Referral hospitals often prove to be a highly effective focal point for disease-specific health promotion and education activities. Bermuda’s diabetes education program serves all levels of care and provides a strong link between the primary, secondary, and tertiary health care levels. The program is centered in the main referral hospital and serves not only diagnosed patients but also families at risk. Overall, hospitals in the Caribbean are recognizing that central coordination of public health programs within hospitals can provide benefits by strengthening coordination with other services.

Valuing the Benefit of Clinical Services

Measuring the improvement in an individual’s health status produced by the combined activities of a referral hospital, whether for patient care in the hospital or for population-based programs, would theoretically be possible, although practically and methodologically demanding. To our knowledge, such an effort has not been attempted at the referral hospital level, though two studies have attempted to proxy the effect of hospital interventions on health outcomes for small district hospitals, focusing on survival only (McCord and Chowdhury 2003; Snow and others 1994). Both studies indicate that district hospitals appear to have a significant positive effect on health outcomes.

Large numbers of patients receive care in referral hospitals, and most survive with their suffering alleviated, having gained substantial benefit from the care they receive. Therefore, the aggregate direct personal health benefits from referral hospital care will almost certainly be high. The question of whether referral hospital care is cost-effective relative to other interventions delivered at lower levels of care is less easy to answer in aggregate. By its nature, appropriate care in a referral hospital will tend to require more complex input mixes and higher skill levels and, hence, will be relatively expensive. Analysis of the costs and cost-effectiveness of individual interventions offered at different levels is tackled directly by the disease-specific chapters in this volume.

Wider Activities and Functions

Aside from direct patient care, referral hospitals serve other functions within the health system, some of which are offered within the facility, such as teaching and research, while others reach out to the lower levels of the health services, such as technical support and quality assurance.

Advice and Support to Lower Levels. The referral process does not simply entail transferring a patient from a lower to a higher level of care, nor does it end when a patient is discharged from a referral hospital. An effective referral system requires good communication and coordination between levels of care and support from higher to lower levels to help manage patients at the lowest level of care possible. Too often, personnel in referral hospitals adopt an insular and inward-looking perspective, focusing exclusively on the patients directly under their care. However, referral hospitals should offer significant support to personnel in lower-level facilities, and specialist staff members should ideally spend a significant portion of their time providing advice and support beyond the walls of their own hospital, either in person or through various modes of telecommunication. Even in poor countries, a steady improvement in communications infrastructure means that such support functions should become easier to provide over time. Key dimensions of this support function include the following:

- availability by telephone or e-mail to advise referring practitioners on whether referral is required
- specialist advice to the patient’s local practitioner on post-discharge care
- specialist advice on the long-term management of chronic conditions
- specialist attendance at lower-level facilities to provide regular outreach clinics
- provision of expert diagnosis or consultation through telemedicine
- coordination of discharge planning between levels of care
- coordination of the development of and training in the use of shared care protocols and referral protocols
- provision of technology support by skilled technicians and scientists.

Quality Assurance and Quality Improvement. Referral hospitals can and do play a pivotal role in quality assurance and improvement. The most important mechanism for quality assurance and improvement is through the training that referral hospitals provide. The other key mechanism is through the setting of standards for treatment. For example, experts at referral hospitals should review evidence of effectiveness and cost-effectiveness applicable to the local context, determine the formularies to be used at each level of the health system, and develop and amend treatment protocols. Referral hospitals can improve the quality of peripheral services by giving advice, offering on-site training, providing clinical services alongside local practitioners, and monitoring the quality of the referrals they receive.

Education and Training. Many tertiary referral hospitals in developing countries are associated with universities and medical schools and may, therefore, also be regarded as teaching hospitals. Any country wishing to train its own doctors will need one or more teaching hospitals. The number of doctors a country needs will be influenced by its level of development, resources, and personnel structure. Many will aim for a ratio of at least...
2 per 1,000 population, though most developing countries have 0.05 to 1.0 per 1,000 (Puzin 1996; WHOSIS 2004). If we assume a 40-year working life and loss through brain drain or other attrition of 25 percent, the number of doctors that must be produced each year is between 16 and 67 per 1 million population, resulting in 0.5 to 2.0 doctors per 1,000 population (box 66.1). A population of 40 million would, thus, need medical schools able to graduate between 640 and 2,680 doctors per year. Medical schools possess economies of scale, and although some extremely small schools train 50 or so students a year, agreement is widespread that a class size of about 150 to 200 is optimal (see, for example, Harden and Davis 1998). A country with fewer than 3 million population would really need to consider whether training doctors locally is justified on economic and other grounds, but for larger countries, the arguments for training doctors locally are strong, and a teaching hospital would, therefore, be required.

