Health Facility Surveys: An Introduction

by

Magnus Lindelow\textsuperscript{a,b} and Adam Wagstaff\textsuperscript{a,c}

\textsuperscript{a} The World Bank, The World Bank, 1818 H St. NW, Washington, DC, 20433, USA
\textsuperscript{b} Centre for Study of African Economies, St Anthony's College, Oxford OX2 6jF, UK
\textsuperscript{c} The University of Sussex, University of Sussex, Brighton, BN1 6HG, UK

Summary

Health facility surveys come in various guises. One dimension in which they vary is their motivation. Some seek to understand better linkages between households and providers. Others seek to understand better provider behavior and performance. Others still seek to understand the interrelationships between providers, while yet others seek to shed light on the linkages between government and providers. Health facility surveys differ too in the data they collect, in part due to the different motivations. Surveys also vary in the way they collect data, some relying on direct observation, some on record review, and some on interview. Some quality data are collected through clinical vignettes. Facility data have been put to a variety of uses, including: planning and budgeting; monitoring, evaluation, and promoting accountability; and research. We review some of the literature under each heading, and offer some conclusions regarding the current state of health facility surveys.

Corresponding author and contact details: Magnus Lindelow, The World Bank, 1818 H Street NW, Washington, D.C. 20433, USA. Tel. (202) 458-0125. Fax (202)-522 1153. Email: mlindelow@worldbank.org.

Keywords: Health facility surveys; household surveys; costs; quality of care; accountability.

Acknowledgements: The findings, interpretations and conclusions expressed in this paper are entirely those of the author, and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.
# Contents

I. Introduction .......................................................................................................................... 2
II. What Motivates the Exercise? ............................................................................................ 3  
   PROVIDER-HOUSEHOLD LINKAGES .................................................................................... 3  
   PROVIDER PERFORMANCE .................................................................................................. 4  
   LINKS BETWEEN PROVIDERS ............................................................................................ 4  
   GOVERNMENT-PROVIDER LINKAGES ................................................................................. 4  
III. What Data Are Collected? ............................................................................................... 5  
   INPUTS ................................................................................................................................. 5  
   COSTS ................................................................................................................................... 7  
   OUTPUTS, ACTIVITIES AND THROUGHPUTS ...................................................................... 7  
   QUALITY ............................................................................................................................... 9  
IV. How Are the Data Collected? .......................................................................................... 10  
   INPUTS .................................................................................................................................. 10  
   COSTS .................................................................................................................................. 11  
   OUTPUTS AND THROUGHPUTS ............................................................................................ 12  
   QUALITY ............................................................................................................................... 12  
V. How Have the Data Been Used? ....................................................................................... 15  
   PLANNING AND BUDGETING ............................................................................................. 15  
   MONITORING, EVALUATION, AND ACCOUNTABILITY ......................................................... 17  
   RESEARCH ............................................................................................................................ 19  
VI. Conclusions ...................................................................................................................... 22  
Appendix: A Review of Health Facility Surveys ................................................................... 24  
   DEMOGRAPHIC AND HEALTH SURVEYS (DHS AND DHS+) ........................................... 24  
   THE LIVING STANDARDS MEASUREMENT STUDY (LSMS) ............................................ 25  
   SITUATION ANALYSIS ....................................................................................................... 27  
   RAND SURVEYS .................................................................................................................. 28  
   MEASURE EVALUATION .................................................................................................... 29  
   WHO MULTI-COUNTRY EVALUATION OF IMCI .............................................................. 31  
   PUBLIC EXPENDITURE TRACKING SURVEYS .................................................................. 32
1. INTRODUCTION

Health facility surveys come in various guises. Unsurprisingly, all have as their focus the health facility. Beyond this, they vary in a number of respects. To understand these variations, it is helpful to have in mind Fig 1, which is a simple schematic showing how health facilities interact with one other, and with other elements in the health system, in the production of health outcomes. The latter are, of course, heavily influenced by households, who combine health services—and other inputs, including their own time but also non-health inputs—to produce good health. The household’s utilization of health services is shaped in part by factors within it (income, education, location, etc.) but also by factors beyond it, including the quality and other characteristics of health services. The quality of health services also influences the impact of utilization on health outcomes. The quality and other characteristics of health service providers are influenced in part by the actions of households and communities, but also by other providers (e.g. public and private providers are interdependent), and by government policies, including spending and regulation.

*Figure 1 – The routes to better health outcomes*

The health facility surveys undertaken to date vary along at least four dimensions. First, they have different motivations. Some aim simply to understand better what is actually happening at the facility level in terms of, say, resource usage, throughput, and the quality of care. Others aim to understand better how health facility characteristics influence health-seeking behavior and health outcomes. Yet others have sought to shed light on the linkages among providers—whether between different levels of the public system, or between public and private providers. In other cases, the motivation has been to better understand the links between providers and government. For example, some surveys have been designed to track public expenditures from the health ministry in the capital to the small facilities in rural areas. A second difference is that different surveys collect different data. While many collect data on inputs, not all collect data on costs or on the quality of care. This reflects in part the different
motivations of different surveys. A third difference is that surveys collect data in different ways. For example, data on quality are sometimes collected indirectly through data on inputs. In other surveys, they are collected directly, sometimes through clinical vignettes (in effect trying to infer the quality of care delivered from responses to hypothetical questions), and sometimes through direct observation. Finally, a fourth difference is that surveys vary in the uses to which the data have been put. Sometimes, the focus has been on research. In others, the principal use has been to design interventions, or to monitor and evaluate programs. For some surveys, the main use has been the feeding back of data to providers in an effort to further improve performance.

In this paper, we seek to provide an introduction to health facility surveys. We show in sections II through V how the surveys conducted to date vary along the four aforementioned dimensions. In section VI we offer some conclusions, focusing on lessons learnt and emerging themes.

II. WHAT MOTIVATES THE EXERCISE?

Provider-household linkages

In some cases, the motivation for collecting health facility data has been to understand better the link between health facilities, on the one hand, and the health-seeking behavior and health outcomes of households, on the other. This strand of work dates back to the 1970s, when the World Fertility Survey (WFS) started collected data to measure the impact of health service availability on fertility and mortality (Turner, et al. 2001). Data were not, in fact, collected initially at the facility level but rather at the community level, and were gathered through interviews with key community informants. This practice was continued in the context of the DHS surveys, which took over from the WFS in 1984.¹ Many LSMS surveys also included community modules that collected information on, inter alia, the availability of public services.

From collecting community data on service delivery, it was a natural extension to visit the actual service delivery units. A number of LSMS surveys (e.g. Côte d’Ivoire and Jamaica) in the late 1980s experimented with health facility and/or school surveys to complement the household data. A more systematic approach—Situation Analysis (SA)—was introduced by the Population Council in 1989, where the focus was on family-planning and reproductive health services. At least in part, the approach was motivated by findings emerging from the DHS surveys, which indicated that weaknesses on the supply side were important in explaining low contraceptive prevalence (Miller, et al. 1998; Miller, et al. 1997). Unlike the LSMS facility surveys, the SA did not also collect household data, although client interviews were carried out at the sample sites. More recently, facility surveys have been implemented in conjunction or coordination with DHS household surveys. These so called Service Provision Assessments (SPAs) are ambitious in scope, and seek to provide information about the characteristics of health services, including extensive information about quality, resource availability, and infrastructure.

¹ In the context of the DHS surveys, these community surveys were referred to as the Service Availability Module.
Provider performance

In other cases, the motivation for collecting health facility data has been to understand better the supply side of the health care system. Health facility surveys have proved a useful means of obtaining basic information on inputs, costs, outputs and quality. In principle, such information should be provided by a well-functioning management information system. Indeed, a number of studies of hospital costs have been carried out in both developed and developing countries on the basis of administrative data. However, management information systems often do not provide such data, and in this context a survey approach can be highly valuable. Once good data have been generated from survey data, they can be used to measure provider efficiency, and to analyze the determinants of and interrelationships between costs, efficiency, and quality (see, e.g. Angeles, et al. 1999; Bitran 1995; Somanathan, et al. 2000; Wouters 1993).

