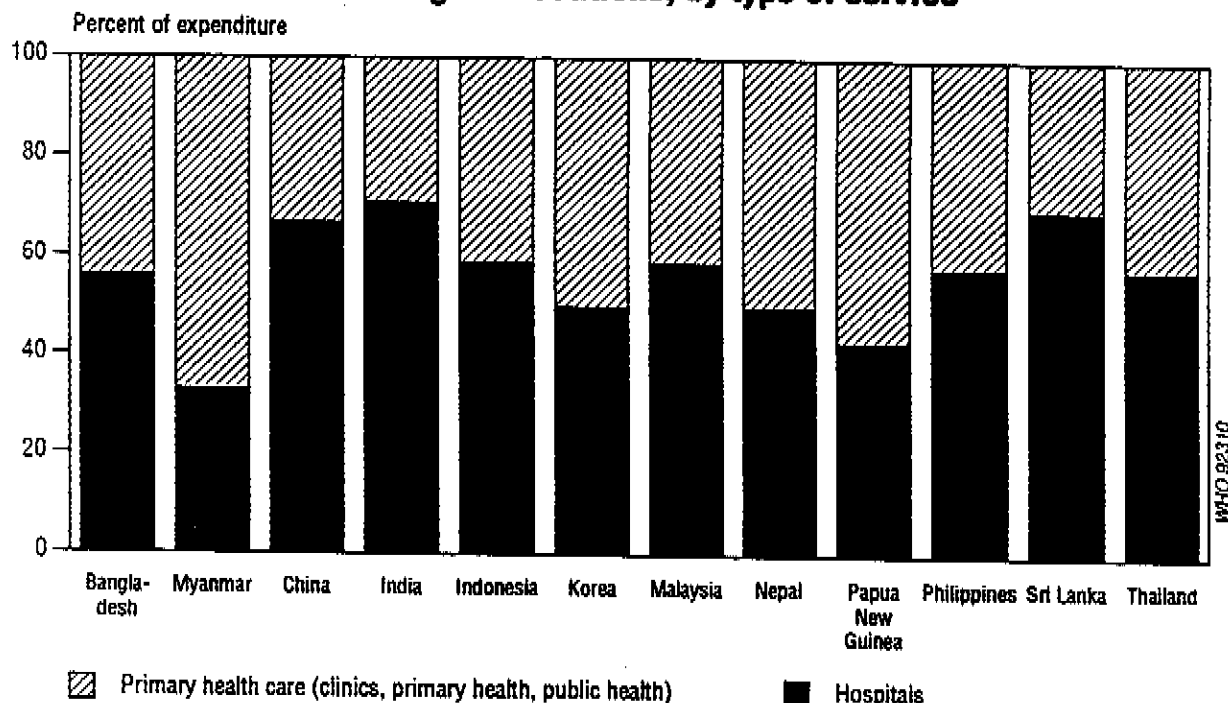
In developed countries the concern about the high costs of health care in light of scarce national and government resources has resulted in the close examination of the hospital sector of the health system. Hospitals warrant such scrutiny, it is argued, because the savings which can be generated by controlling hospital costs are much greater than those which could be generated from adjustments to the non-hospital components of the health system (Evans 1990). As a result, extensive analytic work has been done on the causes of the hospital cost spiral and on experiments and policies attempting to solve this problem.

By contrast, the attention given to health sector financing in developing countries in the 1980s tended to overlook hospitals. Policy makers have almost exclusively focused on primary health care despite the knowledge that the nucleus of the health system's capital, personnel and financial resources is located at hospitals. Figure 1 shows the hospital share of government health sector expenditures of developing countries in Asia (Griffin, 1990). Often the only acknowledgment of hospitals has been made in general statements that they consume a large proportion of the health sector's scarce resources. Few detailed studies of hospital costs are available prior to the 1980s. Only more recently, as the constraint in public revenues available for health has become more acute for developing countries, has there been increased interest in focusing on the public resources allocated to hospitals. A number of hospital case studies carried out since 1986 are examined in Barnum and Kutzin's (1992) recent review of hospital economics in developing countries.

There is also interest in dealing with the costs of operating hospitals amid indications of widespread wastage within the health sector. In the Americas, WHO (1989) estimated that wastage accounts for as much as 40 percent of the available resources for health services. The resources that could be generated from savings through efficiency gains alone are substantial due to the scale of hospital resources and operations. For example, a study in Malawi estimated that simple management correction of inefficient practices could result in a savings of 44 percent of that nation's major hospital non-personnel recurrent costs (Creese, 1990). The magnitude of hospital operational costs and inefficiencies in the health system raise questions about the utilization of resources by hospitals.

Figure 1

Government health budget allocations, by type of service



Note: Data is most recent year available from 1986 to 1989

Source: Griffin, 1990, Table A.14

The hospital sector also has need for additional resources. Developing countries have used user fees and insurance to generate revenues from the provision of hospital services. For public hospitals, the proportion of operating costs recovered from user charges has been very low, usually less than 10 percent (Mills, 1990b). There have also been concerns about the equity of access and utilization of hospital services when such schemes are introduced. Hence, the need for mechanisms to generate additional resources requires financing policies which simultaneously achieve efficiency, equity and revenue objectives.

To address these issues facing hospitals in developing countries requires an examination not only of actual hospital expenditures and their trends but of the policies and management practices which underlie and affect these trends. Economics provides a framework for such explorations, and analysis of the production and costs of hospital services can address the issues of resource allocation, efficiency of hospital operations, and the design of revenue generation policies for hospitals. This framework can help identify options for answering key questions about hospital economics and finances and facilitate understanding of the efficiency, equity, utilization and revenue implications of each. However, generalizable conclusions about hospital expenditure and resource use may be

difficult to derive because of the diversity of experience among developing countries. Thus, there is a need for detailed country level economic analysis of hospital resource allocations and costs to make short- and long-term policy, planning, and management decisions about hospitals.

The Role of Hospitals in Health Systems

In many developing countries, hospitals are not integrated into the primary health care (PHC) system (Van Lergerghe and Lafort, 1991 and WHO, 1987). The consequences of this lack of integration are evidenced in the contrasts in resources and workloads between different levels. These contrasts are experienced in many forms. For example, hospital outpatient facilities are often overcrowded due to nonfunctioning referral systems. Sometimes, the imbalance of human resource allocation among facilities and levels of the system results in the underutilization of nurses in tertiary hospitals while maternal and child health (MCH) nurses in rural health centres cannot meet immunization targets because of the overload of work. Idle, underutilized or inoperable high technology equipment litters the hallways of hospitals while health centres lack basic laboratory or diagnostic equipment due to misallocation of capital budgets. The health centre MCH teams cannot visit remote areas because they have no transport, while at the hospital ambulances are used primarily for administrative tasks rather than servicing patients.

There are numerous reasons for this situation. Some are related to the elitism of hospitals desiring to maintain the status quo and refusing to participate in the PHC approach. In other instances there have been negative reactions on the part of non-hospital elements toward the dominance of hospitals in the health system. Because hospitals compete with PHC health services for scarce health resources they are viewed as opponents rather than allies. Such views fail to recognize that hospitals perform critical functions in the health system that should complement rather than compete with PHC services. Failed attempts at cooperation have also caused deep seeded mistrust. Each of these difficulties has blocked initiation of an honest reappraisal and delineation of roles between the various levels of the health system. As a result, hospitals, with their significant level of human, physical and financial resources, have often continued to operate independently of the other elements of the health system.

When the roles of each element and level are defined in an integrated health system, it is possible to deal systematically with issues concerning the allocation of all resources--finances, personnel, and facilities. The lack of clearly defined roles and responsibilities results in a fragmented system with widespread disparities in

financial and human resources between hospitals and primary health care facilities. Despite awareness of the magnitude of health resources devoted to hospitals, the fundamental question of the role of hospitals in the health systems of developing countries has seldom been examined. The determination of the hospital role, where it has not been clearly specified, will permit changes in the health system that can improve the use of scarce health sector resources. Hence, the issue of the role of the hospital in the health system must be addressed by countries if they are to answer questions about the proper number, type and size of hospitals as well as the human, physical and financial resources which should be allocated to them. The study of hospital economics and financing can facilitate the process of role definition by providing information and an analytical framework to assist in identifying options for integration of hospitals into the health system and rationalizing the resources committed to them.

Hospital Resource Issues

Hospitals in developing countries are confronted with complex problems. The symptoms of these underlying problems include: patient overcrowding in the wards; long waiting queues in outpatient clinics; questionable quality of care; shortages of basic pharmaceuticals and medical supplies; lack of or inoperable diagnostic and treatment equipment; low staff morale; and dirty and decaying facilities. These difficulties are caused by a combination of problems including lack of a clearly defined role and relationship to other parts of the health system; organizational weaknesses; lack of responsiveness to the service population and communities; and management deficiencies including poor planning and monitoring.

All of these problems relate directly or indirectly to the availability and use of resources by hospitals. Experience indicates that there is considerable scope for improving the resource allocation and resource management practices of the hospital sector. The resource issues common to hospitals in developing countries fall within three major areas:

- ◆ **resource allocation to hospitals:** The distribution of resources to hospitals within the health sector, as well as the allocation among hospitals by type of hospital (secondary or tertiary), geographic region, and population group served (e.g. urban versus rural, vulnerable versus nonvulnerable, poor versus rich) must be examined in light of cost-effectiveness and equity concerns to ensure that the objectives of society and the

health system are being served best.

- ◆ **resource management by hospitals:** The use of resources by hospitals must be analyzed by comparing the outputs to the inputs to make assessments of relative efficiency. Such comparisons can be used to monitor performance for individual hospitals over time and across many hospitals, on a national or regional basis.

- ◆ **resource generation for hospitals:** Despite the relatively large amounts currently spent, there is little question that the hospital sector has need for additional resources. Health insurance, user charges and community financing of hospital services are the most common options mentioned for generating additional funds. The design of such systems must balance the revenue raising objective against the distributional issues of access to and equity in the use of hospital services.

Each of these hospital resource issues for developing countries is addressed below by examining the relevant economic issues and the experience of countries, and then suggesting the further work that needs to be done for successful policy development and implementation.