Basic generalist doctors should be trained in a range of facilities across all levels of care, reflecting the facilities in which they will work after graduation. Traditional approaches toward medical education have been widely criticized by educationalists and health planners for being dominated by training in tertiary settings by specialists. Not only is this setting inappropriate, but typical content and clinical experience do not reflect what the doctors will be doing or what they will need to know after qualification. Nevertheless, the university teaching hospital cannot be omitted from the basic training of doctors. If students and faculty were involved only in district-based services, they would miss many important advances in biomedical science and the care of complex problems (Husain 1996). Moreover, doctors need to know enough about what the various tertiary specialties do to be able to refer patients appropriately and to make personal career choices.

The training of specialists, of course, depends far more on the existence and proper functioning of referral hospitals. Again, a particular country will need to decide how many specialists it needs in which specialties and whether it should send its doctors abroad to specialize or train them internally. In developed countries, 60 to 90 percent of doctors are specialists, whereas in developing countries the range is wider (for example, 76 percent of Indian doctors are specialists, 45 percent are specialists in Tanzania, and 31 percent are specialists in Morocco). A World Health Organization expert workshop agreed on a figure of 50 percent (Puzin 1996). Therefore, a country of 40 million would aim to train approximately 300 to 1,300 specialists per year. On average, such training lasts four years. Thus, at any time the academic referral hospital system would need to supply 1,200 to 5,200 residents. A guideline many countries use requires a ratio of postgraduate specialist supervision of not more than two residents per qualified specialist. This ratio can be used to get some idea of the referral hospital capacity required to train specialists.

Although basic doctors could spend most of their training time in primary care and district hospital facilities, with limited exposure to tertiary care hospitals, the training of specialists—as well as of other specialized allied staff members such as nurses for intensive care or specialized psychiatry, physiotherapists specializing in back injuries or burns, and pharmacists specializing in oncology—can take place only in referral hospitals.

In recent years, continuing medical education has grown in importance as the need for professionals to continually update their knowledge and acquire new skills has been more clearly appreciated. The coordination and provision of appropriate continuing medical education depends heavily on the specialists and academics associated with referral and academic hospitals.

**Management and Administration.** Referral hospitals in many developing countries play important roles in providing managerial and administrative support to other elements of the health system. These roles may include managing laboratory services on behalf of the whole health system; serving as the location for drug and medical supply depots and distribution systems and managing procurement systems; hosting and managing health information systems, often including epidemiological surveillance systems; managing centralized transport fleets; and, on occasion, providing financial management, payroll, and human resources.
resource management services to other health units. Our intent is not to consider whether such arrangements are “right” or “wrong”—complex factors would have to be taken into account in every individual circumstance—but to note that making changes to the functioning of referral hospitals may have unintended consequences. For example, moving referral hospitals from funding based on a global budget to reimbursement systems based on patient activity may unintentionally cause hospitals to cease to provide these wider support functions if explicit alternative funding mechanisms are not established.

Research and Innovation. Referral hospitals tend to be where most health research is undertaken. Whereas in developed countries they may often be associated with the development of new technologies, in developing countries they are more often the site of research for the initial piloting and introduction of new technologies developed elsewhere and for the evaluation of their local suitability and field efficacy. Referral hospitals are also the vehicle for disseminating such technologies through the exposure of staff during training as well as through the role that referral hospitals frequently play in continuing professional education.

Research activities are vital in attracting and retaining specialist staff members who are required not just for the treatment of complex patients, but also for the training of new specialists. Research that is responsive to local conditions—that is, local disease burdens and technology constraints—fills a critical gap because researchers in developed countries and pharmaceutical companies do not generally pursue such research questions if they do not foresee sufficient returns to their investments.

Valuing the Indirect Contribution to the Health System. From the enumeration of the many roles of referral hospitals and their indirect effect on health through their contribution to the health system by way of supervision, administration, training, research, and quality improvement, it is immediately evident that these benefits cannot readily be translated into DALYs or any other metric to be used in a relative cost-benefit analysis.

Externalities and Intangible Benefits

The previous sections reviewed the various functions of referral hospitals within the health system, all of which contributed directly or indirectly to the health status of individuals. This section addresses other ways in which referral hospitals contribute to welfare and well-being, and comments on how they complicate the issue of valuing the contribution of referral hospitals in society.

Referral hospitals have a broader effect on overall societal welfare than can be captured by measures of health outcomes. Utility, or welfare, includes health as one of many important outcomes, such as financial security, risk alleviation, and psychological reassurance. However, as Hammer and Berman (1995) note, health policy is typically conducted as if it has a unidimensional objective—namely the maximization of health (DALY) outcomes. Determining the appropriate resource allocation to referral hospitals purely on the basis of the cost of generating health (DALYs) may, therefore, seriously underestimate the optimum level of resources, because such measures will fail to capture the full welfare gains from the availability of higher-level health services. An example will highlight the difference between valuing hospitals on the basis of their contribution to health status alone compared with including wider concepts of welfare in the valuation.