Links between providers

Another potential motivation for health facility surveys is to examine the interactions between providers. This can include the interactions between different levels within the public system, and the interactions between public and private providers. One issue that has received attention in the industrialized world is the impact of competition on provider behavior. Propper (1996) and Propper et al. (1998), for example, have explored the effects of competition on prices in the NHS hospital sector in the UK. They find that in the market for services financed by district hospital authorities, prices are lower in geographic areas where there is competition. In the market for services financed by fund-holding GPs, by contrast, the effect of competition on price is very limited. Soderlund et al. (1997) explored the effects of competition between hospitals on hospital costs (rather than prices), finding that competition, measured in terms of the number of providers within a 25-mile radius, does not have a significant impact. Similar research, which could clearly benefit from health facility surveys, does not appear to have been undertaken in developing countries.

Government-provider linkages

Another motivation for undertaking health facility surveys has been to examine the interface between government and providers. The Public Expenditure Tracking Surveys (PETS), for example, were motivated by a desire to be able to account better for the flows of public expenditures from central government to regions and districts and thence to providers in the periphery. The need for this type of tool has increased in recent years, in the wake of debt-reduction initiatives and a move towards sector and budget support under a “country-owned” policy framework. In large part, this shift in the approach to aid has resulted from increased appreciation of the fungibility issue and from finding that lack of government control over projects and the budget process more generally have seriously undermined long-term impact of aid. The premise of these initiatives is that sustainable development impact requires local

---

2 Examples of studies based on administrative data include Dor (1994b), Bitran-Dicowsky and Dunlop (1989b), and Anderson (1980b). For a general review, see Barnum and Kutzin (1993) and Wagstaff and Barnum (1989a).

3 Some work on competition and strategic interaction in the health care market has also been carried out in the US. For example, the Center for the Study of Health System Change (http://www.hschange.org) has carried out a number of “Community Tracking Studies”, some of which address these issues.

4 The World Bank (1998, p. 12) notes: “Aid agencies have a long history of trying to ‘cocoon’ their projects using free-standing technical assistance, independent project implementation units, and foreign experts – rather than trying to improve the
governance capacity, and an effective public expenditure process. In practical terms, this means that governments now face new pressures to demonstrate that resources are being channeled into increased and improved delivery of public services.

On other occasions, the motivation has been to better understand the influence that governments have—through the various policy instruments—over provider performance. There are few examples of surveys specifically addressed to these questions. A recent case is a survey implemented in Paraguay, carried out with support from MEASURE Evaluation, to identify and quantify the effect of transferring management control for basic health services provision from central to municipal government in Paraguay on costs, efficiency, basic health service quality, and patterns of health service use and equity (Angeles, et al. 1999).

III. WHAT DATA ARE COLLECTED?

Health facility surveys vary in the data they collect. Inevitably, this reflects, in part, the different objectives surveys have.

Inputs

A host of different inputs go into the production of health care services. It is customary to distinguish between capital and recurrent inputs, where capital inputs can be defined as inputs that last for more than one year, including vehicles, equipment, as well as buildings and other infrastructure. In contrast, recurrent inputs primarily comprise personnel, supplies (drugs, vaccines, other medical supplies, non-medical supplies), and operations and maintenance (primarily for vehicles and buildings).

Many facility surveys have collected data on inputs. There are three primary reasons for collecting these data. First, information about the availability and quantity of inputs can be used to assess the adequacy of resources at facility level. Second, and related to resource adequacy, the availability and characteristics of inputs at facility level affects the range, quality, and effectiveness of services provided. Finally, in combination with price information, data on inputs are required for the analysis of costs. It is of course the case that data on some facility inputs are often available from administrative records at national or sub-national level. However, a survey approach is motivated by a lack of confidence in these records, or by insufficient richness and detail.

Inputs may be divided into capital and recurrent inputs. The former are inputs that last for more than one year—e.g. vehicles, equipment, buildings, other infrastructure. Recurrent inputs primarily comprise personnel, supplies (drugs, vaccines, other medical supplies, non-medical supplies), and operations and maintenance (primarily for vehicles and buildings). The scope and level of detail of input data collected at facility level have varied considerably, in part reflecting differences in study focus. For material inputs, data can be collected on, *inter alia*, (i) quantity; (ii) quality; (iii) condition or state of repair; (iv) cost of acquisition; (v) price (in case of sale to public – e.g. drugs); (vi) criteria for allocation of shared inputs to programs or output
categories. For recurrent inputs, the data can cover both stocks and flows. For health facility personnel, the nature of the data is different.

**Capital inputs—buildings, vehicles, and equipment**

Many facility surveys collect some data on facility infrastructure and means of transport. In some cases, attention is restricted to the availability and functioning of basic infrastructural characteristics such as type of building, source of water and electricity, and toilet. In other cases, a wider range of infrastructure characteristics is considered, and some direct observation concerning hygienic conditions is also included. Finally, the availability of transport serves to reduce access time in emergency cases, permits outreach services, and may also reduce interruptions in supplies. With this in mind, information about the availability and state of repair of means transport (van, car, motorcycle, bicycle, etc.) has been collected in some surveys.

On the premise that the absence or non-functioning of a certain piece of equipment has an adverse effect on the ability of the facility to provide effective health services, many surveys collect data on the quantity or availability of medically related equipment. In some surveys, the list of equipment has been limited, including items such as refrigerator, sterilizing equipment, blood pressure monitors, scales, microscope, thermometer, stethoscope, beds. However, in other surveys, a more comprehensive list of equipment has been considered, covering, for example, lab equipment and clinical tools. In part, this has reflected the fact that the surveys have included larger and more complex health care providers. For the purposes of costing, quantities are required. In addition, it should be noted that clinical importance of a certain piece of equipment may not always be related with cost. As a result, the “minimum” list of equipment may differ depending on whether the primary concern is the analysis or costs or quality.

**Drugs and vaccines**

Most surveys have collected information about the availability of considerable range of drugs and vaccines. In some cases, questions have been limited to whether the drug is normally held in stock, whether it is currently in stock, and whether the facility had experienced stockouts (week or more without supply) during a recall period (e.g. six months). However, in other surveys, the drug related data are more comprehensive, and includes information on amounts usually prescribed and on the costs charged to patients. In surveys where costs are of primary interest, data collection must go beyond questions about availability to consider quantities and prices.

**Personnel**

Staff numbers and the staff mix are an important characteristics of the health facility, with an impact on the range of services that can be provided, the technical competence with which these services are delivered, as well as overall cost. With this in mind, most facility surveys have collected information on the number of staff in different categories, sometimes with additional information, e.g. qualifications, education and training, and terms of employment (full

---

5 For example, in the case of drugs, a survey can investigate drug availability at the time of the interview. This can be a useful quality indicator insofar as the current situation is representative of the typical state of affairs. However, in order to analyze costs, flow information (supplies to facility and use of drugs over a certain period) must also be collected.
or part-time). For the purposes of costing, data on salaries and other benefits are also required. Moreover, information on the allocation of staff time to different programs or categories of outputs extends the scope of possible cost analysis. Beyond staff numbers and basic characteristics, some surveys have sought to assess the clinical and interpersonal skills of health workers, as well as the perceptions of patients of the process of care. These are clearly important factors for the effectiveness of the care provided, as well as for the propensity for individuals to seek care from the respective health care provider. Approaches to collecting these data have varied, but include clinical vignettes and consultation observation.

Costs

Costs arise from the use of various scarce resources - including, patient’s time, volunteer services, in kind donations, etc. - in the production of health services. Information on costs is an important management tool, useful in determining resource requirements in the process of planning and budgeting. In addition, cost information is a necessary input into the analysis of cost effectiveness of health care interventions, and may serve as a guide for cost recovery policies. Finally, information on costs is also important for the evaluation of service delivery performance, in particular efficiency. In general, it is possible to observe five categories of resources at facility level: (i) staff time; (ii) supplies, medicines and other consumables; (iii) facility level administration and overhead expenses (e.g. utilities); (iv) equipment; and, (v) buildings and other physical infrastructure. In addition, there are costs incurred at higher level - e.g. training, supervision, central and sub-national administration, etc. – the evidence of which is not observed at facility level. We can further distinguish between recurrent and capital costs. Capital costs typically refer the use of resources whose useful operating life is more than one year (i.e. categories (iv) and (v) above). Costing is based primarily on input and price data. However, for some inputs, market prices may not be available, and some values must be imputed in order to find total cost estimates. Joint inputs – i.e. inputs that are used in the production of more than one type of service – comprise another challenge in costing. These inputs can be attributed to specific outputs on the basis of information collected in the survey, or on the basis of ad hoc assumptions.