RESOURCE ALLOCATION TO HOSPITALS

The goal of resource allocation decisions is to maximize output and promote equity. Allocation issues are at two levels. The first, the system level, deals with the allocation of health resources to hospitals relative to other areas of the health sector and involves making choices among alternative means of providing health services. For instance, decision makers at the national level determine what proportion of government health resources go to hospitals and what proportion go to PHC. Figure 1 shows the choices made at this level by several Asian countries. The second, the institutional level, concerns allocations between hospitals. Decisions are made (usually also at the national level) concerning the portion of total hospital resources that will go to secondary versus tertiary level facilities, the per capita allocation to hospitals in region A compared to region B, urban versus rural hospitals, or between districts 1 and 2, for example.

Economic Concepts and Methodology

The tools of applied economics can be used to assist in allocation decisions and to promote efficiency within the health system. Allocation decisions must be made both within the health sector as a whole as well as more narrowly within the hospital sub-sector. Within the health sector for example, the health ministry may be faced with a choice of allocating a 5 percent increase in its budget between hospitals and primary health care. It may decide that the best use of these additional resources would be to allocate the entire increase to PHC services since this is more likely to reduce the incidence of morbidity than if these resources were provided to hospitals. Alternatively, the ministry may be faced with a choice of scale in construction in the hospital sub-sector or a choice of skill mix within the hospitals. The concepts of the production and cost functions provide a framework for consideration of such allocation decisions.

(a) Production function

A production function describes the mathematical relationship between the outputs and the quantities and combinations of inputs required to produce them. It indicates how resources can be combined to produce various levels of output in a technically efficient manner: in effect, it describes the productivity of the inputs. The current state of technology (medical art) determines how these inputs are used to produce the case weighted output. For example, an early hospital production function, used by Martin Feldstein (1967) to estimate production functions for National Health Service hospitals in Britain, specified hospital output (Feldstein defined output as weighted case mix) as a function of inputs consisting of hospital beds, medical staff, supplies, nursing and housekeeping. Subsequent production functions have improved on Feldstein's specification, but the principle remains the same. The hospital is the unit of analysis and the production function specifies the minimum quantity of inputs required to achieve a given output or, conversely, determines how output can be maximized with the given inputs.

One can also discuss, somewhat more abstractly, the production function for the health system as a whole where health status or some aggregate measure is used as the output and the various types of health services are considered the inputs. In this case the question is the allocation of ministry resources (inputs) among the various programmes or production processes (hospitals and PHC services) to achieve the greatest output (e.g. reduction of morbidity or mortality).

An example can be used to illustrate the significance of the production function concept. If 100 newly qualified nurses will soon become available, a

ministry may need to decide the number to be placed in health centres compared to hospitals. The soundness of the decision would be judged based upon its allocative efficiency. Allocative efficiency is achieved if, after deciding the number to be assigned to hospitals and to health centres, the ministry cannot increase output any further by reallocating the nurses any differently. For example, allocative efficiency has been achieved if the decision maker decides to place 40 of the nurses in hospitals and 60 in health centres because the additional hospital output gained by assigning a forty-first nurse to a hospital is less than the extra output obtained by assigning a sixtieth nurse to a health centre. Allocation decisions in this context are based on marginal analysis, that is, changes in total output. It is the marginal relationships between inputs and outputs that are important in determining the allocation of resources to different health services that will maximize total output.

(b) Cost function

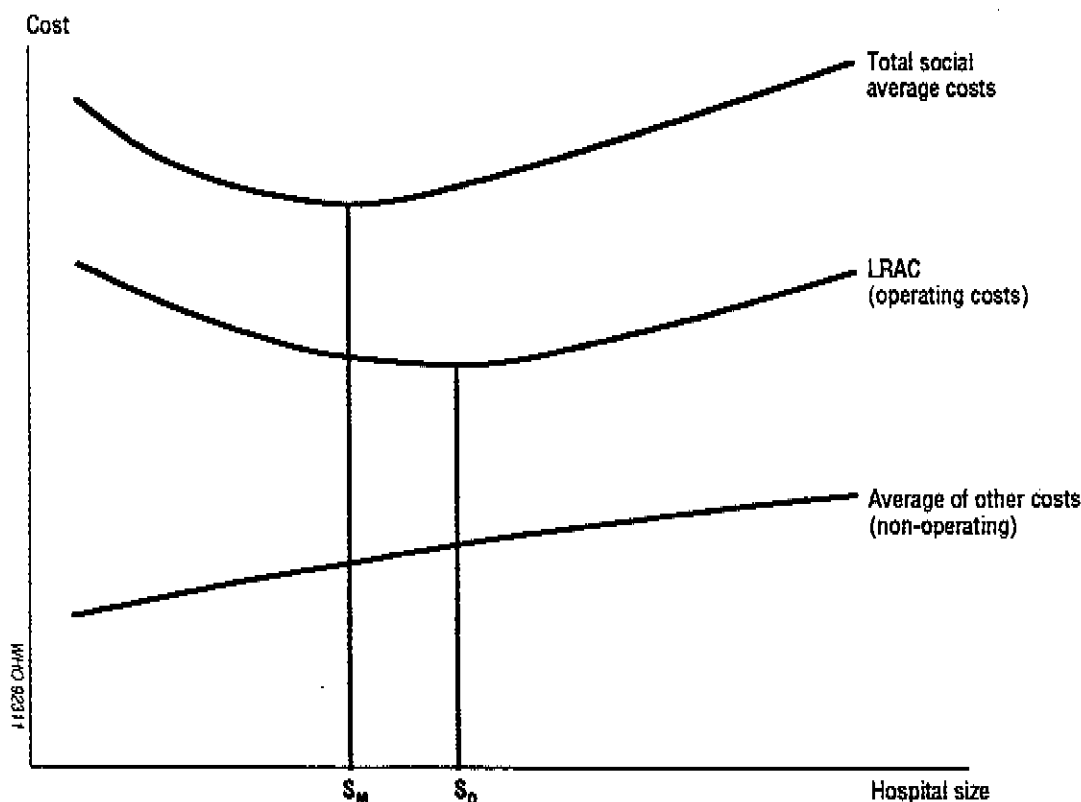
A concept closely related to the production function is the hospital cost function. This expresses the relationship between the hospital's costs and output levels. Cost functions can establish the relationship between hospital costs and size, or economies of scale. If the objective is to minimize operating unit costs, the concept of economies of scale is important for planning facilities in health systems.

Economies of scale have three possible forms: constant, increasing or decreasing. If inputs are increased in equal proportional amounts, and output increases at a constant rate, there are constant returns to scale. For example, if physician time, nursing hours, supplies, equipment and the number of beds are increased by 10 percent, and hospital patient discharges (output) increase by 10 percent, the hospital is experiencing constant returns to scale. If outputs are increasing at an increasing rate (patient discharges are increasing by 15 percent and more while the physician time, nursing hours, and other inputs have increased by 10 percent, there are increasing returns to scale (and thus decreasing unit costs). Conversely there are decreasing returns to scale if output is rising but at a decreasing rate, while the inputs are being increased in equally proportional amounts (e.g., patient discharges are increasing by 5 percent while the inputs have increased by 10 percent). The reasons for increasing returns to scale are the division and specialization of labor as well as technical factors of advantages related to larger size. Diseconomies of scale may be encountered due to the limitations for efficient management caused by the size of the facility which makes coordination and control much more difficult.

The relation of hospitals' operational costs to output are shown in a long run average cost (LRAC) curve by combining information about the costs associated with the inputs and their productivities with data on output. These relationships are shown in LRAC in Figure 2, which relates the scale of the inputs to the output or total product. The LRAC first declines, reaches a minimum and then rises. The optimal size for hospitals would be at point S_0 since that is where average costs are minimized. There are increasing returns to scale realized by increasing the size of hospitals up to point S_0 . Hospitals larger than that size will experience decreasing returns to scale beyond that point (i.e. to the right of S_0). Thus, a long-run average cost curve (LRAC) for hospitals would be useful for planning the proper scale of operations for hospitals (i.e., their size, number and distribution).

Decisions made solely on the LRAC of hospital operations would be incomplete if only the operational costs of the hospital were considered. There are patient and social costs other than the financial costs incurred by the hospitals,

Figure 2
Average cost curves



such as travel costs for the time required of patients, the social costs related to access, or ease of use costs associated with patients feeling more comfortable receiving care at facilities in or closer to their community than those at a distance. As the average size of a country's hospitals increases (in terms of hospital beds), the number of hospitals needed will decrease. This will result in other costs, related to travel time and access, increasing. These considerations are combined and reflected in Figure 2 as the "average other costs" line. The shape of the curve would be related to population distribution and may be empirically determined. When these costs are added to the LRAC for hospital costs, the result is a Total Social Average Cost curve which reflects the hospital and patient costs. The information on average costs can be used for the planning of facilities in the health system. For example, a ministry of health study may show that its district hospitals of 200 beds have the lowest operating costs, S_0 in Figure 2, but those of 150 beds, S_M , have the lowest total social average costs. This finding may result in a decision to allocate capital funds to build more hospitals of 150 beds rather than a smaller number of larger 200 bed hospitals in order to minimize social average costs.

In principle, hospital cost functions can be estimated using data from a number of hospitals, but in practice data and statistical deficiencies limit the number of functions that have been estimated. Although the number of estimated functions has been small, the concept is important in demonstrating differences between average and marginal costs and economies of scale. Resource allocation based only on average costs can be problematic because hospitals with higher average benefits may actually have lower marginal benefits when current programme size is expanded. For example, in allocating an additional \$10,000 between two hospitals, a ministry may decide to give it to hospital A, because it had an average cost per patient day of \$150 compared to hospital B with an average cost of \$175. Information from the cost function may indicate, however, that hospital A is on the decreasing returns to scale portion of its LRAC curve and can produce only 55 additional patient days. This means the marginal cost of each additional patient day is greater than the average cost of a patient day. Hospital B is on the downward slope of its LRAC curve and may produce 62 additional patient days with the additional money because it is operating where the marginal cost of each additional patient day is less than the average cost of all patient days. Hence, to make the proper allocation decision the changes in relative output (i.e. additional patient days of similar quality) must be compared with changes in inputs (i.e. additional costs or resources). The decision rule for allocating resources among programmes or activities with different benefits and costs is to select the activity or institution which has a greater marginal benefit per currency unit spent or input utilized.