Renal failure leading to the need for dialysis is relatively rare, and certainly rare in comparison to many other infectious and chronic diseases in lower- or middle-income countries. Treatment is lifesaving, but must continue indefinitely (involving visits two or three times every week) and is, therefore, extremely expensive. In many cases, dialysis can be justified only if it is linked to a renal transplant program, which terminates the need for dialysis and frees the equipment for someone else. The proportion of the total population who will benefit from such a referral hospital program is small; therefore, the DALYs generated are low, and the program would not rank high among the priorities given a limited budget. However, every member of the population is at risk of renal failure and, if affected, would find that, in the absence of a publicly funded program, he or she would either die or face extremely high costs to secure treatment in the private sector or abroad.

Even in poor countries, patients’ price elasticity of demand is low when faced with life-threatening illnesses, particularly when treatment can change the outcome. Studies on poverty have shown that a significant proportion of households that have become poor did so as a result of serious illness, which resulted in their liquidating assets to pay for health care (see, for example, Liu, Rao, and Hsiao 2003). Thus, people seek the peace of mind of knowing that they can obtain lifesaving treatment should they need it without the risk of incurring catastrophic costs of care. This additional welfare derives both from the financial security of not having to spend more than people can afford to save their lives and from the direct health benefits of treatment itself. The utility from the former (financial security) increases with the cost of the intervention required, whereas the utility derived from the latter (direct health benefits) is unrelated to the cost of the intervention. Paradoxically, one could, therefore, argue that the rarer a particular illness is—and the more costly the intervention required—the greater will be the welfare gain from public spending on that intervention.

This argument, of course, is likely to stand in direct contrast to the conclusions drawn from prioritization based on cost-effectiveness. For most individuals, willingness to pay is far less
than the costs of the procedure to them; however, because the whole population benefits from the security of knowing that each individual would be entitled to referral hospital care should he or she need it, in the aggregate the welfare value generated by public provision or funding may be many times greater than the value of the DALYs generated directly for those few patients who do receive treatment. This literature review did not find evidence of studies on national willingness to pay for referral hospital care in developing countries, but this area could be of interest for future research.

In practice, too, the public—particularly an urban, middle-income public—expects the government to provide care of last resort for complex trauma or diseases, especially for natural and man-made disasters. Thus, even though referral hospitals may provide care to a small number of people, often with limited health benefits, politicians and the public alike may value and prioritize them simply because they meet the public’s expectations for what the government must provide. In addition, politicians and the public often regard a country’s ability to provide the kind of complex, high-tech care offered in a referral hospital as a measure of that country’s level of development and sophistication, and it is a source of national pride. Whether economically rational or not, this nonhealth benefit appears to drive public choices to some extent.

Negative Impacts

The “negative” impact of referral hospitals is largely attributable to their potential to exert distortionary effects on the health system by diverting resources from peripheral areas and from lower levels of care (Fiedler, Schmidt, and Wight 1998; Filmer, Hammer, and Pritchett 1997) for the following reasons:

- Tertiary hospitals and specialists have a high political and public profile.
- Urban and political elites are more likely to use referral hospitals than rural primary care facilities or district hospitals.
- Harmful competition with lower levels of care may result from the maintenance of higher-level referral hospitals in many poor countries, lowering use of the former.
- Referral hospitals can be entry points for the introduction into the health system of inappropriate and unaffordable technologies.
- Skilled personnel frequently find referral hospitals far more attractive to work at than rural and district hospitals for such reasons as preferences for a metropolitan location, better hospital resources allowing for a more rewarding professional experience, and better opportunities for private practice (official or unofficial). However, given the huge problem of global migration of health workers from poor to rich countries (Bundred and Levitt 2000), one could argue that well-functioning referral hospitals might provide local health professionals with a good incentive to remain at home, whereas the absence of referral hospitals would increase the propensity of local professionals to emigrate.

Determinants of an Appropriate Balance of Referral-Level Care

When one considers the ideal level of resources to be provided for referral hospital care and the appropriate balance between resources for referral hospitals and for other levels of the health care system, no simple formula is available that can be applied to different countries and contexts. However, certain key factors have an important influence on the need and demand for referral-level care, the resources that may be available to the health sector, and the ability of the health sector to provide adequate and effective care in different settings.

General Determinants

Arguably the most important determinant of demand for and ability to pay for referral hospital care is a society’s level of economic development and wealth, captured (albeit imprecisely) by measures of GDP per capita. Extensive international evidence indicates that national health expenditure displays an unambiguously positive income elasticity both across countries and over time; that is, as a country gets richer, it spends relatively more on health (see, for example, Getzen 2000; Schieber 1990). Studies in developed countries indicate that in the United States, every 1 percent long-run increase in GDP leads to a 1.6 percent increase in health expenditure, and in other countries the increase is between 1.2 and 1.4 percent (Getzen 2000). Therefore, expecting developing countries to spend a higher proportion of their GDP on health care as they become wealthier seems to be reasonable. If the poorest countries were to focus their limited resources on highly cost-effective interventions in primary health care, somewhat better-off countries might be expected to spend progressively more on the referral hospital level as resources became available.