Outputs, activities and throughputs

It is possible to conceptualize service delivery as a production process, where resources (inputs) are transformed into outputs. The “technology of production” refers to the relationship between inputs and outputs, and also determine costs. Information on outputs is clearly important from a management perspective, for example in establishing input requirements. However, the planning and management function is typically filled by administrative data collected on a routine basis. In contrast, survey data on outputs are typically collected with a view to analyze costs and the relationship between inputs and outputs.

How should the output of a health care provider should be appropriately measured? At one level, we may think that the appropriate facility output is the capacity to provide services, rather than the actual number of cases seen. This is because households, as demanders of health services, play a part in influencing utilization levels, so the amount of services delivered—utilization—is not entirely under the control of providers. For example, low levels of utilization may be due to low population density or a good overall health status in the catchment area.
Insofar as variation in output across providers is driven by this type of factors, service capacity would appear a theoretically appealing measure of output. However, in many cases, a divergence between service capacity and rendered services will be due to factors that are within the discretion of the facility. Here, of course, there is some scope for disagreement as to precisely what is under the control of providers. Insofar as there are influences that are genuinely beyond providers’ control, one could focus on the things that are directly under their control—for example, one might measure potential inpatient days rather than actual days, just as one might measure school places rather than school enrolment or attendance. Or one might stick with utilization-based measures but try to remove the confounding effects of influences that are genuinely beyond the control of providers.

“Readiness-to-serve” capability may be theoretically appealing, but may be difficult to measure. Indeed, most surveys collect data on the number of units of services delivered—e.g. outpatient visits, immunizations, antenatal visits, deliveries, inpatient days, etc.—as measures of output. Is this a satisfactory approach?

Problems with rendered services as output measure

Most health service outputs are highly heterogeneous. There are different sources of this heterogeneity. First, quality may differ considerably over time, across health care providers, and even between individual clients at a given health care provider. For example, a thorough outpatient consultation with a doctor is a very different service than a rushed consultation with a poorly trained nursing aide. While this is more important for some types of services than others (e.g. immunizations may be considered reasonably homogenous in terms of quality), this issue can create serious problems of comparability across time and space. Although it is in principle possible to control for quality differences in analysis, this has proved difficult in practice. Second, within a particular service category, there can be a notable variation in case mix and case complexity. Again, this causes problems in terms of comparability. For example, inpatient days can range from cases requiring simple interventions and limited monitoring, to highly complex cases where a broad range of material and human resources are required. Differences in case mix can arise from the differences in the socio-economic characteristics of the population in provider catchment areas, or because more complex cases may seek out providers with particular characteristics. The problem can be partly overcome. For example, output can be disaggregated into more detailed diagnostic categories. Alternatively, we can stratify providers by “service-capability groups” and assume homogeneity in terms of case mix and severity.

In addition to problems relating to the comparability of output measures in specific categories, most health care providers provide a wide range of services (service mix). Even with a small number of aggregated categories of interventions/services—e.g. inpatient days and outpatient visits—the issue arises of how to compare the quantity of outputs across providers with different service mix. A standard technique for dealing with multi-output production is to

---

6 For example, if households fail to take advantage of a free child immunization program, providers presumably can be blamed for a part of the non-utilization if they fail to publicize it and fail to employ delivery mechanisms that respond to the reasons for non-utilization (e.g. outreach rather than clinic-based services).
7 See e.g. G. Premushi, G. Rubio, and K. Subbarao, “Monitoring and Evaluation”, in PRSP Sourcebook.
8 See Thatchell (1983).
9 This was pointed out by Feldstein (1967a), who found that differences in case mix was an important explanation for variation in ward costs per case among hospitals.
construct an output index using market prices as weights. This approach is not appropriate in the health sector, where output prices do not exist or do not necessarily reflect production costs. Still, it is possible to construct composite output indices, although the weights are inevitably *ad hoc*. Moreover, for certain types of analysis the issue can be avoided. For example, if we are concerned with the estimation of costs, multiple outputs can be included as RHS variables.\textsuperscript{10} Similarly, for the estimation of a production function, Wouters (1993) suggests as an estimation strategy that output(s) can be included as RHS variable(s). For example, the number of outpatients can be regressed on the number of inpatient days and other explanatory variables. In this framework, the number of outpatient visits is estimated *given* the level of inpatient admissions at a facility.

**Quality**

As noted in the previous section, one of the main uses of survey data on facility inputs has been to assess health care quality, and, where household data are available, the effect of quality on health outcomes and health seeking behavior. But quality is clearly a concept that extends well beyond inputs.

According to a commonly used definition, health care quality pertains to the management of a personal health problem by a health care practitioner (Donabedian 1980). Following previous literature, Donabedian distinguishes between a technical and an interpersonal domain of management, where the first refers to the application of the science of medicine and other health sciences, and the second the social-psychological interaction of the health care provider and the patient.\textsuperscript{11} In this framework, the basis for making assessment about the quality of technical care pertains largely to professional norms based on a body of scientific knowledge. It is more complicated to evaluate the quality of interpersonal care, but criteria must clearly relate to socially defined norms of acceptability and good practice.

Donabedian suggests that the most direct approach to assessment of health care quality is to focus on the “process” of care – i.e., the activities of health care provider – and evaluate this against established technical, professional, and ethical norms. In contrast, quality of health care can be assessed more indirectly through an analysis of “structure” and “outcome”. Structure refers to the characteristics of the health care provider, and is relevant through the impact of the respective characteristic on the probability of a good process and outcome. Structural dimensions of health care include the tools and resources that health care providers have at their disposal, and the physical and organizational setting in which care is delivered (i.e. inputs). In contrast, outcome refers to the change in a patient’s current or future health status that can be attributed to antecedent health care, as well as changes in patient attitudes (including satisfaction), health knowledge, and health related behavior. This “structure-process-outcome” trilogy has been influential in structuring both health sector research and operational approaches to assess and improve the quality of health care in developing countries.

\textsuperscript{10} If we are willing to remain in a standard micro-economic framework, we can in principle derive the production function using the duality principle.

\textsuperscript{11} Donabedian underlines that the two domains are interrelated and often difficult to distinguish. He also distinguishes the “amenities” offered by the health care provider as a possible element of quality, but considers this part of the interpersonal management of an illness episode.
Each of the three approaches to assessing quality has strengths and weaknesses.\textsuperscript{12} The problem of the “trilogy” approach lies in the tenuous links between different dimensions of quality. For example, quality of structural inputs by no means assures good process. Similarly, the link between process and outcome is often not clear, and may not be visible for a long time. Also, favorable outcomes are often affected by factors not directly under the control of the health worker. This then raises the question of how to weight different dimensions of health care quality in an overall assessment.\textsuperscript{13} Are some dimensions more important than others? Are there important complementarities? Gilson et al. (1995) argue that quality of primary health care in developing countries is best assessed by looking at structure and process dimensions of quality. This is because primary health care is largely based on interventions for which effectiveness has already been demonstrated. In other words, it is more important to know why clients do not attend facilities and why interventions are not properly executed than to know their potential impact on health status.

\section*{IV. How are the Data Collected?}

Health facility surveys vary too in the way they collect data. The methods of data collection have varied across surveys, but include—depending on the data—interviews, record reviews, direct observation, and clinical vignettes.

\textbf{Inputs}

As noted in the previous section, the scope and nature of input data have differed across surveys, and the methods and tools for collecting the data have varied accordingly. Basic input data – e.g. on infrastructure characteristics, the availability of certain types of equipment, or the frequency and duration of drug stock-outs - can be gathered through an interview with the facility in-charge. However, information reported by facility staff may be imprecise, particularly in respect of detail. Direct observation of inputs may therefore be advisable. This permits verification of what drugs are actually in stock at the time of the interview, as well as an assessment of the state of repair and level of functioning of essential inputs (e.g. temperature of fridge). A record review is required if data on the flow of resources to the facility (e.g. number of drug kits received in a six month period) or historical data are required. This approach however relies on the availability and reliability of records – for example stock cards – at facility level, a condition that is often not met. Finally, because of the paucity of records at facility level, some input data are best collected at administrative levels above the actual facility. This may include payroll data, as well as information on the supply of medical and non-medical supplies to the sampled facilities.