Experiences of Countries

Resource allocation decisions in developing countries are made at both the aggregate health expenditure level (i.e., hospitals compared to PHC) and within the hospital sector. The latter array of allocation decisions is reflected in the distribution of total hospital resources by type, level, and geographic location of hospitals.

(a) Total health expenditure

The allocation of resources to hospitals in developing countries presents a clear picture: they receive the largest share of health resources overall. In examining total public sector health spending, Barnum and Kutzin found that for 29 developing countries the proportion of government recurrent health expenditure utilized by hospitals ranged from 25 to 81 percent and had a cross country mean of 60 percent. They noted that only four of the countries studied had less than 40 percent going to hospitals while eight had at least 70 percent of public recurrent expenditure allocated to hospitals. If this is further divided to look at government expenditure on health facilities only, there is a distinct contrast between hospital and primary care facilities: hospitals account for anywhere from 60 to 80 percent of the government expenditure for health facilities (Mills, 1990a).

However, hospitals are found to receive a much smaller share of total capital expenditure. Examination of four developing countries showed a range of 8 to 39 percent of health capital expenditure being for hospitals (Mills, 1990a). The spread of the range can be explained by the capital intensiveness of hospitals which results in the unevenness of the flow of those major outlays over time. Why is the hospital share of capital expenditure so much less than recurrent expenditure share? Several explanatory factors were identified by Mills. First, the hospital share should be lower because a much higher percentage was noted in many developing countries in the 1960s when many hospital construction projects were initiated through foreign donors. The hospitals constructed were intended to last from two to three decades. With the useful life of most of these facilities expended, capital expenditure for hospitals should surge in the 1990s if the facilities are replaced, rehabilitated or renovated on schedule. Also, priorities among donors in the past decade, as well as for the foreseeable future, have been for extending PHC coverage. As a result, donors have generally been willing to only fund capital projects related to PHC. This influx of capital for PHC tipped the scales and resulted in a lower percentage of total capital expenditure going for hospitals. Another factor is that the governments of developing countries have not had the large amounts of money required to build hospitals but have been able to build

PHC facilities since much less capital is involved.

(b) Within the hospital sector

In assessing the allocation of resources in the health sector, and in particular to hospitals, information must be gathered on budgets and expenditures for hospitals by type of facility, level of service and geographic location. This information can be used to assist allocation decisions by enabling examination of the geographic distribution of hospital expenditure and the distribution of users by socio-economic group. It could provide a basis for making decisions concerning equity in the distribution of and access to hospital services.

Geographically, there are many disparities noted in the distribution of hospital expenditure. This tends to mirror the disparity in the distribution of all health resources whether viewed from a per capita or a need basis. For example, on a regional basis in Papua New Guinea, the hospitals for one of the four regions accounted for nearly 26 percent of the country's total hospital expenditure though only 16 percent of the nation's population resided in the region, while a region with roughly the same percent of total hospital expenditure had more than double the population, (Newbrander, 1987). In developing countries, in general, larger and higher level hospitals at the central and regional level account for 60 to 80 percent of hospital expenditure with the remainder expended for district hospitals (Mills 1990a). Some disparity in allocation is justified based on factors such as health status, dispersion of population, and difficult topography. Each of these would increase the costs of providing the same level of health services in other areas of the country. A suitable means for determining the appropriate allocation of health and hospital resources that considers such factors has not been found for developing countries, though some developed countries (e.g. the United Kingdom and Spain) have developed such formulas to determine allocations.

Studies of economies of scale (how size affects average costs) for hospitals in developing countries have shown that diseconomies of scale or at best constant returns to scale were noted as hospital size increased (i.e. a flat average cost curve), although the data are sketchy and the number of studies small. Barnum and Kutzin noted that cost functions necessary to determine economies of scale had been derived for hospitals in only five developing countries. Three of these studies had been undertaken as research specific for their book. In one of these, the hospital cost function was estimated for Ethiopia in order to determine marginal and average incremental costs (Bitran-Dicowsky and Dunlop, 1989). The marginal cost of an inpatient day was not significantly different from average costs. It was concluded that the hospitals were at the point of constant returns to scale for

inpatient days, laboratory exams and delivery outputs, so no major changes were called for in the planning of hospital size for these services.

Directions for the Future

Improved allocation of resources for hospitals in developing countries will require that there be analysis of economic data concerning hospitals and an understanding of the underlying economic concepts, processes and interactions. In addition, decision makers will have to be familiar with the national budget process and political context if the outcomes of planning are to be realistic and viable. Several countries have found their national health plans to be unrealistic in the programmes outlined and the finances planned for hospitals because they were made in isolation without consideration of the political environment or the involvement of the planning and finance ministries (Thomason and Newbrander, 1991a).

The review of the information on the allocation of resources to hospitals in developing countries reveals a paucity of data on health and hospital expenditures. In addition, very little effort has been devoted to understanding the allocation of resources within the political and epidemiological environment. There will have to be the development of methods by which resources can be allocated based on the existing situation and national priorities. This type of analysis will require examination of several categories of resources: financial, capital and human. The decisions made about these will have short and long range implications for the health system. The most obvious consequence of this process is the development of norms or targets regarding the proportion of health sector recurrent and capital expenditure which should be utilized for hospitals compared to other facilities and programmes. This will influence the direction of the health system in the long term.

Little study has been done on how the planning of the allocation of hospital resources influences access to and equity of hospital services. For instance, the number, location and type of hospitals financed have an effect on the population's ability to use those services, and the degree of specialization will define how large a population is served (e.g., the more specialized the services financed, the fewer the population groups which benefit from those services). These factors can and should be included in allocation decisions if they are specified as policy objectives for the health system.

An understanding of the relationship of costs to the size and structure of hospitals is required to plan the number, size and type of hospitals. Developing an

information base, whether from accounting based studies or estimated cost functions, is important in light of the number of countries that will begin replacing their hospital infrastructures over the next few years. Such decisions and investments must be made wisely. The use of economic information on hospital costs of operation is essential not only in making capital investment decisions but also for taking into account the recurrent cost implications of these decisions over the long term.

RESOURCE MANAGEMENT BY HOSPITALS

Poor management of hospitals leads to wastage of resources, including money, staff, buildings and equipment. Such wastage means the existing level of service or output could be achieved with far fewer resources. Alternatively, the same level of resources could provide a greater quantity or range of services. Savings from reduction of this waste would result in resources being available for services other than those provided in hospitals (e.g. PHC) or for expansion of access and improvement of the quality of hospital services. Although there have been few, if any, studies clearly identifying the magnitude of inefficiencies in developing country hospitals, many developing countries have identified better management of hospital resources as one of their major needs (WHO, 1987 and 1991).

Economic Concepts and Methodology

The term "efficiency" has become common place in health ministries, in the literature and with international donors and agencies. The issue customarily addressed by use of this term is the question of whether health expenditures are higher than they need to be because the cost of producing services is not minimized. For example, a regional hospital which has average costs of \$330 per admission compared to \$300 for two other regional hospitals may be identified as an inefficient hospital. However, the use of unit costs to compare individual hospitals requires that the products of the facilities be similar, in terms of the type of patients treated (the case mix) and the quality of care delivered. If the regional hospital in question is treating patients with much greater health problems (i.e., a more complex case mix), then the greater unit costs do not imply wastage of resources because its output is qualitatively different (i.e., treating sicker patients). However, if the other regional hospitals have lower unit costs because they are treating patients and conditions which could be treated at a district hospital which is less costly, they could be, in fact, the inefficient facilities. To make such direct

efficiency comparisons (the input/output ratio of hospitals) requires that the input and output units of measurement be the same (or adjusted to provide comparable units) for all hospitals being compared. There are three types of production efficiencies: technical, economic and scale.

(a) Technical efficiency

Technical efficiency deals with the mix of inputs which will produce a given output. The basic question is, given the health personnel, supplies and equipment employed by the hospital, is it obtaining the highest output possible? Alternatively, given the number of services produced, is the hospital employing the fewest possible people, supplies and equipment? There can be several technically feasible solutions to producing a given output. For example, to "produce" an inpatient stay (the output), the clinical case could be managed by using various combinations of inputs of physician time, nursing care, diagnostic services, and hotel services. The inpatient episode may use a longer stay in hospital with more nursing care and diagnostic services and less physician time, or it may make more intensive use of physician and therapeutic services with a shorter stay and less total nursing hours employed to treat the patient. Any of these possible means of managing the clinical case are technically efficient if they achieve the same desired output; in this case, a completed inpatient episode with an acceptable outcome (remedying the health problem).

An example of technical inefficiency would be the inappropriate over prescribing of pharmaceuticals to outpatients. This is technically inefficient because the inputs used (drugs) were greater than were necessary (more drugs than were medically required) for successfully treating the patient. Technical inefficiency also deals with achieving a smaller output than is possible with the given inputs (e.g., fewer outpatients seen than possible with the given staff, supplies and examining rooms). The result is an unnecessary increase in the cost of producing the output since more inputs are utilized than is necessary to achieve the same output level.

(b) Economic efficiency

Economic efficiency deals with the least-cost combination of the inputs which will produce the desired output, given various possible combinations of inputs. Given the hospital's fixed budget, is output (e.g. patient days, admissions, surgeries, etc.) maximized? Or alternatively, given a fixed quantity of services to be produced by the hospital, is total cost minimized? For instance, it may be found that to produce one patient discharge, the least costly solution for the hospital is to use less physician and nursing care time (costly inputs) and more laboratory and

diagnostic services (a less costly combination of inputs). This assumes there are various combinations of inputs which will achieve the same qualitative output. These may or may not be technically efficient, as will be shown below. As a rule, when the term "efficiency" is used it is economic efficiency (the least cost solution to achieve a given output) to which reference is being made.