An overlapping set of demographic and geographical factors also plays an important role in determining the balance of referral care—namely, population size, population density, terrain, distances between main urban centers, and access. Populations of some millions are required to justify a major tertiary hospital with a full range of tertiary services. Small countries with populations of less than 1 million will certainly not be able to provide a full range of tertiary hospital services because of the need to achieve minimum volumes to ensure service viability and to attract a critical mass of specialized personnel. Countries with fewer than 100,000 inhabitants (generally island states) may find even secondary hospital services
beyond their means and capabilities. Supranational referral, reliance on larger neighbors, or regional collaboration may be unavoidable for smaller countries, especially for tertiary care provision, with the Caribbean and southern Africa providing clear examples of many smaller states relying on referral facilities in larger or wealthier neighbors. Within larger countries, population density can complicate the planning of referral services. Compact countries or regions with dense populations can typically provide high levels of access to referral care at a relatively small number of sites, whereas countries or regions with more dispersed populations face more complex tradeoffs regarding number of sites, volume thresholds, and transportation systems.

The other main influence on the appropriate balance of referral services for a given country is its particular pattern and burden of disease. Although referral-level services will always be needed, as a society passes through epidemiologic and demographic transitions, it is likely to require more of those services typically found at referral hospitals. For example, rapidly increasing rates of heart disease and cancers are typically encountered in industrializing nations and aging populations, and these are diseases whose effective management requires access to the interventions, skills, and equipment that will typically be concentrated at the referral hospital level.

**Health System Determinants**

A number of factors specific to the particular context of a country’s health system will also influence the appropriate balance between referral hospitals and lower levels of care. These factors are especially important in considering the appropriateness of plans to change the balance of care between levels. Broadly, they can be summarized as follows:

- capabilities of lower levels
- availability of specialized personnel
- training capacity, organization, and needs
- cultural issues, political issues, and traditions.

The first three factors are closely interrelated. If primary health care and district hospital services are weak, cutting resources for referral hospitals without destabilizing the system will be more difficult. In such circumstances, rapid rebalancing of resources is unlikely to be possible because careful efforts will be required to develop lower-level services first, while still maintaining the referral service. Where lower-level services are strong, devoting relatively fewer resources to referral hospitals may well be possible. However, even though an effective district health system will be able to treat a large proportion of patients at lower levels of care, it will also be better able to identify patients who require referral for more complex care and, thus, may generate a greater appropriate demand for referral hospital care.

Referral hospital services require a specialized staff to fulfill their mission. If specialized personnel are not available in a country, then attempting to develop referral hospitals on a large scale will clearly be infeasible. However, many countries arguably have too many specialized staff persons and too few well-trained generalists. Where large numbers of specialists exist, their presence will likely tend to draw resources disproportionately toward the referral level and away from district health systems. Wherever such imbalances exist, positive changes will require a substantial training or retraining agenda. The feasibility of such efforts is closely linked to the professional and social status of different professional groups and subgroups—for example, whether medical specialists are viewed as having a higher status than general practitioners—and to the premium a society places on having access to “advanced” medical care.

**CURRENT BALANCE OF CARE IN PRACTICE**

In this section, we summarize data on the current balance between referral and lower levels of care. We first look at the share of total health expenditure going to these different levels, but given that referral care normally has much higher unit costs, we recognize that the balance also needs to be viewed in terms of volume of cases and access and equity.

**Share of Health Expenditure**

Different health systems categorize hospitals and services rendered differently. Methodologies in national health accounts in developing countries during the 1990s and early 2000s have tended to use a simple, catch-all category of “hospitals” or “acute hospitals” (for example, WHO 2002). Even high-income countries following the Organisation for Economic Co-operation and Development’s system of health accounts provider classification (OECD 2000, 136) distinguish only between “general” hospitals and “mental health and substance abuse” and other “specialty” hospitals in their national health accounts. Consequently, making valid cross-country comparisons of spending by levels of hospital care remains extremely difficult.

Mills (1990a) reviews published data on hospital expenditure patterns in developing countries, and Barnum and Kutzin (1993) provide a comprehensive survey of expenditure on hospital services in a number of developing countries, drawing their information largely from World Bank sector reviews. These analyses remain the most authoritative assessment of the proportion of public health expenditure absorbed by secondary and tertiary hospitals, even though their data represent only a handful of countries at different points in time.