\textsuperscript{12} In a review of studies of quality of health care in developing countries, De Geyndt (1995) notes that outcome measures are generally absent in the literature. Ackerberg et al. (2001) provide an exception in their comparison of the performance of different providers in Maine’s alcohol abuse treatment program. In contrast, most studies have focused on structural or process indicators (or both). Notably, researchers from the economics discipline have tended to focus on quantifiable and measurable structural variables, while process variables are more often used by medical practitioners.

\textsuperscript{13} We may of course decide that the concept of overall quality is a meaningless concept.
Costs

Many studies of costs have relied on administrative cost data. We may refer to this as a “top-down” approach. Top-down costing relies on administrative data from central or lower level. These data can take many forms, including disaggregated budget allocations or financial records at hospital or sub-national administrative level. While this may be a convenient and relatively cheap way of gathering data, the scope for analysis can be limited by poor quality and insufficient disaggregation. These problems are likely to be particularly severe in developing countries, and for primary health care providers. Also, records of this nature may not exist or be available for NGO or for-profit providers.

In contrast to the top-down approach, the bottom-up approach proceeds by quantifying the use of resources at facility level. There are a number of issues to consider here. First, there is the question of which costs to consider. In facility based costing, a trade-off has to be made between detail and sample size. At one end of the spectrum, a survey can collect data on basic output indicators and on key input use in the production of services. At the other end, a full resource costing considers a more comprehensive range of inputs, and allocates resources used at provider level to specific service categories. This typically includes some form of time and motion study, and a recording of all goods and services (direct and indirect) associated with different categories of services. In this way, the resource costs needed to treat the average case can be calculated. This type of study can provide interesting perspectives on efficiency. However, the limitations of this approach are obvious; detailed costing is very time intensive, and attention must typically be restricted to one or a few health care providers. For example, Lewis et al. (1996) studied only a single hospital, although in considerable detail. One of the advantage of the bottom-up approach is that it is easier to establish a link between the use of a particular resource and a specific category of output (service).

Second, many inputs are shared across activities. For example, most facility staff are involved not only in the provision of outpatient services, but also in immunizations, deliveries, and inpatient care. Similarly, certain pieces of equipment may be shared across many categories of services. In terms of costing specific services, this raises the question of how these respective inputs should be allocated across the services in which they are used. This can be dealt with on the basis of ad hoc assumptions, or using information collected in the survey – for example on the proportion of time staff spends on different types of activities.

Finally, there is a question of how to treat capital inputs. Capital costs arise from the use of assets in the production of services, and include the opportunity costs of the initial investments and the depreciation of the asset over time. Due to the difficulties involved in measuring capital costs in a meaningful way, they are often not considered in facility surveys.

14 Lewis et al. (1996) comprise a rare example of this approach being used in a developing country. Through a detailed study, they estimate costs for outpatients, inpatient and emergency care, taking into account case mix, clinical norms, indirect and imputed costs, as well as depreciation of physical infrastructure and equipment. They find that although personnel represent only 2.5 11.5 and 5.1 percent of total hospital costs for emergency patients, outpatients, and inpatients respectively, 84 percent of hospital budget in the Dominica Republic pays salaries. This indicates a high degree of waste in the system, and clearly demonstrates the possible divergence between expenditures and actual use of resources in the delivery of specific services.

15 “Step down analysis” is a standard methodology for distributing aggregate costs across department (cost centers), and, ultimately, to final service categories (with measurable outputs). For a review of issues in the allocation of overhead costs, see, e.g. Drummond et al. (1997)
Outputs and throughputs

Pragmatism may dictate that we measure facility outputs by the units of services delivered—i.e. outpatient visits, immunizations, is an appropriate measure of output. The level of detail (i.e. the number of service categories) is to some extent context specific, and will depend on the way in which patients are registered at facility level. There is however some choice in the source of output data. Information on service outputs can be collected from either the patient register at facility level, or from administrative records at district or provincial level. In principle, these sources should be consistent. However, results from a recent service delivery survey in Uganda indicate that there may be considerable discrepancy (Figure 2 below shows the tendency towards over-reporting) (Lindelow, et al. 2001).

![Figure 2 – Comparison of facility and district records in Uganda](image)

Why do we observe this discrepancy? There are several possible explanations. First, it is possible that health facility staff know that submitted patient number will not have any real effect on the activities and resources of the facility, and that the numbers will not be checked. If so, facility staff may decide that it is not worth expending effort on accurately reporting patient numbers. Second, if resources are allocated on the basis of reported patient numbers, or if these records are used to define user fee revenue targets against which the facility is held accountable, there may be incentives to over- or under-report patient numbers. In either case, it is clear that output data must be collected with some care.

Quality

Although there is little doubt that health care quality has potentially large effects on both health outcomes and behaviors, there is clearly much disagreement about the appropriate way of defining and measuring it. For example, Bessinger and Bertrand (2001) describe the difficult process of reducing the number of indicators of quality in family planning services from 200 to 25.\(^\text{16}\) Even after this considerable simplification, the recommended approach requires data collection through consultation observation, a facility “audit”, and interviews with clients. This

---

\(^{16}\) The “Quick Investigation of Quality” (QIQ) approach is described in Bertrand et al. (Bertrand, et al. 2001).
then raises the question of how to best measure quality in a more general sense. This question has two components: what dimension of quality should we try to measure, and how do we best go about doing it? Most likely, it is not possible to reach full agreement on this, at least not in the short term. Still, findings from research on quality to date can provide some guidance.

**Quality measurement through inputs**

The most common approach to measuring the structural dimension of quality is to collect data on the availability and quantity of inputs considered important for the delivery of services. Quality can then be proxied by focusing on particular inputs, for example drugs. Alternatively, composite quality indices can be constructed. Quality data of this kind permit an analysis of the relationship between quality and health outcomes and behavior.\(^\text{17}\) Observed variation in “structural quality” also raises the question of what explains these differences. While some of these indicators are clearly important and have the benefit of being easy to measure, they are likely to be inadequate as measures of quality.\(^\text{18}\)

**Quality measurement through clinical vignettes**

Good health care quality, as pertaining to the overall management of a health problem, clearly requires more than material inputs. In this regard, clinical skills are of obvious importance, and many surveys have tried to collect data concerning the process of care. Here, the general principle is to assess case management against established diagnostic and treatment norms. The most common ways of assessing “process quality” is through clinical vignettes and consultation observations.\(^\text{19}\) In addition, some surveys (e.g. DHS SPA) have asked questions about general facility procedures, for example in respect of sterilization, disposal of needles, etc.

Clinical vignettes assess clinical skills by presenting health workers with hypothetical cases, and seek to elicit information about the way in which they would handle such a case. This approach has the advantage of permitting a considerable degree of structure in the survey instruments, thus limiting the need for technical skills in the enumerators. It has been used in a number of surveys, including the Jamaica LSMS and the IFLS. A more ambitious approach entails using highly qualified enumerators, who observe the case management of health workers in the respective facilities. While the skill requirements for this type of survey can be prohibitive in many contexts, the approach has the advantage of permitting a more comprehensive assessment of the process of care. For example, beyond the strictly clinical aspects of the case management, the enumerator can assess the social-psychological interaction between the health care provider and the patient. This can include the general attitude and courtesy of the health worker, but also communication of information about treatment regimes, medication, return visits, etc. On the other hand, direct observation of patient-provider interactions may suffer from bias due to the behavioral response of the health worker to being observed. In addition, differences in case mix across providers can confound the assessment of quality. Observation of

\(^{17}\) Evidence from this literature is discussed below.

\(^{18}\) One reason for this is that we may expect only a weak link between structural dimensions of quality and health outcomes. As noted by Gilson et al. (1993), process quality is more likely to be validated in terms of effectiveness than structure.

\(^{19}\) Consultation observation is often combined with re-examination of the client. It is also possible to carry out ex post reviews of patient records, or chart abstractions. This approach however relies on records that are often inadequate in the case of outpatient care and/or unavailable in developing countries (Peabody, et al. 2000).
case management has been a feature in surveys such as the WHO multi-country evaluation of IMCI, SPAs, and Situation Analysis.