(c) Scale efficiency

Scale efficiency deals with whether a system as a whole is producing services at least cost. Economies of scale are important not only for planning the size and number of hospitals from a macro planning point of view, but also for the individual hospital since the size of the facility will have an impact on its operating costs and the efficiency that can be achieved in the short and long run. In the short run the hospital's bed capacity and facilities are fixed. This limits the range of input combinations (staff, supplies, equipment, buildings) that can be used to produce the desired output and, thus, limits the efficiency that the hospital can achieve in the near term.

The long run is the time period where all inputs, including the size of facilities, can be varied. Thus there is a greater range of combinations of inputs to produce a given output. Scale efficiencies are achieved as size increases by spreading large fixed costs over a greater number of patients and by the greater degree of specialization among staff and departments which can be achieved through the divisibility of functions. Scale diseconomies are encountered when the size of the facility begins to make management and control of the hospital as an organization more difficult. For example, it may be determined that the optimal size for a district hospital is 140 beds. If one district is operating a 200 bed hospital and another is operating a 80 bed hospital, both will be subject to less than maximum efficiency. In the first case it is because decreasing returns to scale have been encountered while the second district hospital could take advantage of further increasing returns to scale if it were of a larger size. The total costs for hospital services is not minimized in either case.

(d) Efficiency interrelationships

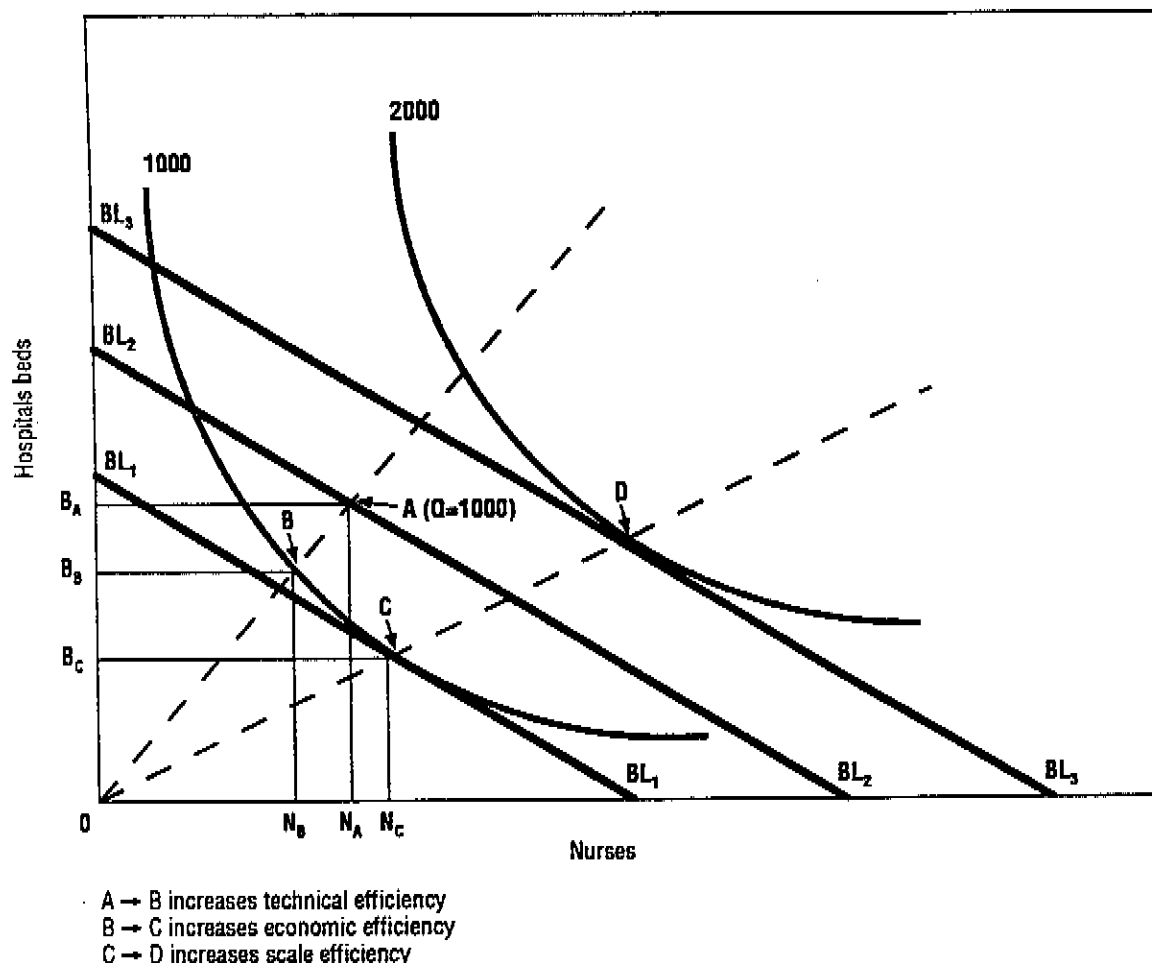
Understanding the interrelationships among these different types of efficiency is important because the solutions to efficiency problems will vary according to the type of inefficiency facing the manager. For example, consider two hospitals. Both have laboratory services with the same equipment. In one hospital technicians perform laboratory tests manually despite having autoanalyzers, while in another autoanalyzers are used and physicians prepare and run the tests and report the

results. The first hospital has a problem with technical efficiency because it is not combining its equipment and staff in the right proportions to do the tests. The second hospital is technically efficient in using staff and equipment in the right proportions but is economically inefficient because the costs to do the tests would be lower if there was labor substitution by using a technician rather than a physician to run the tests. If one decides to increase staff levels at the first hospital in order to run more tests, the solution chosen will magnify the technical inefficiencies rather than solve them. There is also scale inefficiency in the system because neither hospital is operating with the correct or least cost combination of inputs for operating a laboratory service. Thus, it is necessary to understand that the inputs must be combined in the right proportions, as in the second hospital, and that the least costly combination of inputs must be used, as was done in the first hospital (but which did not use the equipment and staff in the right proportions). Understanding the underlying differences in types of efficiency is necessary to achieve the objective of using the most economically "appropriate" or productive input combination.

The interrelationships between these types of efficiency are illustrated in Figure 3. This graph is a simplification of the hospital production process because it assumes only two inputs are needed, nurses and hospital beds, to produce the hospital's output, patient days. Each curve is termed an "isoquant" and depicts the many different technically efficient combinations of the two inputs which can be combined in varying proportions to produce the same number patient days of hospital care. For example, each point on the 1000 isoquant represents a different combination of nursing care and beds which will produce 1000 patient days of hospital care. Point A, though not on the 1000 isoquant, represents another combination of the two inputs which also produces 1000 patient days of care. Point A is not technically efficient because it uses more nurses and beds to produce 1000 patient days than is necessary (i.e., there are points on isoquant 1000 that produce the same output but use less inputs, and a greater output is possible using the combination of inputs defined by point A). Moving from point A to point B represents increased technical efficiency because the same output is produced at B as at A but with fewer inputs used.

The straight lines in Figure 3 are termed budget or isocost lines. They represent the various combination of inputs (nurses and beds) which can be purchased with a fixed total budget. Because any point on a given budget line, such as BL₁, shows the various combinations of nurses and hospital beds which could be purchased for the same total cost or fixed budget, budget lines are straight. The further out budget lines are from the vertex the greater the total budget to purchase those combinations of inputs. For hospitals, this would mean

Figure 3
Technical, economic and scale efficiency



the further out their budget line the greater their fixed budget. Thus, the combinations of inputs on budget line 1, BL₁, all cost the same total amount but this is less than the total cost of the various combinations of inputs purchased on budget line 2, BL₂. And the total cost of any combination of inputs shown on budget line 3, BL₃, is more than the cost of nurse and bed combinations shown on BL₂. The relative prices of each unit of input determines the slope or angle of the budget lines.

Movement from point B to point C reflects improved economic efficiency. While the output is the same for both, 1000 patient days, the budget line passing through B is further out than the line passing through C, reflecting a more expensive combination of inputs (or greater hospital budget) to produce the same number of patient days. The tangency of point C with the isoquant 1000 indicates it is the least cost combination of nursing time and hospital beds which can produce 1000 patient days. The movement from point C to D reflects an increase

in scale efficiency because the hospital would be operating at increasingly higher output levels at the points of tangency between various isoquants and budget lines.

Experiences of Countries

(a) Cost function estimation

Barnum and Kutzin report estimated hospital cost functions for five developing countries (Kenya, Ethiopia, Nigeria, Columbia and China). Their basic conclusion is that the variability of the results indicates the caution that should be used in estimating and interpreting such functions. The country specific context within which hospitals operate must be considered. The authors cite an example from China where the basic data results showed great inefficiencies since high occupancy rates were maintained despite marginal costs exceeding average costs. However, closer examination of hospital service statistics revealed extremely long lengths of stay. So rather than expand hospital size or the number of facilities to reduce occupancy rates, a reduction in the length of stay was recommended to expand capacity and also lower unit costs.

Other findings were that in general there were constant or decreasing returns to scale. The implications of these findings would be that the hospitals constructed in developing countries should not be too large. In addition, findings of diseconomies of scope (i.e., increasing marginal costs due to the range of services offered) indicate that the large hospitals, usually in urban areas, should not expand their outpatient facilities. Perhaps urban primary health facilities, satellite clinics and outpatient surgeries should be considered rather than expansion of existing hospital facilities to meet demand. Overall, Barnum and Kutzin believe the limited data suggest that there should be increased interest in strengthening the capabilities of the first-referral level hospital to provide better integration into a referral and support system for hospitals.

(b) Average operational costs

Comparing the outputs of hospitals requires basic workload indicators to be determined so they can be combined with financial information. The common measures of hospital inpatient output or workload are the numbers of patient days, admissions (or discharges), and hospital beds. These measures can be used to generate performance indicators--the average length of stay, the occupancy rate and the bed turnover rate. These indicators were used in addition to hospital cost analyses to form the basis for Barnum and Kutzin's review of efficiency of hospitals

in developing countries. However, they recognized that their conclusions were only suggestive because neither the cost studies nor the inpatient performance indicators adjust for variations in case mix and quality across hospitals.