Overall, Mills (1990a) finds that hospitals in developing countries appear to absorb from 30 to 50 percent of total
health expenditure. Public hospitals of all types absorb some 50 to 60 percent of public health expenditure, and secondary and tertiary hospitals absorb about 60 to 80 percent of public hospital expenditure, with the remainder going to district hospitals. Her results are broadly similar to those of Barnum and Kutzin (1993, 26–33), who find that public hospitals at all levels absorb a mean of approximately 60 percent of recurrent public health expenditures. Across five countries (Belize, Indonesia, Kenya, Zambia, and Zimbabwe), they find that tertiary hospitals account for between 45 and 69 percent of total public expenditure on hospitals. In South Africa, Thomas and Muirhead (2000) find that tertiary hospitals account for 28 percent of hospital expenditure and 17 percent of total public health expenditure, but when taken together with regional referral hospitals, constitute 59 percent of hospital expenditure.

Unit Costs of Care

One of the explanations for the high share of expenditure that flows through higher-level referral hospitals is, of course, that the unit costs of a referral hospital are necessarily higher than the unit costs of a district hospital. This difference results from the more complex case mix they treat, their more expensive inputs, and the additional costs of their teaching functions (Barnum and Kutzin 1993, 26). Mills (1990b) reports that her analysis of available data indicated that secondary-level hospitals were typically twice as expensive per bed day as district hospitals and that tertiary hospitals were typically between twice and five times as expensive per bed day as district hospitals. Barnum and Kutzin (1993) find similar relationships between unit costs by level of hospital in a variety of different countries. This upward gradient in unit costs has also been found in econometric studies of hospital costs (Adam, Evans, and Murray 2003) and has been explicitly incorporated into the regression-based unit cost estimates used in other chapters in this volume.

Table 66.3 shows data on unit costs by level of care from Mauritius and highlights a commonly encountered contradiction of the preceding paragraph—namely, that costs appear similar at all levels. This phenomenon is explained by average bed occupancy in Mauritius that was around 45 percent in 1995, compared with average bed occupancies of 90 percent or more in tertiary hospitals. Thus, the high cost of district hospital care in this case reflects not inputs, which are much less extensive than in a tertiary hospital, but the effect of low levels of utilization. Such a difference in utilization between levels of hospital tends to be the norm in many developing countries (Barnum and Kutzin 1993, 91–96). Note that the regression-based unit costs of district hospitals used in the cost analysis for this volume reflect an “optimized” bed occupancy of 80 percent (Mulligan and others 2003, 29). This assumption is entirely defensible from a long-run perspective, assuming cost-minimizing behavior is necessary and appropriate. It does, however, reflect quite a substantial shift from the levels of utilization and productivity commonly seen in rural district hospitals in most developing countries.

The use of a simple unit cost hides important cost differences between services and specialties within the same hospital, as demonstrated by the breakdown for Mauritius tertiary hospitals in table 66.4. Differences in length of stay for different specialties and conditions also obviously contribute to radically different costs per admission or patient; however, these differences should be captured by the condition and procedure costs used in the disease chapters in this volume.

Table 66.4 Cost Per Bed Day for Selected Specialties, Tertiary Hospitals, Mauritius, 1995 (2001 U.S. dollars)

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>Orthopedics</td>
<td>18</td>
<td>23</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>29</td>
<td>43</td>
</tr>
<tr>
<td>Cardiothoracic surgery</td>
<td>36</td>
<td>39</td>
</tr>
<tr>
<td>Burns</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Intensive care unit</td>
<td>106</td>
<td>120</td>
</tr>
</tbody>
</table>

Source: Murray and others 1996.

Table 66.3 Cost Per Bed Day in a Medical Ward by Level of Hospital, Mauritius, 1995 (2001 U.S. dollars)

<table>
<thead>
<tr>
<th>Level of hospital</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>District</td>
<td>17</td>
</tr>
<tr>
<td>Regional</td>
<td>21</td>
</tr>
<tr>
<td>Tertiary</td>
<td>20</td>
</tr>
</tbody>
</table>

Source: Murray and others 1996.

Appropriateness of Utilization of Referral Hospitals

Perhaps the most frequent theme in the research literature on referral hospitals in developing countries is the inappropriate utilization of higher-level facilities and the apparent failure of most referral systems in developing countries to function as intended. Broadly speaking, hospitals of all levels, up to and including national tertiary centers—especially in their outpatients departments—are overwhelmed by patients who could have been treated successfully at lower-level facilities, many of
whom have self-referred, bypassing primary health care or district hospitals in the process (Holdsworth, Garner, and Harpham 1993; London and Bachmann 1997; Omaha and others 1998; Sanders and others 2001).

Atkinson and others (1999) describe an extreme manifestation of this phenomenon, whereby the University Teaching Hospital is actually the only public hospital in Lusaka. Combined with the bypassing of primary health clinics in the city, this situation results in the University Teaching Hospital’s functioning primarily as a glorified health center and first-contact provider for most of Lusaka’s population. The problem of bypassing typically seems to be driven by a number of factors, including patients’ perception of superior quality of care and resource availability at referral hospitals, which often may be entirely well founded and rational (see, for example, London and Bachmann 1997; Nolan and others 2001); the desire to avoid delays in care if referral to a higher-level facility proves to be necessary; and the fact that for many urban populations a referral hospital may simply be the closest health facility. Grodos and Tonglet (2002) argue that many countries’ failure to develop an adequate urban equivalent of the district health concept greatly exacerbates inappropriate utilization of hospitals. The urban phenomenon of widespread bypassing and self-referral is frequently accompanied by low rates of formal referral from rural and outlying facilities (see, for example, Nordberg, Holmberg, and Kiugu 1996; Omaha and others 1998).