What is the best approach – clinical vignettes or consultation observation? Either approach has its advantages. Clinical vignettes are easy to administer, and to not necessarily require enumerators with sophisticated medical training. Consultation observation, on the other hand, provides more direct evidence from a clinical setting and permits observation of non-clinical aspects of the provider-client interaction. That said, the presence of an observer may lead the provider to adjust his or her behavior. Assuming that the provider effectively applies clinical knowledge in case management, and that observation bias is not significant, we would expect a close correspondence between quality measures from vignettes and clinical observations. There is indeed some evidence from the US that this is the case; Peabody et al. (2000) report evidence that vignette scores in an outpatient setting appeared to reflect physician practice as recorded in observation episodes. On the other hand, it is not clear that these findings can be generalized. As noted by Peabody et al., the costs – in terms of time and effort – of complying with diagnostic and treatment procedures in a clinical vignette are negligible. This is not the case in a clinical setting. As a result, we would expect vignettes to be biased in favor of high quality in a context where external and/or internal incentives for compliance are low. Reflecting this concern, Leonard et al (2000) report evidence from Tanzania that while most clinicians knew how to deal with certain tracer conditions (on the basis written examination), many did not actually do the right thing when observed in a clinical setting.

Quality measurement through beneficiary perceptions

The assessment of quality through structural and process characteristics is premised on a set of established norms for clinical effectiveness. An alternative approach to assessing quality of care is to ask beneficiaries about their perceptions. Patient satisfaction is important in its own right as an objective of health care, but also as an indicator of the structure, process and outcome of health care, and through the effect of satisfaction on client behavior. Patient satisfaction is typically assessed through interviews with actual or potential patients. Data on actual patients can be collected through exit polls, whereas a sample of households or individuals in the community must be surveyed in order to capture the views of non-users. Multi-topic household surveys with added facility modules, such as the LSMS surveys, have typically not collected detailed data on patient perceptions, whereas exit polls have often been a feature of more focused health surveys (e.g. WHO MCE, SPAs, and Situation Analysis surveys).

There is some evidence to suggest that observations and client exit interviews provide similar results for many (family planning) quality indicators (Bertrand, et al. 2001). However, it is also known that there are considerable problems in interpreting subjective perceptions of health care quality. In part, this is due to “courtesy bias”, such that individuals may provide responses that they deem socially acceptable. However, it is also the case that subjective

---

20 Data can also be collected through focus group discussions. This option is not discussed here.
21 For example, Bitran (1995) found that regardless of facility scores on quality of diagnostic and treatment norms, and availability of drugs and supplies, patients generally reported being satisfied with the treatment they received, and indicated that they would return for future treatment. Many other studies have noted an inconsistency in responses, where high levels of overall satisfaction are combined with widespread complaints about specific aspects of care or the health care system. With a view to reconcile these findings, it has been suggested that positive and negative comments about health care are typically not reducible to satisfaction and dissatisfaction, and a more careful methodology is required to collect meaningful qualitative data on patient perceptions (Atkinson 1993).
perceptions of health care quality depend on both needs and expectations, which in turn are underpinned by the explanatory models held by different individuals regarding the functioning of the body, etiology, and the appropriateness of different remedies and interventions in response to a given symptom (Atkinson 1993). There may be important systematic differences across demographic and socio-economic groups in respect of these explanatory models, possibly making client perceptions poor proxies for objective assessments of different dimensions of quality.

There are two other approaches relating to client perceptions and satisfaction that, despite not relying strictly on facility surveys, deserve mention in this context. First, following the publication of the 2000 World Health Report (WHO 2000), the concept of health system responsiveness has gained increased currency. The concept of responsiveness is different from patient satisfaction in that it is seen to pertain solely to non-health enhancing aspects of the health system, and in its focus on health systems rather than individual providers (de Silva 1999). In addition, the methodology for measuring responsiveness seeks to overcome some of the perceived weaknesses of measures of patient satisfaction, in particular in respect of relativity arising from differences in expectations and experiences. While the methodology is far from uncontested (see, e.g. Williams 2001), the framework may contribute to improved measurement of client satisfaction in the context of facility surveys.

Second, there is a host of approaches that seek to gather “community-based evidence” on service delivery. This includes service delivery surveys, “report cards” on public services, and other related approaches. These approaches cannot in general be described as facility surveys, nor are they always limited to the health sector. Rather, the purpose is to use a range of different tools – e.g. household interviews, focal groups, institutional reviews of service providers, interviews with service providers – to elicit information about awareness and experiences of services, and opinions about prices, quality, waiting times, courtesy, etc. Despite their simplicity, the approaches can usefully inform the formulation of hypotheses about service delivery, as well as the design of quantitative surveys. Moreover, they have proved effective (at least from a monitoring and advocacy perspective), and may in the future complement and support facility surveys.

V. HOW HAVE THE DATA BEEN USED?

Planning and budgeting

Planning and budgeting in the health sector concerns decisions about the use of resources in the future – for example, in respect of mix of services or the location and means of provision. Although inherently a political process, information is a fundamental input in all the stages of planning and budgeting, including situation analysis, priority setting, option appraisal, programming, implementation and monitoring, and evaluation. Much of this information will be

22 The different aspects of responsiveness considered in the WHO methodology are: dignity, autonomy, confidentiality, prompt attention, quality of basic amenities, access to social support networks during care, choice of care provider (de Silva 1999).
23 For service delivery surveys, see e.g. Cockcroft et al. (1999) and CIET international (1996). See also www.ciet.org. For report cards, see, e.g. Paul (1992; 2000) and World Bank (2001a).
generated by routing health information system. However, facility surveys can provide valuable non-routine information to support decision making.

First, facility surveys can provide a detailed picture of resource availability and adequacy and costs at facility level. The premise in assessing “resource adequacy” is that a certain range, quantity and quality of inputs is required in order to provide the intended service package. In this sense, resource or input adequacy can be assessed by analyzing how resource or input availability relates to resource requirements. The resource requirements clearly relate to the package of services that the respective facility should be delivering, as well as to the amount of clients. Hence, in order to make a meaningful assessment of resource adequacy, a careful estimate must be made for each type of input on the basis of information about the service range and the activity level of the facility in question. For example, an assessment of staff adequacy may be based on a minimum standard for any facility, and on the patient load for staff (Hanson and Gilson 1996). Similarly, there are established methods for estimating drug requirements (WHO 1988). In addition to looking at specific inputs, composite indicators can be constructed to provide information on the capacity of the health facility to perform basic clinical tasks. This type of exercise can focus on defined “tracer conditions”. In general, this type of information can be used to estimate total resource needs and to support decisions about the allocation of resources within the sector. In principle, surveys can be used to estimate unit costs for the delivering different types of services in different contexts. This could then form the basis for calculating total resource needs to meet certain service standards, and for making decisions about the allocation of resources within the sector. In practice, however, surveys may not be the most appropriate tool for costing basic service packages. However, survey information on the availability of human and material resources at facility level can point at egregious imbalances in how resources are allocated across inputs.

Second, it should be noted that while resource inadequacy at the level of health care facility is likely to be the consequence of a general shortage of resources in the sector, these problems may be aggravated by weaknesses in the budget execution system, including systems of procurement and distribution of in-kind resources. For example, Public Expenditure Tracking Surveys have shown how delays and leakage in budget execution can result in budget outturns that depart considerably from intended allocations. Results from these surveys are discussed in greater detail below.

24 Most facility surveys have taken a “checklist approach” to assessing resource availability, focusing on the availability of certain types of equipment, staff, drugs and other inputs. However, in the analysis of input data, the availability and quantity of inputs are often not related to output variables, making it difficult to make any statements about the adequacy of resources. There are only few examples of surveys that have used tracer conditions. One example is a survey of rural health facilities in Papua New Guinea, aimed at assessing the “potential quality” of each unit (Garner, et al. 1990). In this case, a range of items on a checklist was scored to construct a “quality indices”, reflecting the capability of each facility to perform certain common clinical tasks. The study analyzed differences in resource adequacy across, *inter alia*, agency types and regions.