Studies across a number of countries have confirmed that the hospital costs of operation are much greater at tertiary care facilities than at secondary facilities. Barnum's study of Indonesian hospitals viewed the inpatient costs on a per patient day, admission and bed basis. Using the district hospital as the base, the provincial and tertiary hospitals costs per patient day were 2.1 and 2.7 times greater, respectively (Barnum and Kutzin, 1992). Similarly, Mills (1990b, p. 208) concluded from her review that "inpatient care in general hospitals can be up to twice as expensive as that in district hospitals, and in central hospitals between two to five times as expensive as in district hospitals."

Differences in the cost of an outpatient visit among the different levels of hospitals were of the same magnitude as for inpatient cost differences. For example, Barnum's study in China found that outpatient costs in tertiary and provincial level hospitals were 2.0 and 1.3 times greater, respectively, than those in district hospitals. Many of the cost differences for inpatient and outpatient services are attributed to a number of factors, such as the greater resource intensiveness of services, the complexity of the case mix of hospital patients and the teaching function incorporated into the hospitals' mission. On a per bed basis, therefore, larger hospitals cost more to operate and take a larger share of the resources allocated to hospitals.

(c) Technical efficiency

Technical inefficiency may result from lack or improper use of inputs in hospitals: personnel, supplies and medical equipment. For example, rigid staff assignments which prevent hospital personnel from performing different functions can result in technical inefficiency. If hospital radiology technicians who are not fully occupied with their duties are not redeployed to do other tasks for which they are qualified, there is a lack of productive use of all of their available work hours. With medical equipment, a China study indicated diagnostic (CAT scanners) and treatment equipment (dialysis and ICU equipment) were unused or used inappropriately (Barnum and Kutzin, 1992). In some cases the equipment, such as ultrasound, might have provided a useful diagnostic tool at a lower level hospital, but the lack of trained technicians or the skills to interpret the results meant that it was not used properly. Some of the hospital equipment problems in developing countries emanate from donors who contribute equipment that is inappropriate, does not include proper training for staff, cannot be maintained due to costs or

availability of spare parts or technical support, or is incompatible with existing hospital equipment. Poorly designed or managed medical maintenance programmes can render hospital equipment useless.

Pharmaceuticals is another area where considerable technical inefficiency occurs. Barnum and Kutzin identify both clinical and administrative management deficiencies as sources of this inefficiency. Clinically, there may be inappropriate prescribing of drugs to outpatients and inpatients (i.e., the wrong drug may be prescribed or improperly prepared) so as to render it useless or harmful to the patient. There may also be unnecessary prescribing of multiple drugs. Over- or under-prescribing may result from the wrong strength or quantity being prescribed to patients. The management system may be a cause of pharmaceutical technical inefficiency problems as well through such problems as pilferage of drugs, wastage due to improper storage, or inadequate procurement procedures. In Mali, for example, savings of up to 40 percent of existing pharmaceutical expenditures could be achieved through better management practices (Creese, 1990).

(d) Economic efficiency

The hospital administrator may operate a technically efficient process, but it is not economically efficient if the combination of inputs used are not the least cost means, as shown by the contrast between points B and C in Figure 3. Input substitution and improved management techniques can often improve the economic efficiency of hospitals.

Input substitution achieves economic efficiency when less costly inputs, which are equally effective, are substituted for more expensive inputs. Barnum and Kutzin discuss two major categories of input substitution: among different types of health workers (e.g. nurses for physicians) and between labor and labor saving equipment and pharmaceuticals (e.g. additional diagnostic equipment for physicians). Because labor substitution relates to the roles that health care providers assume, its feasibility is determined by the flexibility hospitals have for substituting one category of health worker for another to perform certain functions. It is often a country's professional associations or civil service regulations which restrict this flexibility for hospital managers.

Management procedures, such as staff scheduling, can improve efficiency. Examples include shifting underutilized staff from wards with few patients to those that are busy and improving work flow, such as having the next surgical patient prepped and the surgical suite ready with instruments so as to eliminate idle time for surgeons and nurses in the operating theatre. Some countries are attempting

to increase staff productivity by establishing norms for the staffing of facilities and evaluating their output based on workload (Kolehmainen-Aitken and Shipp, 1990).

In Indonesia, Fiji and Jamaica the larger hospitals have a greater concentration of physicians and other health workers per bed (Barnum and Kutzin, 1992). This could be due to higher occupancy levels, a more complicated case mix treated by the hospital, teaching responsibilities or inefficiencies arising from lack of substitution or misallocation of workers among hospitals. Because the case mix, intensity and quality of services of the larger hospital is assumed to be greater than district hospitals, however, one cannot assume the hospital with higher staffing levels per bed is operating less efficiently. The assessment of relative efficiency among different hospitals must adjust for differences in case mix, quality, and teaching function.

Another means of improving the economic efficiency of the health sector would be to search for alternatives to delivering care on an inpatient basis that are lower in cost yet achieve the same outcome. One example is in Cali, Columbia, where satellite day surgery units for minor surgical procedures were established as extensions of the hospital as a less costly means of providing the service. Shepard et al. (1991) found that the average cost of performing an inguinal herniorrhaphy on an ambulatory basis was less than one-fourth that of the same procedure performed for hospital inpatients, and clinical outcomes were better for the ambulatory patients.

Pharmaceuticals or medical equipment may be substituted for other inputs such as nursing time, which may result in fewer hospital days per patient. Improved laboratory or radiology equipment can increase productivity of staff so they have a greater output. Pharmaceuticals may also reduce the number of return outpatient visits to the provider. Poor pharmaceutical ordering systems may necessitate the substitution of more expensive or less efficacious drugs for needed drugs that are out of stock.

A key to achieving basic management improvements that increase efficiency is tailoring the incentives under which hospitals operate. In many countries hospital managers have been very restricted in their ability to reallocate their budgets among different line items. If they achieve savings in the use of supplies or pharmaceuticals, their only option is to purchase greater quantities of these items since savings can not be reallocated to other budget lines. Unexpended funds revert back to the central government treasury at the end of the fiscal year. Thus, there is no incentive for managers to make efficiency improvements since they and their hospital receive no direct benefit from their actions.

(e) **Scale efficiency**

As mentioned above, analysis of the hospital cost functions reviewed by Barnum and Kutzin suggests that in developing countries there should be more hospitals of a smaller size due to the constant or decreasing returns to scale found in the countries studied to date. These findings are consistent with those for developed countries. Carr and Feldstein's study (1967) in the U.S. found that the optimal size for minimizing costs was 190 beds. Their general conclusions were that scale economies are present but are not large (a shallow LRAC curve); costs are higher as complexity of service offered by the hospital rises; and diseconomies exist only for the largest hospitals (greater than 300 beds).

More recent studies have found constant or slightly diminishing returns to scale, especially for larger hospitals. In their survey of developed country cost functions for hospitals, Cowing, Holtmann and Powers (1983, p. 276) concluded that moderate and large hospitals can be characterized by constant returns to scale. Such information on scale efficiencies is useful for planning the optimal size, number and distribution of district and regional hospitals.

Directions for the Future

Hospital unit costs are derived by relating a hospital's inputs to the outputs produced over a certain period of time. It may be done through a detailed analysis of a single hospital or for a group of hospitals using statistical techniques. Estimates of the relation of hospital costs to service outputs can be important for not only assessing efficiency but also for decisions concerning the financing of hospitals and pricing of their services. The data can be used to determine average and marginal costs by combining the total cost information with various types of output such as admissions, patient days, and outpatient visits, or by type of patient (surgical, medical or by diagnosis). The information, if it covers a sufficient number of hospitals, may also give some indication of the extent of economies of scale. This is important for planning decisions about the size of new hospitals compared to the option of expanding existing hospitals. However, the developing country studies and data on hospital average costs are extremely limited. The cost information which has been available has been at an aggregate level only. There is a need for better financial and cost data about hospital operations if their efficiency and management is to be improved.

First, hospital financial data should be gathered and included in the routine information system of hospitals in developing countries. Accurate financial data

concerning the total cost of hospital services is needed. This goes beyond merely examining the hospital's allocated budget categories to determining the actual total costs of operation. For example, capital costs or maintenance of buildings or vehicles, which are carried out by other ministries, must be costed and included to determine the total costs of hospital services. Extensive work must be done with countries so they may develop the appropriate methods for such studies as well as proper financial information systems.

Second, there is a need to have adequate measures of hospital output to be combined with the cost information. This is needed so marginal costs, average costs and economies of scale may be determined. Output measures that adjust for case mix and quality differences will enable the efficiency of hospitals to be assessed by evaluating and monitoring each hospital's performance relative to other hospitals.

Third is the need for training of hospital managers in the use of this information for the day-to-day operation of their institutions. Cost analysis can be used for planning hospital services and capacity as well as for establishing appropriate fee levels under cost recovery schemes. The use of input and output information can serve as a management tool to identify deficiencies of various types in the hospital sector or simply problems which arise within a single institution concerning the costs of production. This would be extremely useful for a regional or national hospital manager to review and compare the operation of various hospitals in terms of their efficiency. Lack of data on inputs and their costs as well as outputs may be given as a reason for hospital managers being unable to make assessments of their efficiency. Very basic analysis can occur, however, with existing information, but it is often the lack of managerial skills on the part of hospital administrators and managers in how to use the data which is the barrier.

Finally, there is the need to study and better understand the incentives under which existing health system and hospital structures operate and their effects on hospital efficiency.

RESOURCE GENERATION FOR HOSPITALS

Many developing country hospitals have relied almost exclusively on government financing of recurrent and capital costs for hospitals. The Papua New Guinean government, for example, provides 97 percent of total hospital recurrent costs (Thomason and Newbrander, 1991b). Similar patterns exist in many other countries. Despite the need for additional finances for hospitals, the future

availability of existing levels of public finances are more uncertain now than ever before, as illustrated in Indonesia (Prescott, ed., 1991). As a result of economic realities and the limitations of government financing, many countries have recognized the need to find alternatives to public general revenues for financing health services. User fees and health insurance are the most common options mentioned for hospital services.