These problems have a number of negative impacts and consequences. Simple conditions are unnecessarily treated in a high-cost environment; outpatient departments are congested by patients requiring primary care, thus causing long waiting times; scarce staff time is diverted from specialized areas and into inappropriate care; and more complex cases requiring specialized care are crowded out by more urgent but less technically demanding cases that could be cared for at lower levels. The latter has been a particular concern in those countries with more serious HIV/AIDS epidemics. As the number of patients falling sick with AIDS increases rapidly, they start to occupy a significant proportion of beds in hospitals at all levels (Gilks and others 1998), inevitably crowding out patients requiring other forms of care. Although AIDS cases may well require hospitalization, only a small proportion of cases require specialized or tertiary care. Gilks and others (1998) find that this crowding-out effect may fall over time as the health system adjusts to the pressures of AIDS, but countries facing impending AIDS epidemics should be prepared for its initial appearance.

Taken together, this complex of problems undermines the effective delivery of both specialized care and appropriate primary health care. Specialized care is pushed to the background by the human wave of demand for primary care, while hospitals unwittingly further undermine the credibility of the primary health care system through one-sided competition (Stefanini 1994), which reinforces the cycle and ensures that primary health care facilities remain underused and inefficient.

**Access and Equity**

By their nature, referral hospitals must be limited in number and will inevitably be sited in major towns and cities. As a result, a significant portion of the population, especially people living in rural areas, will tend to live at some distance from the nearest referral hospital. Studies of the accessibility of referral hospital care in countries such as Ethiopia (Kloos 1990) and Nigeria (Lyun 1983; Okafor 1983) have repeatedly confirmed the existence of a steep distance-decay function, indicating that—other things being equal—individuals with a given need for a clinical service will be less likely to access that service the farther away from the referral center they live.

Compounding the impact of distance, investigators find that problems relating to the availability, regularity, and cost of transportation to referral centers also affect service utilization (Kloos 1990; Martey and others 1998). The same authors indicate that prohibitive hospital fees are often a significant barrier to utilization, especially among poorer patients. Other important barriers included perceived lack of drugs and essential supplies, even at referral centers; negative staff attitudes; and cultural and linguistic differences (for example, where the staff at a referral center does not speak the language of a patient from a remote area). As noted earlier, peripheral district hospitals also tend to have low rates of referral. These barriers, which all disproportionately affect rural patients, must be contrasted with the phenomenon noted earlier of excessive and inappropriate use of referral hospitals for primary care by urban residents.

In addition to finding that public hospitals favor urban residents over rural dwellers, a number of studies have indicated that public hospitals in many poor countries disproportionately benefit the better off, leading their authors to argue that diverting public funds from hospitals and toward primary health care would be pro-poor (see, for example, Castro-Leal and others 2000; Filmer, Hammer, and Pritchett 1997; Mahal and others 2002). Other studies find this tendency in some countries but not in others (Makinen and others 2000). By contrast, in Latin American countries, Barnum and Kutzin (1993) find strong evidence that public hospitals are pro-poor in their distributional effect. Even if referral hospital services are not currently pro-poor, policy makers face two contradictory alternatives: either to retarget public funds toward primary health care for the poor, hence greatly reducing or abandoning public funding for referral hospitals, or to attempt to remove the barriers that prevent the poor from using higher-level services, which would probably require increased spending on all levels of care.
GETTING BETTER VALUE FOR MONEY FROM THE HOSPITAL SYSTEM

Although prescribing how resources should be allocated across levels of care is hard, at least they should be efficiently used, wherever they are spent within the hospital system. The preceding analysis has highlighted how deficiencies at the lower levels of the hospital system render referral hospitals less efficient and how factors that affect access lead to skewed benefits and inequity. Here we look more specifically at three areas for improving the efficiency of the hospital system: interventions within the referral hospital, the use of public-private partnerships, and strengthening of the referral chain.