25 Comprehensive costing is typically beyond the scope of most surveys. In addition, survey based costing is likely to result in considerable variation in unit cost across unit. This variation may have many explanations, including differences in efficiency, quality, case mix, and input prices.
Monitoring, evaluation, and accountability

A number of survey initiatives have for many years served as bases for continuous assessment of health status and living standards more generally. These include household surveys such as the Demographic and Health Surveys (DHS), the Multi-Indicator Cluster Surveys (MICS), and the Living Standards Measurement Study (LSMS). Data from these surveys have proved a useful supplement to data gathered from surveillance sites, such as the ICDDR,B Centre for Health and Population Research in Bangladesh\(^\text{26}\) and the Indepth Network\(^\text{27}\), as well as from more traditional sources, such as censuses and vital registration systems. The attraction of household surveys is that they can be stratified by, for example, gender, location, ethnicity, and household living standards.\(^\text{28}\) In addition, surveys provide more representative information that data collected routinely at facility level in contexts where access to health care is limited.

Similarly to surveys carried out at household level, semi-standardized surveys of health care providers – for example, Service Provision Assessments (SPAs) or Situation Analysis (SA) - have provided valuable assessments of health care quality, infrastructure, utilization and availability. For example, a recent survey in Kenya (Macro International 2000) found that although contraceptive services, drugs and materials are widely available, compliance with infection control procedures, including sterilization of equipment and sterilization of needles was often poor. Also, consultation observations revealed that the diagnostic process was often unsatisfactory, clients were frequently provided with incomplete or insufficient information to chose between interventions, and to ensure compliance with treatment regimes. Aside from providing data for research on the determinants of health outcomes and health related behavior, this type of information is valuable in the design of health sector strategies and interventions. Moreover, repeat surveys permit the monitoring of changes in quality over time.

In other cases, household and/or facility surveys are implemented as components of the monitoring and evaluation of specific projects or programs. As noted by Mills (1997), many health sector reforms have been proposed by donor agencies and health sector specialists in recent years, but there is little evidence that can be put forward on the impact of past reforms in terms of quantitative measures of efficiency or equity. In this context, efforts aimed at continuous assessment and ex post evaluation are of particular importance. Under ideal circumstances, this will entail (i) a baseline survey; (ii) subsequent policy change in a random sub-sample of units; and (iii) a follow-up survey of both treatment and control facilities. However, in many cases evaluations have to be carried out under less than perfect conditions.

Several examples of surveys implemented in the context of project or program evaluation can be mentioned.\(^\text{29}\) For example, two household and facility surveys have been carried out in Uganda (1995 and 1997) as components of the ongoing DISH project (reproductive health). The surveys have been implemented to measure changes in reproductive, maternal and child health knowledge and behavior (Kantende, et al. 1999). An ongoing Multi-Country Evaluation (MCE)

\(^{26}\) http://www.icddrb.org
\(^{27}\) http://www.indepth-network.org/
\(^{28}\) See http://www.worldbank.org/poverty/health/data/index.htm for a 42-country study that disaggregates various key MCH outcomes by poverty status.
\(^{29}\) In many cases these efforts are ongoing, so it is hard to say how successful this approach has been in analyzing the determinants of facility performance. Many of these surveys are discussed in greater detail in the appendix.
of the Integrated Management of Childhood Illnesses (IMCI) comprises another example of the role that facility surveys can play in the monitoring and evaluation of health sector interventions. Worldwide there are 30 countries at different stages of implementation of IMCI among which Uganda, Peru, Bangladesh and Tanzania will participate in the MCE. The purpose of the MCE is to document the effects of IMCI interventions on health workers performance, health systems and family behaviors; determine the extent of measurable impact on health outcomes; and, describing costs of implementation (WHO 2001). Health facility surveys also formed part of the evaluation strategy for the Bolivian Social Investment Fund (SIF), which was established in 1991 to direct public service investments to poor communities. Surveys were collected in 1993 and 1997, and included households, communities, and service delivery units. The surveys demonstrated clear improvements in infrastructure and equipment, as well as increased utilization rates and reduced mortality. As a final example, a survey was carried out in Paraguay in 1998 with a view to identify and quantify the effect of transferring management control for basic health services provision from central to municipal government on costs, efficiency, basic health service quality, and patterns of health service use and equity. The study uses pre- and post-decentralization design, with a control group, and focused on family planning, maternal health, and infant and child health (Angeles, et al. 1999).  

Monitoring information can also serve to strengthen accountability in the delivery of services. This can include the use of survey methods both to shed light on the systems through which budget resources are transformed into services, as well as to give users an opportunity to express their views on public services. The Bank’s Public Expenditure Tracking Surveys (PETS) and Quantitative Service Delivery Surveys (QSDS) provide examples of the former type of surveys. The focus of these non-standardized surveys have varied. On the one hand, some surveys have been concerned with the actual flow of resources - in financial or real terms - from central level administration to the actual service delivery unit. Indeed, it was in response to these concerns that the first Public Expenditure Tracking Survey (PETS) was designed and implemented in Uganda in 1996 (see Ablo and Reinikka-Soininen 1998; Reinikka 1999). This survey asked why it was that service delivery outputs in the health and education sectors had failed to respond to considerable increases in public spending. Data problems were severe, particularly in the health sector (a finding of some interest in its own right). Yet, the survey provided powerful evidence of how weak governance and lack of accountability can hamper the process of budget implementation. For example, on average only 13 percent of the total yearly capitation grant from the central government reached the school in 1991-95. Eighty-seven percent either disappeared or was used for purposes unrelated to education. Interestingly, the access to funds varied in systematic ways with school size, income, and the extent to which teachers were qualified. Information asymmetry was seen to be the primary factor behind these findings. In response to the survey results, the government began publishing the monthly transfers to the districts in newspapers, broadcasting on radio, and requiring primary schools to post information on received funds in public places in the school.  

---

30 A second round of data collection is planned, but has been delayed due to slower than expected implementation of decentralization.  
31 In the health sector records were poor or non-existing, and little quantitative data were collected. On the basis of interviews, it was however estimated that on average, almost 70 percent of medical supplies and drugs were expropriated by staff. Similar findings are reported by McPake et al. (1999)  
32 Follow-up surveys have documented a considerable increase in the proportion of funds received by schools. This can at least in part be attributed to improved access to information, although other factors are also likely to be important.
In other surveys the focus has been on the service delivery unit, although with an eye on the relationship between the facility and the administrative system upstream. This is in recognition of the fact that even if budget outturns reflect allocations, ultimate impact in terms of development objectives also requires that financial resources are efficiently transformed into appropriate public services.\(^{33}\) This can include facility surveys, but also citizen surveys or scorecard approaches. For example, the original “Bangalore Scorecard”, designed and implemented by the Public Affairs Centre (PAC), a local NGO, questioned a sample of both rich and poor users on their perceptions of the city’s public services. The survey was used to rate different agencies in terms of staff behavior, quality of service, and information provided, and proved a powerful tool to generate feedback on public services, identify weaknesses in service delivery, and advocate for change (Paul 2000).

Both tracking surveys and service delivery surveys are increasingly used as monitoring and accountability mechanisms in contexts where routing management information systems are weak. Indeed, Uganda intends to place the conduct of regular service delivery surveys at the centre of its strategy for improving public services, with baseline surveys to establish strategies and targets, and public dissemination of survey findings.\(^{34}\)

**Research**

**Quality and health outcomes**

As data on health care quality has become increasingly available, the body of empirical research on the link between health care quality and health outcomes has grown. There is now evidence that certain characteristics of health care providers and their relationship with clients are important determinants of health outcomes. For example, a number of studies have found a positive correlation between health service infrastructure and characteristics, such as the number of clinics or doctors per capita, and health indicators, including child mortality, fertility, nutritional status and anthropometric indicators (see e.g. Beneo and Schulz 1994; Hossain 1989; Lavy, et al. 1996; Rosenzweig and Schultz 1982; Rosenzweig and Wolpin 1982; Thomas, et al. 1996).\(^{35}\) There is also a literature based on DHS data, focusing specifically on family planning and contraceptive use (see, e.g. Beegle 1995; Feyisetan and Ainsworth 1996; Mroz, et al. 1999; Steele, et al. 1999).

**Quality and health seeking behaviour**

One of the conduits through which health care quality affects health outcomes is client behavior, including both care seeking behavior (utilization of health services) and adherence behavior (i.e. compliance with treatment regimes, follow-up visits and referrals). However, concern with quality was limited in early studies of health care demand.\(^{36}\) To the extent that health care quality was considered at all, it was treated as an unobserved variable, pertaining to

---

\(^{33}\) These surveys have been referred to as Quantitative Service Delivery Surveys (QSDS) (Dehn, et al. 2000).