Economic Issues and Methodology

The introduction or modification of user charges or insurance for hospital services would mean a change in the price of hospital services facing users. Changes in price will affect demand and utilization of those services. The extent of the effect is related to the consumer's elasticity of demand. Elasticity measures the change in the quantity of the good or service demanded with a change in its price. If a user's demand for hospital services is elastic, an increase in the price of hospital services will result in a larger percentage decrease in his or her demand. Conversely, inelastic demand by the user means that the decrease or increase in demand will be less, in percentage terms, than the increase or decrease in the price of hospital services.

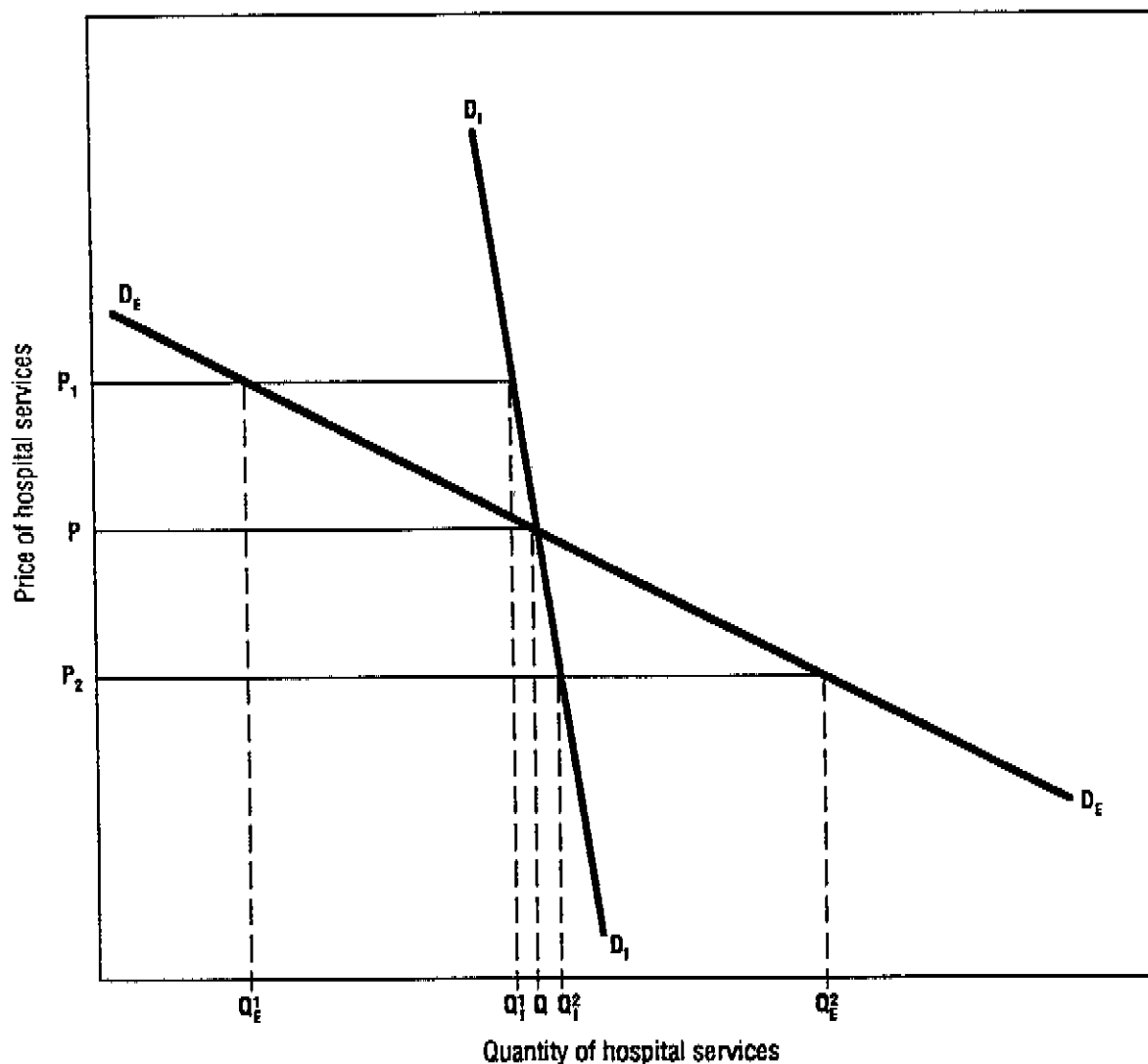
Aggregate changes in demand resulting from user fees or insurance will have implications for the revenues generated for the hospital as well as for equity and efficiency. For example, if demand for hospital services in a poor district is very elastic, small increases in fees may result in a large drop in demand, and the additional revenue gained may be very small. Do the fees provide an efficient means of generating revenues? Equity relates to whether the fees result in the poor being inhibited from receiving necessary services or whether adjustments are made based on ability to pay. Efficiency issues include reductions in demand for unnecessary care and whether the charges bias the selection of inputs chosen.

(a) User charges

The primary justification provided for introducing user charges for hospital services is their ability to achieve efficiency, equity and revenue goals (Griffin, 1988). There are possible efficiency gains because of the incentives for improvement in the way in which hospital services are utilized. Equity may be enhanced by cross-subsidizations to improve utilization by low income groups and correct any regressive elements in current fee structures or facility locations. Likewise, fees may have adverse effects on equity by creating access barriers for the poor.

The implications of basic price changes on user demand for hospital services are shown in Figure 4. The demand for hospital services may be elastic, as shown in the demand curve D_E , or inelastic, D_I . The initial equilibrium is where price P is charged and quantity Q is consumed. If demand for hospital services is inelastic (i.e., considered a necessity which cannot be foregone) as shown in the demand curve D_I then a price increase to P_1 will result in a very small drop in quantity demanded to Q_1^1 . The same price increase for consumers with elastic demand (demand curve D_E) will result in a large decrease in quantity of hospital services demanded to Q_E^1 . Conversely, a price decrease to P_2 will increase demand only slightly to Q_I^2 for those with inelastic demand and to Q_E^2 for those with elastic demand.

Figure 4
Elasticity of demand



WHO 92313

In addition to the issue of how user fees influence demand, there are efficiency concerns in terms of the administrative ease with which fees can be collected and equity exemptions applied. The incentives for the hospital to be efficient in fee collection must also be considered. For instance, if government hospitals are expected to collect fees but retain only a small portion of those collected with the balance of the fees being returned to the central treasury, there is little incentive for hospital managers to ensure that their clerks are collecting fees to the greatest extent possible.

(b) Insurance

Insurance is a means for spreading the risk for the high cost of hospital services due to the unpredictability of the need for hospitalization of any particular individual. Designing an insurance scheme that meets certain social welfare objectives is extremely complicated because of the incentives created on both the demand and supply sides of the market for hospital services. For example, an insurance scheme which pays total charges for all inpatient and outpatient services will result in those who are insured demanding more services than if they had to pay a portion or all of the charges. It may also encourage physicians to create demand in their patients in order to increase the fees they may collect from insurance (e.g., physicians may prescribe an insufficient quantity of drugs which will require patients to make numerous follow-up visits for the same episode in order to increase their level of reimbursement). Insurance, by reducing the costs of services to individuals, may also induce behavioral change that results in higher demand for services. This phenomenon is termed "moral hazard".

Certain mechanisms, such as deductibles, copayments or limits on coverage, are usually incorporated into insurance schemes in order to counteract moral hazard and deter unnecessary demand for hospital services. Each mechanism has specific equity and efficiency implications and affects hospital revenues as well. A deductible, for instance, requires that the insured pay the first set amount of hospital costs, and after these are met the remainder of costs are covered by insurance. The advantage of this is that it provides full coverage of high cost events which have a low probability of occurring. The disadvantage is that it is not flexible enough to deal with moral hazard. If it is too small, the deductible will not dissuade unnecessary use; but it would hinder access to necessary care if it is too large.

Coinsurance requires the insured to pay a fixed percentage of all hospital expenses from the first dollar incurred. This is more flexible for dealing with the issue of moral hazard and should prevent premiums for insurance from becoming

excessive as a result. The disadvantage is the inefficiency associated with having to deal with frequent, small losses. In addition, a fixed coinsurance rate may result in large out-of-pocket costs for high cost services.

With limits on coverage, the insurer pays all costs up to a defined maximum, after which the insured pays all costs. This shifts the burden for catastrophic illnesses to the insured in an attempt to reduce the cost of insurance, even though the additional premium to cover such illnesses, which are characterized by high costs but a low probability of occurrence, would be small. This puts the very small number of families which exceed the limit in a precarious position. Key features in the design of insurance programmes are the degree of consumer choice to allow regarding the risks covered (to choose whether to insure for events with low or high probability), the magnitude of the loss covered, the degree of coverage (the specific events and services covered by insurance), and the size of the premium, deductibles and coinsurance rates.

The design of the insurance scheme will also create incentives for hospitals as well as the insured. For instance, insurance which covers only inpatient services would result in the insured demanding inpatient care for minor surgical procedures which could be done equally effectively on an outpatient basis because it would reduce the cost of the procedure to the user. Or if the payment by insurance for inpatient episodes is a fixed amount regardless of the procedure or the length of stay, there would be a bias for hospitals to minimize the length of stay for such covered patients or to not accept complex cases which would have a higher cost. In general, the more comprehensive the insurance, in terms of services covered, the less distortion introduced into the choice of treatment.

Experiences of Countries

(a) User charges

The evidence on the use of fees in government hospitals to generate revenue or the effects of user fees on hospital utilization is limited. Review of existing experiences concerning user fee collections revealed several basic trends: public hospital cost recovery was low though the experiences varied widely, even within countries. Secondly, the mission and private sector hospitals recover a much greater proportion of their recurrent costs through fees. Information on the impact of user fees in terms of utilization, access and equity, however, is limited.