Improve the Efficiency of Referral Hospitals

Although space does not permit a lengthy discussion of approaches to improve efficiency in the context of referral hospitals, this aspect is nonetheless important in planning and system strengthening (for a more detailed discussion see Barnum and Kutzin 1993; Hensher 2001; Walford and Grant 1998). In summary, the key areas on which planners and managers should focus are as follows:

- reducing inappropriate outpatient and inpatient use of referral
- improving systems to allow early discharge from the hospital
- ensuring that bed occupancy rates can be maintained as close as possible to optimal rates—namely, 85 percent for referral hospitals
- developing systems for booked outpatient appointments, admissions, and procedures to permit better planning of activity and staffing
- undertaking as much activity as possible on an ambulatory rather than an inpatient basis, supported by the use of “step-down” beds and patient hotels
- evaluating the staff skill mix and the potential for skill substitution, as well as efficient remuneration strategies, on a continuous basis
- evaluating and improving processes and systems, including cost-effective clinical guidelines for patient treatment, on a continuous basis
- ensuring that new or replacement referral hospitals conform as much as possible to available evidence on economies of scale—that is, that hospitals with fewer than 200 beds are likely to be scale inefficient and that diseconomies of scale are likely to become increasingly evident in hospitals with more than 600 beds
- adopting intelligent procurement processes and engaging in effective negotiations with suppliers in relation to prices and service levels
- ensuring effective ordering, stock control, and distribution systems to minimize theft and wastage of key supplies
- undertaking planned preventive maintenance and programmed replacement of equipment and buildings.

Can Public-Private Interactions Improve Efficiency?

In the context of this discussion, privately owned hospitals that provide subsidized care to public patients, such as nongovernmental organization and mission hospitals, are regarded as public hospitals. Private refers to for-profit hospitals that are generally funded by paying patients and are minimally subsidized. Few studies have been undertaken of how private hospitals operate in developing countries (see, for example, Muraleedharan 1999). Although the exact balance of and relationship between the public and private health sectors varies greatly from country to country at all levels of the health system, a common theme in almost all low- and middle-income countries is that private hospitals do not follow the pyramidal referral form that public hospital systems have adopted almost universally. Most private health sectors do not clearly delineate district, secondary, or tertiary hospitals. Different private hospitals may offer different services and facilities on a more or less idiosyncratic basis, with independent medical specialists practicing and admitting patients at various different hospitals.

In most systems, scope exists for both positive collaboration and competition between public and private hospitals, especially for secondary and tertiary services. Competition between public and private sectors obviously has the potential to be beneficial by driving quality up and costs down, but it may also have negative effects by encouraging a duplication of services and resulting in the underutilization of fixed capital by creating perverse incentives for physicians and patients and by competing with the public sector for scarce human resources. In some settings, the private sector may be able to offer services that the public purse cannot afford to provide, thus allowing patients who could not afford private care some chance of accessing sophisticated treatments through the government’s paying private providers or by some pro bono provision of treatment for poor patients.

In many countries, government hospitals are establishing private wards as a vehicle for income generation. The fees for such units are lower than those at private hospitals, offering access to private facilities to patients who may not be able to afford private hospitals. The link with academic medicine often adds to the appeal of such facilities. However, as is the case in South Africa, effectively only tertiary hospitals and a handful of secondary hospitals are felt to be attractive enough to private patients to offer genuine opportunities as preferred providers. The mass of district and regional hospitals are unlikely to be attractive to private patients; therefore, the positive spinoffs of these initiatives may be limited in their scale and reach.
Contracting out services to private providers, particularly high-cost, low-volume services, may be an efficient way to offer such services to public patients. For example, the government of Barbados contracts out surplus demand for dialysis to a private facility on the island. In some provinces of South Africa, expensive imaging such as MRI has been contracted out to private radiology practices. South Africa is also experimenting with contracting out the management of some academic referral hospitals to a private hospital group that is assumed to have greater management expertise and is free from certain public sector constraints, such as salary scales for senior managers. It is too early to judge the success of this arrangement, but in all cases it is imperative that contracts be carefully regulated, monitored, and enforced. For a comprehensive review of contracting, see Bennet, McPake, and Mills (1997).

Particular problems may arise where the same doctors provide care in both public and private hospitals. Under fee-for-service arrangements, physicians may focus on their more lucrative private patients to the disadvantage of public hospital patients, refer patients with adequate insurance to their private practices and private hospitals, and transfer patients with expensive diseases or inadequate insurance to public hospitals.

**Improving the Functionality of Referral Systems**

An ideal referral system would ensure that patients can receive appropriate, high-quality care for their condition in the lowest-cost and closest facility possible, given the resources available to the health system, with seamless transfer of information and responsibility as that patient is required to move up or down the referral chain. Although few referral systems anywhere in the world live up to this ideal fully, it does provide a target in relation to improving the current situation. Improving the effective functioning of referral systems broadly requires progress in three areas: referral system design, facilitation of the smooth transfer of patients and information between levels, and what Walford and Grant (1998, 38) refer to as effective “referral discipline.”