\(^{36}\) Several contributors (e.g. Akin, et al. 1984; 1986; Litvack and Bodart 1993) have noted that excluding quality from the purview of analysis is problematic on methodological grounds. Higher price may simply reflect higher quality, and, insofar as clients are willing to pay for quality, the failure to control for it in estimation of health care demand will result in underestimation of the price effect.
the provider type rather than the individual provider (e.g. Gertler, et al. 1987; Gertler and van der Gaag 1990). More recently, there have been studies that have used cross-sectional data to examine the effect of a series of structural quality variables on health seeking behavior. In some of these studies, multicollinearity across quality characteristics and service type prevented the examination of independent impact of each characteristic. Generally, the studies demonstrate a significant and sometimes large statistical correlation between quality and utilization. However, as Gertler and Hammer (1997) note, with the exception of the experimental study by Litvack and Bodart (1993), this may reflect the effect of utilization on pricing and quality policy rather than the other way around.

There is plenty of survey evidence of low quality in the process of care. This relates both to clinical (clinical knowledge, adherence to diagnostic and treatment protocols, etc.) and non-clinical (inter-personal relations) aspects of the process. There is however less evidence on the effect of different aspects of the process of care on health outcomes and behavior. An exception in this regard is Peabody et al. (1998), who use the LSMS Jamaica data to study the relationship between quality and birth weight. While none of the structural measures of quality had a statistically significant effect, they found that women who had access to complete examination (process) had infants that weighed an average of 128g more at birth.

Researchers have also used facility and household data to study the phenomenon of bypassing – i.e. when patients reject a closer facility in favor of a more distant health care provider. There is only a limited literature on this issue, mostly focusing on the case where households bypass public facilities in favor of the private sector (Akin and Hutchinson 1999; Leonard, et al. 2000). In large part, this reflects the complexity of the data required to analyze this aspect of health seeking behavior. In general, studies have found that bypassing is associated with facility quality (e.g. number of doctors, drug availability, opening hours). However, the studies have also demonstrated that care seeking behavior is often a sophisticated response to the type and severity of illness of the client.

Econometric estimation of cost functions

Cost functions can be useful in shedding light on a number of questions. Are health facilities under- or over-capitalized? Are facilities inefficient in other respects? Are they technically inefficient in that they fail to produce the maximum output from their inputs? Do they use inputs in the wrong proportions? Do facilities vary in their degree of inefficiency? Are public facilities less efficient than others? Should facilities specialize or provide a broad range of services? Could costs be reduced by concentrating cases in fewer facilities?

In the industrialized world, there is a large literature on these topics. A much smaller literature exists in the developing world. Wagstaff and Barnum (1992) review the four principal studies up to 1992 (Ethiopia, Kenya, Nigeria and Peru). Barnum and Kutzin (1993) report

---

37 Mwabu and Ainsworth (1993) consider drug availability and staffing; Lavy and Germain (1994) measure the quality of various provider types through drug availability, staffing, infrastructure, and service range; Akin et al. (1995) use operational cost per capita as reported by facility as the preferred proxy for quality, but also consider measures relating to the physical conditions of facilities and availability of drugs (percentage of year) to control for differences in efficiency. Litvack and Bodart (1993) focus on drugs as the primary component of quality. Although the scope of the study is narrow, its methodological strength is due to the experimental design.

38 The welfare implications of bypassing are not immediately obvious, but it is clear that under some circumstances, bypassing can lead to low utilisation of available public facilities (resulting in low efficiency) and impose unnecessary costs on the clients.
estimates of cost functions for hospitals in Colombia and China. The published literature during
the last 10 years includes a large volume of cost function studies from industrialized countries
but apparently none from the developing world. Some of the studies (e.g. Nigeria, Colombia,
China) that have been done in developing countries—in contrast to those undertaken in the
industrialized countries—have relied on survey data, which is unsurprising in view of the poor
quality of data in most developing country health facilities (Lewis, et al. 1996; Mills, et al.
1993).

Few of the studies to date shed light on the issue of whether health facilities have too
little or too much capital. The test proposed by Cowing, Holtman and Powers (1983) is to see
whether the slope of the variable cost function with respect to capital is equal to the price of
capital. Wagstaff and Barnum (1992) suggest an alternative test similar to Feldstein’s (1967b),
namely to test the hypothesis that the slope of the short-run total cost function with respect to
capital equals zero. Anderson’s (1980a) study of Keynan hospitals in inconclusive on the over-
capitalization issue. The results of the studies from Ethiopia (Bitran-Dicowsky and Dunlop
1989a) and Nigeria (Wouters 1993) are consistent with over-capitalization, but Wagstaff and
Barnum warn against taking the results too literally.

In many studies the primary focus has been on efficiency. Broadly speaking, the concept
of efficiency pertains to the relationship between inputs and outputs.39 In modern efficiency
measurement it is customary to distinguish between technical and allocative efficiency (Farrel
1957).40 Technical efficiency refers to the maximization of output with a given set of inputs. In
contrast, allocative efficiency reflects the substitution between inputs with different prices to
achieve minimum costs. These two measures can be combined to provide a measure of total
economic efficiency. However, applying the concept of allocative efficiency in the public sector
is often fraught with methodological difficulties.41 In consequence, many studies of efficiency in
this context restrict attention to technical efficiency.42 Much of the recent literature on efficiency
of health care providers in the industrialized world has employed a statistical frontier model (see
e.g. Dor 1994a; Førsund, et al. 1980; Li and Rosenman 2001; López Casanovas and Wagstaff
1996; Rosko 1999; 2001; Rosko and Chilingarian 1999; Vitaliano and Toren 1994; Wagstaff
1989b; Zuckerman, et al. 1994). The traditional cost function links costs to outputs, input prices,
and—in the short-run—capital stock. The cost frontier model extends this by introducing a term
that represents the excess above the minimized cost, given the facility’s outputs, input prices, and
so on. Kutzin and Barnum use a frontier model to test for inefficiency in their sample of Chinese
hospitals, but find none. Dor (1987), in his study of Peruvian hospitals, does not use a frontier
model, but includes in his cost function a dummy variable to indicate whether the hospital is
operated by the Ministry of Health (MoH) or the Social Security Institute (SSI). His findings
suggest that the MoH hospitals are more efficient than those operated by the SSI. Inefficiency in

39 In the broadest sense, efficiency can be seen to concern the relationship between inputs and outcomes. Using this broad concept
of efficiency would however require data that are not typically available.
40 The efficiency concepts were originally developed in relation to firm performance. See Fried et al. (1993) for a comprehensive
treatment.
41 First, the choice of inputs is often beyond the control of the individual facility individual facility (at least in the public sector).
where there is discretion, price signals may be weak. Second, cost minimization may not be an appropriate behavioral
assumption. For example, staff allocation may be driven by a policy to ensure minimum service standards. As a result, we should
not necessarily expect to observe allocative efficiency.
42 In cases where the allocative efficiency in the health sector has been addressed, this has typically been done by comparing the
ratio of marginal products to the ratio of staff remuneration for the respective categories. Marginal products are calculated on the
basis of an estimated production function with, inter alia, different categories of staff as its arguments. Ref? Grossman &
Goldman.
a cost frontier model can be due to technical inefficiency, or input-mix inefficiency, or both. Wouters (1993), in her study of Nigerian facilities, explores the issue of input-mix inefficiency. She finds evidence in her sample of under-employment of health workers relative to non-health workers, and that inefficiency is greater in the private sector.

Another important issue confronting policy-makers is whether hospitals should be encouraged to specialize or to provide a broad range of services. Should hospitals, for example, provide both inpatient and outpatient care? Should hospitals providing inpatient services aim to treat most casetypes, or should they specialize? This is the issue of economies of scope—whether costs can be reduced by producing two or more products jointly. In many studies of health facilities, the functional form used in the regression analysis is too restrictive to explore this issue. Of the four studies surveyed by Wagstaff and Barnum (1992) only one employ a specification that is sufficiently general not to prejudge the issue of economies of scope. Both Anderson’s study of Keynan hospitals and Wouters’ study of Nigerian facilities employ a multiproduct Cobb-Douglas production, which, as Baumol et al. (1982) note, implicitly assumes cost anti-complementarities and hence assumes diseconomies of scope unless there are sufficiently strong offsetting fixed costs. Dor's specification, in his study of Peruvian hospitals, is less rigid but is consistent with economies of scope only in the implausible case where the marginal cost of an outpatient visit is negative. Only in the Bitran-Dunlop study of Ethiopian hospitals is the model specification sufficiently general not to prejudge the issue of economies of scope. The results imply mild economies of scope—a result that is similar to that reported by Kutzin and Barnum in their analysis of Chinese hospitals.