The structure of hospital user fees varies from a general fee to cover a

hospitalization episode, to daily charges, to fees for specific services such as diagnostic tests or pharmaceuticals. Wide variations exist in the amount of hospital recurrent costs recovered through user fees across countries, but most public facilities recover only a small amount. Mills (1990b, p. 211) summarized her literature search by stating: "It seems safe to conclude that, for public hospitals, charging practices vary greatly between countries and that it is very unusual to recover more than 20 percent of hospital expenditure: most hospitals recover far less." The fees that are charged are most often found to be substantially below the unit operating costs of the hospitals (Barnum and Kutzin, 1992).

A study of public hospital fee revenue in three Latin American countries found the portion of the recurrent budget covered by fees (excluding the value of drugs and supplies provided to hospitals outside their budget) ranged from less than 1 to 8.5 percent (Lewis, 1990). This low level of hospital expenditure recovered through fees is accounted for by several factors: poor administrative collection mechanisms; lack of incentives for hospitals to collect fees; static fee schedules; fee exemptions for many patient groups; and abuse of the system. Changes to any of these factors could generate a significant amount of resources which are not currently realized. Often, however, all or a large proportion of the fees must be turned over to the general government treasury (Barnum and Kutzin, 1992). There is no incentive for hospitals to improve fee collections unless they are permitted to retain all or most of the fees collected.

Honduras attempted to overcome this difficulty by decentralizing authority for disbursement of the fees collected. Revenues from fees were used to purchase supplies and drugs and to pay for personnel and building maintenance. The incentives from this scheme resulted in the hospitals' fee collections growing 8 percent per annum over a three year period (Overholt, 1987). Yet the proportion of operating costs covered by the fees collected remained small. They represented 3 and 5 percent of operating costs for the large national hospital and district hospitals, respectively. These recovery levels were far below the target of 30 percent of hospital costs set by the government. Some of the difficulties were attributed to there being few guidelines established on the services for which fees would be collected, the amount of fees, the basis of fees (per admission, per episode, per service, etc.), exemptions which would apply or the allowable uses of the fees collected. Though Overholt felt the exemption mechanisms had ensured access for the most disadvantaged, there were no data to determine the actual effects on access by income groups.

Mission hospital facilities often have a much higher proportion of their costs recovered through revenues from the community. Among various denominations in

Africa, the proportion of income derived from user fees ranged from 27 to 70 percent (Mills, 1990b). Though these are small scale endeavors and provide only a small proportion of national hospital revenues needed, the apparent higher level of willingness-to-pay by patients and families for hospital services from missions is due, in part, to the community's perceptions that the quality of care is higher in private and mission facilities than in government hospitals. In addition, the community can more often identify with mission hospitals as their community facility than is the case with public hospitals. These observations do raise interesting questions of how public hospitals will have to adapt if they expect more revenues to be generated from user fees.

The low levels of revenue collections from user fees for hospitals in developing countries reflect two related but separate issues. First, the fees collected represent a very small percentage of hospital operating costs. This problem is a result of pricing policies in which the prices set are not related to the real costs of providing the service nor are prices updated to maintain fee levels in real terms. For example, in Papua New Guinea the fee schedules for hospital services were set in 1978 and were not changed for over ten years. As a result, the fees decreased by more than 50 percent in real terms. Only recently have fees been increased, but the problem will recur because the new fees have not been indexed to inflation. Few countries regularly increase fees to keep pace with inflation; consequently, fee schedules quickly become outdated.

The second issue concerning user fee problems is that actual revenues from fees collected are much lower than the potential or expected revenues from fees if all patients utilizing services had paid for services for which there are charges. This is a result of problems with exemption policies, administrative procedures for collection, and collection incentives. For example, the guidelines provided by government on those categories of patients that are exempt from fees are often multiple and confusing. There are exemptions for certain categories of patients, but it may be difficult to verify a patient's status for exemption, or there may be pressures on the fee collectors to include the noneligible in the exemption category. Thus, the application of exemptions is not uniform and tends to err on the side of permitting numerous nonexempt patients to receive care or pharmaceuticals without payment of the required fee.

In many countries poor administrative methods and guidelines result in low fee collection levels because there are few mechanisms in place for monitoring, control and auditing the fees collected. Poor administrative procedures and supervision have led to lax application of fee collection guidelines, noncompliance with exemption regulations, confusion as to the types of services for which fees are

to be charged or pilferage of funds.

Another problem causing low fee collection levels is the lack of incentives for hospitals to increase cost recovery efforts by collecting fees. As a result, hospital managers do not closely supervise and monitor fee collection efforts. The case in a number of countries is that the revenues revert to the general government treasury or the hospital is permitted to retain only a very small proportion of the fee. Vogel (1988) found that the greater the proportion of fee revenues reverting to central or regional government the lower the amount collected relative to what should have been collected. Barnum and Kutzin suggested that some proportion of fees collected should be retained at the local hospital level to encourage efficiency in fee collection, but the appropriate level is a matter requiring empirical study.

(b) Insurance

Private and public insurance schemes, including social security, exist in a number of developing countries. Private health insurance is becoming more prevalent but covers only a small segment of the population. Kutzin's study of Jamaican health insurance (Barnum and Kutzin, 1992) found less than three percent of payments going to public hospitals. The public hospitals were heavily subsidized, and thus the financial exposure of patients for hospitalization was minimal, as evidenced by the low percentage of public hospital revenues generated by fees. Eleven percent of claim amounts was for private hospitals. He thus concluded that the motivation for health insurance is to provide access to private physicians and pharmacies. Often insurance schemes develop to allow a certain group (such as civil servant employees or other employment groups) to utilize non-public hospital facilities, such as the public servants health insurance scheme in Papua New Guinea. These are viewed as preferable to public hospitals because of higher quality of care, better accessibility, shorter waiting times, or longer and more convenient hours of operation.

Health insurance under social security may provide hospital care at facilities operated by the social security organization, such as in Turkey or some Latin American countries, but more commonly it merely finances care received at public or private hospitals. When hospital services are provided by the insurance programme's own hospitals, the revenue for hospital operations comes from the insurance scheme's collection of prepayment premiums. General government revenues may also be used to subsidize the hospital's operation. Social security based health insurance programmes are widespread in Latin America.

McGreevey (1990) noted a number of equity (access) and efficiency (duplication of services, diseconomies of scale) problems with the social security schemes he studied. The inequities often arise because of the dual health systems. In Columbia, for instance, social security accounts for 62 percent of total government health expenditures though it only covers 18 percent of the population. There are also geographic inequities in the available services. Inefficiency due to the duplication of facilities by the social security and the ministry of health exist, as well. The public health insurance scheme's incentives encourage economic inefficiency since it is predominantly hospital-based, high technology services that are covered. The types or location of services covered often result in economic inefficiency because they bias the selection of inputs away from the least costly methods toward the most expensive treatment modalities (hospitalization).

Barnum and Kutzin present a wealth of information on various insurance programmes in developing countries, the populations covered, financing mechanisms, and means of reimbursing hospitals, but they found little evidence on the extent to which insurance contributes to hospital revenues. To increase the financing of public hospitals through servicing of more insured patients would require a study of the feasibility of such a programme and the magnitude of the revenues that could be generated. Generation of revenues in this fashion would be most useful if it permitted the charging of services above cost in order to cross subsidize services offered to those unable to pay. This would only be possible if the quality of care at government hospitals increased, and if improvements in the other amenities of hospital care, such as private wards with higher levels of nursing care per patient or more pleasant hotel type services, were offered.

Directions for the Future

A number of key areas require further work for countries to be able to evaluate and select the most appropriate revenue generating options. One is to better understand the impact on utilization and incentives which result from user charges and insurance. Information and experiences from other countries about key resource generation issues could be useful for those countries in the process of addressing and designing their systems. Some of the key questions to be addressed are:

- ◆ the equity implications of the structure chosen
- ◆ the potential and actual revenues generated from alternative financing mechanisms
- ◆ the effects of the payment system on utilization of hospital services,

by patient groups

- ◆ the incentives that the payment or insurance system have on hospital and provider behavior
- ◆ the means for determining the level of fees and exemptions
- ◆ the appropriate fee retention rate at the hospital level to promote efficiency in fee collection.

Countries have experienced problems in making such choices because of the lack of clearly identified priorities and principles to guide them. For example, creation of user fee mechanisms for hospital services should have some basic guiding principles established by countries to help guide their development. Principles for establishing hospital user fees might include:

- ◆ fees should recognize ability to pay and not prevent access to essential services
- ◆ fees should provide the proper incentives for utilization of hospital services
- ◆ fees and the quality of hospital services should be linked
- ◆ services that have important benefits beyond the patient being treated (e.g., communicable disease treatment) should be subsidized.

Similarly, there is a need for countries to define their priorities and develop criteria for health insurance schemes before they are designed. This will preclude the need to adjust the schemes numerous times because of unintended or undesired results. Some possible objectives for insurance schemes may include:

- ◆ production of improved health outcomes
- ◆ encouragement of efficiency in the use of resources:
 - ◆ economic efficiency: selection of inputs not biased away from the least cost method
 - ◆ technical efficiency: reimbursement mechanism does not discourage provision of certain effective services such as preventive care, immunizations, or prenatal care
- ◆ equity in financing
- ◆ prevention of denial of care or financial hardship
- ◆ reduction in access inequalities
- ◆ promotion of quality care
- ◆ acceptability to patients and the public
- ◆ administrative simplicity.

THE NEED FOR FURTHER HOSPITAL ECONOMICS WORK

Changes are occurring in the hospitals of developing countries. The above review of the economic issues and experiences of developing countries concerning resource allocation, resource management and resource generation for hospitals has identified a number of substantive issues to be addressed. What needs to be done relates to both positive economics, which is descriptive and deals with questions of what is, and normative economics, which considers the question of what should be. A number of actions may be taken by governments and international agencies to ensure that the process of change in the management and financing of hospitals achieves its intended objectives.