Improving referral system design must start with a detailed attempt to assess which services should be provided at which level of care, encompassing community- and home-based care, primary health care, district hospitals, secondary hospitals, tertiary hospitals, and specialized hospitals. Such an assessment must take local circumstances into account, requires a significant analytical and consultative effort by planners and clinicians if it is to be credible, and must explicitly be open to revision in light of practical experience. After such an exercise has identified which services can appropriately be provided at each level of care, adequate resources must be dedicated to strengthening lower levels of care to make them attractive and credible in the eyes of patients. This effort will require significant investment and funding to ensure the availability of appropriate staff members and supervision, to ensure continuous drug supplies, and to provide basic laboratory tests (Walford and Grant 1998, 38). Given the pervasiveness of inappropriate use of referral hospitals for primary health care problems by urban residents, both urban and rural primary health care and district health systems must be adequately strengthened. Financing strategies that redistribute funds from urban to rural regions may unwittingly hamper such strengthening of the referral system by failing to allow for the development of appropriate lower-level facilities for urban residents. This risk is especially high when a country is pursuing a redistributive agenda against a background of limited or zero overall growth in expenditure.

From a physical planning perspective, planners should consider providing primary health care and district hospital walk-in ambulatory services (emergency and general outpatients) in a physically distinct facility sited immediately next to the referral hospital. This arrangement not only enables triage and filtering of less severe cases (while proximity ensures that severe emergency cases can be transferred rapidly) but also enables rigorous enforcement of a referral-only policy within the referral hospital.

The development of effective patient transportation arrangements is also critical, not only to ensure that patients from remote areas have a fair chance of being successfully referred to a center of excellence (bearing in mind that most referral systems will almost certainly need to increase referral rates from rural areas), but also to ensure that patients can be discharged in a timely and well-planned fashion.

Perhaps more challenging is the concurrent need to align the incentives of referral hospitals, district hospitals, and primary health care services. This goal may or may not be achievable by means of an integrated management structure, but it certainly requires a good deal of communication, collaborative planning, and collaborative development of shared care protocols, and senior personnel need to be given responsibility for coordination and liaison across key interfaces of the referral network. A single, global budget controlled by an authority that is concerned with optimizing the cost-effectiveness of health care delivery would seem to be a necessary condition to achieve alignment across service levels; however, a consideration of financing mechanisms is beyond the scope of this chapter.

At the patient level, a number of mechanisms to improve referral discipline can be considered. In situations in which eliminating nonreferred patients entirely from the referral hospital is impossible, queuing systems should be redesigned to separate referred patients from nonreferred patients so that referrals can be fast-tracked. Explaining to nonreferred patients why other patients are being fast-tracked past them is important to encourage them to seek referral in future. Ideally, they should be diverted to an on-site primary health care facility where they can be treated more quickly than in the referral
hospital. Another possibility may be to institute bypass fees for nonreferred patients, charging them a penalty fee for failing to use the referral system. Such a decision requires careful consideration and planning. Credible lower-level care must be readily available, and substantial efforts to communicate the new policy to the public will be required if this approach is to be seen as fair. More broadly, intensive public communication and education will be essential to inform the public how, where, and when they should seek health care at different levels and to build their confidence that lower-level facilities really will be able to offer acceptable quality care when they need it.

CONCLUDING COMMENTS

This review of the available evidence indicates that referral hospitals frequently do command a large share of health sector resources and expenditure, yet no simple way exists of assessing what an appropriate share would be. Strong referral hospitals can distort priorities and undermine basic services, but they also provide important health benefits to large numbers of patients whom they treat successfully. Referral hospitals provide essential support to lower levels of the system, which cannot function effectively without access to upward referral, and they are frequently the most functional component of the health system, paying greatest attention to quality of care.

Overall, we have argued that both national and international policy makers should be cautious before demanding the reallocation of resources away from referral hospitals and should be still more cautious in allowing themselves to believe that such a reallocation is likely to be achievable in practice. In particular, this chapter has made the case that a unidimensional focus on cost-effectiveness analysis and cost per DALY gained will fail to capture the importance of referral hospital services adequately. In reality, in most developing countries, the scope for reallocation of resources from referral hospitals to lower levels of care is limited, and the managerial demands of achieving a successful reallocation are great. Lower levels of care certainly require strengthening, but this need is more likely to reflect inadequate financing of the entire public health system than a grossly excessive allocation to referral hospitals. Instead, referral hospitals should perhaps be seen as the capstone of the referral pyramid: they should not be too heavy, but if they are too light, the levels below them will lose cohesion. A restructuring of referral hospital services is certainly called for to improve appropriate referral and utilization, especially by remote and rural populations; to transform the inappropriate use of referral hospitals as primary health care providers; to improve efficiency; and to provide much better outreach and support to lower levels of care.

This restructuring should not be confused with wholesale demolition. Undermining referral services will be far more likely to undermine and destabilize the entire health system than to liberate resources for primary health care. Clearly, countries must critically evaluate their health priorities and their balance of care and resources between levels, but they should do so carefully and thoroughly, with a clear understanding of the analytical effort required to draw meaningful conclusions, of the planning and managerial capacity that they will require to bring about successful change, and of the long time frames required to develop and implement robust plans for major system changes.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the crucial assistance of Etienne Yemek in undertaking literature reviews.

REFERENCES