Country Level Requirements

While the issues and questions concerning hospitals may have similarities across countries, the answers will be different for each country. The hospital sector's problems will be solved and its resources used efficiently and effectively only with country specific strategies. Enlightened decisions about national policies related to these matters as well as local institutional management decisions will depend upon the ability to reflect on experiences of others and to clearly understand the implications of the decisions made regarding resource matters. The ability of developing countries to address these issues will be contingent on several factors:

- ◆ defining the role of hospitals
- ◆ improving information about hospitals
- ◆ developing hospital performance indicators
- ◆ enhancing hospital management capacity.

(a) Defining the role of hospitals

Though the magnitude of health resources devoted to hospitals is known in many developing countries, rarely has it led to an examination of whether that level of resources is consistent with the role hospitals should have in the health system. The clear delineation of roles of the different elements of the health system will facilitate changes that can improve the allocation of health sector resources. Hence, the issue of the role of the hospital must be addressed by countries if they are to plan rationally for the proper number, type and size of hospitals and allocate the appropriate human, physical and financial resources for their operation.

(b) Improving information about hospitals

In most developing countries there are large gaps in the available information on hospitals, their resources and their output. The pressing need is for reliable country based information as well as cross-national information gathering and sharing. Improved information about hospitals is needed for policy, planning and management decisions. The general categories of information needed, adapted from Mills (1990b), are:

- ◆ hospital resource allocation patterns
- ◆ hospital income and expenditure patterns
- ◆ hospital recurrent and capital costs
- ◆ hospital behavior relative to allocation and payment systems
- ◆ hospital performance relative to ownership (public, private not-for-profit, private for-profit).

Information is needed concerning the aggregate national income and expenditure patterns of hospitals, and management information systems at the individual hospital level are needed as well. Hospital managers need regular information about their hospital's income, costs and outputs. To enable hospital managers to use this information effectively, tools will have to be developed and training programmes organized to assist them in learning how to assess what information they need, how to gather it efficiently and how to use it to improve their hospital's operation.

(c) Developing hospital performance indicators

There is a need to identify quantity, cost and quality indexes that can be used to measure total hospital performance and facilitate both inter- and intra-country comparisons. Such an effort must define and develop uniform measures of hospital performance which are well understood and common. This would provide decision makers and managers with comparative data at district, regional and national levels, permitting hospital management, financing and planning decisions to be based on achievement of performance standards. Such a system of measurements should not deal merely with revenues, costs and quantities of service but should also include measurements of service to communities (equity) and of the quality of service provided. Such an effort would promote the improved use of resources. This could serve as a basis for training hospital managers as well as a national planning and monitoring tool for decision makers. The information needed to develop performance indicators might include:

- ◆ community population and services measures
- ◆ quantity or output measures
- ◆ inpatient and outpatient case mix measures
- ◆ cost and revenue measures
- ◆ quality measures.

(d) Enhancing hospital management capacity

To undertake many of the actions outlined above concerning hospital issues will require that countries have a core of skilled hospital managers available. They will need to have the broadest of managerial skills necessary to plan, implement, monitor and evaluate their hospital's resource usage. Planners at the national level must also be skilled in hospital management so as to be able to analyze and address the financing and operational issues facing hospitals. This would include at the national level people with capabilities in health economics, information system development and analytic skills. At the district and regional levels managers need to have basic management skills in planning, monitoring and evaluating. There is also the need to incorporate basic financial management skills in training of mid-level managers and regional and district health office managers so they can make use of information on health resources.

International Agency Actions

In order to bring about rational and efficient hospital resource allocation and improved management, many of the recommended country level actions require countries to share experiences and compare the results of using different models to facilitate change. International agencies have a role in this process by collaborating with countries and providing support through technical and financial assistance as well as providing fora for sharing and disseminating information and experiences among countries. They can actively work with countries in several areas.

- ◆ Training in basic hospital management skills
- ◆ Developing in the hospital setting a critical mass of people with analytic and problem solving skills
- ◆ Exchanging of experiences and lessons learned among countries
- ◆ Encouraging experimentation in hospital financing, efficiency and management improvements
- ◆ Facilitating development of hospital performance indicators
- ◆ Fostering development of improved information gathering techniques and technologies
- ◆ Promoting applied research which finds practical solutions to hospital management and financing problems common to many countries.

REFERENCES

- Akin, J., N. Birdsall, and D. de Ferranti. 1987. *Financing Health Services in Developing Countries: An Agenda for Reform*. World Bank Policy Study. Washington, D.C.: The World Bank.
- Barnum, H. and J. Kutzin. 1992 (forthcoming). *Public Hospitals in Developing Countries: Resource Use, Costs, Financing*. The World Bank. Baltimore: Johns Hopkins University Press.
- Berki, S. 1972. *Hospital Economics*. Lexington, Mass.: Lexington Books.
- Bitran-Dicowsky, R. and D. Dunlop. 1989. "The Determinants of Hospital Costs: An Analysis of Ethiopia." Policy, Planning and Research Working Paper No. 249. Washington, D.C.: The World Bank.
- Bloom, G., M. Segall, and C. Thube. 1986. "Expenditure and Financing of the Health Sector in Kenya". Report. Washington, D.C.: The World Bank, Department of Population, Health and Nutrition.
- Carr, W. and P. Feldstein. 1967. "The Relationship of Cost to Hospital Size", *Inquiry* 4(2): 45-65.
- Cowing, T., A. Holtmann, and S. Powers. 1983. "Hospital Cost Analysis: A Survey and Evaluation of Recent Studies." *Advances in Health Economics and Health Services Research* 4:257-303.
- Creese, A. 1990. "User Charges". *Current Concerns Paper Number 1*. Geneva: World Health Organization, Division of Strengthening of Health Services (WHO/SHS/CC/90.1).
- Evans, R. 1990. "Tension, Compression, and Shear: Directions, Stresses, and Outcomes of Health Care Cost Control," *Journal of Health Politics, Policy and Law*, 15(1): 101-128.
- Feldstein, M. 1967. *Economic Analysis for Health Service Efficiency: Econometric Studies of the British National Health Service*. Amsterdam: North-Holland Publishing Company.
- Feldstein, P. 1983. *Health Care Economics*, 2nd ed. New York: Wiley and Sons.

Griffin, C. 1990. *Health Sector Financing in Asia*. World Bank Report No. IDP-68, Asia Regional Series. Washington, D.C.: The World Bank.

Griffin, C. 1988. *User Charges for Health Care in Principle and Practice*. EDI Seminar Paper Number 37. Washington, D.C.: The World Bank.

John Snow Inc. 1990a. "Final Report-Volume I: Health Sector Financing Issues and Options". Papua New Guinea Health Sector Financing Study Project. Unpublished report. Boston: John Snow Inc.

John Snow Inc. 1990b. "Final Report-Volume II: Hospital Cost Study". Papua New Guinea Health Sector Financing Study Project. Unpublished report. Boston: John Snow Inc.

Kolehmainen-Aitken, R.-L. and P. Shipp. 1990. "Indicators of Staffing Need: A Tool for Assessing Staffing Requirements and Equity in Papua New Guinea, *Health Policy and Planning*, 5(2): 167-176.

Lewis, M. 1990. "Cost Recovery in Developing Country Public Hospitals: The Determinants of User Fee Revenues", Working Paper 2201-01. Washington, D.C.: The Urban Institute.

McGreevey, W. 1990. "Social Security in Latin America: Issues and Options for the World Bank". World Bank Discussion Paper No. 110. Washington, D.C.: World Bank.

Mills, A. 1990a. "The Economics of Hospitals in Developing Countries. Part 1: Expenditure Patterns", *Health Policy and Planning*, 5(2): 107-117.

Mills, A. 1990b. "The Economics of Hospitals in Developing Countries. Part 2: Costs and Sources of Income", *Health Policy and Planning*, 5(2): 203-218.

Newbrander, W. 1987. "Papua New Guinea's Expenditure on Hospitals: Policy and Practice since Independence", *Health Policy and Planning*, 2(3): 227-235.

Overholt, C. 1987. "User Fees in Honduran Hospitals and Health Centers, Policy and Experience", REACH Project Report. Cambridge, Massachusetts: Harvard Institute for International Development.

Prescott, N., ed. 1991. *Indonesia: Health Planning and Budgeting*. A World Bank Country Study. Washington, D.C.: World Bank.

Shepard, D., J. Walsh, W. Munar, L. Rose, R. Guerrero, L. Cruz, G. Reyes, G. Price and C. Solarte. 1991. "Cost-Effectiveness of Ambulatory Surgery in Cali, Colombia." Working Paper No. 1. Department of Population and International Health, Harvard School of Public Health. Boston, Massachusetts. October.

Thomason, J. and W. Newbrander. 1991a. "Health Financing and Budgeting" in *Decentralization in a Developing Country: The Experience of Papua New Guinea and its Health Services* by Thomason, J., Newbrander, W. and Kolehmainen-Aitken, R.-L. Canberra: Australian National University.

Thomason, J. and W. Newbrander. 1991b. "A Survey of Papua New Guinea's Health Sector Financing and Expenditure", *Papua New Guinea Medical Journal*, 34(2): 129-143.

Van Lergerghe, W. and Y. Lafort. 1991. "The Role of the Hospital in the District: Delivering or Supporting Primary health care?" *Current Concerns SHS Paper Number 2*. Geneva: World Health Organization, Division of Strengthening of Health Services (WHO/SHS/CC/90.2).

Vogel, R. 1988. *Cost Recovery in the Health Sector--Selected Country Studies in West Africa*. World Bank Technical Paper Number 82. Washington, D.C.: The World Bank.

World Health Organization. 1987. "Hospitals and Health for All." Report of a WHO Expert Committee on the Role of Hospitals at the First Referral Level. Technical Report Series No. 744. Geneva: World Health Organization.

World Health Organization. 1989. "Summary Record". Document EB83/SR/5, Proceedings from the Eighty-third Session of the Executive Board.

World Health Organization. 1990. "Support to Countries in Rationalizing the Financing of Health Care". Annex 2, Document EB85/35 for the Eighty-fifth Session of the Executive Board.

World Health Organization. 1991. "The Hospital in Rural and Urban Districts." A Report of the Study Group on the Functions of Hospitals at the First Referral Level. (Draft) From a meeting held 30 October to 5 November 1990. Geneva: World Health Organization.