A further important issue facing policymakers in developing countries is whether the current number of facilities should be increased or reduced to cope with the existing workload. This depends on how far economies of scale have been exploited—the conventional wisdom being that unexploited economies of scale imply too many producers. In the multiproduct case (where, for example, hospitals are treating inpatient and outpatient cases), one needs to work with the concept of ray economies of scale, which show how costs change when all outputs are increased in the same proportion. Ray economies of scale depend in part on product-specific economies of scale, and in part on economies of scope. Two of the studies reviewed by Wagstaff and Barnum prejudge the issue of economies of scale. Dor's specification, for example, implicitly assumes that the product-specific economies-of-scale index is equal to one for all outputs. Given this, the link between ray economies of scale, economies of scope and product-specific economies of scale, and that the specification in effect rules out economies of scope, it follows that the specification forces ray diseconomies of scale. Bitran and Dunlop find slight product-specific diseconomies of scale but find ray economies of scale. The implication is that these stem from the economies of scope noted above. Wouters also finds ray economies of scale but in contrast to Bitran and Dunlop finds product-specific economies.

VI. CONCLUSIONS

This paper has reviewed facility survey approaches. From the foregoing discussion it is clear that “facility surveys” have varied extensively—in terms of motivation, type of data collected, methods of data collection, and use of the data. Some surveys have focused specifically on the health care provider, which receives resources, delivers services to the community, interacts with other providers, and reports on activities and outputs. In other
surveys, the primary object of analysis has been the household, which is the ultimate beneficiary of the services, who may make payments to the provider, and who may have a role in holding the provider accountable. Finally, some surveys have sought to shed light on the administrative and logistical system “upstream” from the service delivery unit. This system is responsible for channeling resources to the service delivery unit, and provides the institutional framework for regulation, supervision, provider payment, etc.

What are the lessons that emerge from the experiences to date, and what should be the empirical agenda for the future? One point to emerge is that since many of the issues of interest concern the interaction between different units in the system, it is necessary to look beyond a single part of the system. Indeed, this is precisely the rationale behind the first community and facility surveys implemented in conjunction with the DHS and LSMS surveys. However, while these surveys have taught us a lot about how health behaviors and outcomes relate to service availability and provider characteristics, many questions remain unanswered. This is true not only for the relationship between households, the community, and the provider, but also in respect of the strategic interaction between different providers, and between the administrative and logistical system and providers.

Another point to emerge is that there are difficult methodological issues to resolve concerning the measurement of quality and costs. For example, the paper has noted discussed the disparate approaches to measuring health care quality. Is there an emerging best practice? Will it be possible to design simple ways of collecting meaningful and consistent data on quality? Is it even reasonable to view health care quality as a facility characteristic? Actual client experience may be endogenous to client characteristics and previous client behavior. For example, both structural (e.g. attending staff, access to equipment and drugs) and process (e.g. staff attitudes, time spent with the patient) characteristics of care may depend on client characteristics such as income, social standing, education, assertiveness of the client, or on how much was paid. This would mean that different individuals do not only perceive their visit to the health care provider differently, but also receive different treatment. Progress in the measurement of quality is not only important for the purpose of monitoring, but also for increasing the value of research on health outcome and health related behavior. To date, the disparate approaches to modeling health care quality, and the sometimes inconsistent findings, largely reflect differences in data availability and quality. Similar methodological concerns exist in respect of the measurement of costs and outputs.

Surveys have been better at documenting variation in facility performance—both in terms of quality and efficiency—than in explaining differences across facilities and countries. We still do not know enough about the possible merits of private and NGO providers, or about whether general prescriptions can be made in respect of the institutional context in which public providers operate.

One of the challenges in drawing generalizable conclusions from survey work is to improve consistency across surveys – both over time and across countries. This would require agreement about some core methodological issues, possibly in the form of some core facility modules that will permit international analyses of health system performance.

Finally, it is probably fair to say that some health facility data are under-used. How do we ensure that survey data are thoroughly analysed, and that findings feed into policy making?
APPENDIX: A REVIEW OF HEALTH FACILITY SURVEYS

This appendix provides an overview table of both institutionalized health facility surveys, as well as surveys focusing on costs and efficiency that have been carried out on an ad-hoc basis. In addition to the tables, the appendix offers brief descriptions of the respective surveys. This review does not purport to be comprehensive, but seeks to cover important and accessible facility survey initiatives in the health sector. Where appropriate, references to background reading and research output based on survey data are provided.

Demographic and Health Surveys (DHS and DHS+)

The Demographic and Health Surveys (DHS) have been an important source of individual and household level health data since 1984. The design of the DHS drew on the experiences of the World Fertility Surveys43 (WFS) and the Contraceptive Prevalence Surveys (CPS), but included an expanded set of indicators in the areas of population, health, and nutrition.44

Similarly to the some of the World Fertility Surveys, many of the DHSs included tools to collect community level data. These Service Availability Modules (SAM) were not facility surveys per se, but rather collected information from “community informants”. They were aimed at collecting “objective” information on the facilities and services available to women in the community, focusing, in particular, on family planning services.

More recently, detailed facility surveys—Service Provision Assessments (SPAs)—have been collected within the scope of DHS+ activities. The objective of the SPA is to provide information about the characteristics of health services including their quality, infrastructure, utilization, and availability. The surveys can be national or subnational and can be implemented as a stand-alone assessment of services or can be linked to household survey data. Also, the SPAs are not necessarily limited to government facilities, but may be used to compare services of various providers, including private for-profit, NGO, and community health workers. Examples of SPAs include Guatemala (1997), Kenya (1999), Bangladesh (1999), and Mexico (2000).

Examples of SPAs and Findings: Kenya and Tanzania

The 1999 Kenya SPA (KSPA), focused on the availability and quality of services in the areas of family planning, sexually transmitted infections, or maternal and child health. The data were collected through community interviews, visits to health facilities, and observations of consultations. The facility data complements household data from a DHS implemented in 1998.

---

43 The WFSs were a collection of internationally comparable surveys of human fertility conducted in 41 developing countries in the late seventies and early eighties. The project was conducted by the International Statistical Institute (ISI), with funding from USAID and UNFPA.

44 In 1997 DHS changed name to DHS+ to reflect the integration of DHS activities under the MEASURE programme. Under this mandate, DHS+ is charged with collecting and analysing demographic and health data for regional and national family planning and health programs.
The survey found that although contraceptive services, drugs and materials are widely available, compliance with infection control procedures, including sterilisation of equipment and sterilisation of needles was often poor. Also, consultation observations revealed that the diagnostic process was often unsatisfactory, clients were frequently provided with incomplete or insufficient information to chose between interventions, and to ensure compliance with treatment regimes. Other findings include poor preparedness to deal with STI sufferers due to drug stockouts; lack of equipment to deal with obstetric complications; and poor prescription practices (Macro International 2000).

The Tanzania Reproductive and Child Health Facility Survey was carried out with assistance from MEASURE DHS+. The objective was to provide information on the availability and use of selected reproductive and child health services in Tanzania. The survey included government, NGO and for-profit providers, and covered hospitals, health centers, and dispensaries. Questionnaires were administered to collect information on the community, facility, facility inventory, service provider (staff), pharmacy inventory, and district health management team. The sample was designed to capture the “market of facility services”. This was done by sampling all facilities within a concentric ring of the enumeration area, rather than the nearest facility within 30 km of the EA (as in previous surveys). Summary of findings... (National Bureau of Statistics (Tanzania) and MEASURE Evaluation 2000).

Research and Publications


The Living Standards Measurement Study (LSMS)

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by government statistical offices in developing countries. LSMS surveys are multitopic surveys, designed to permit four types of analysis: (i) simple descriptive statistics on living standards; (ii) monitoring of poverty and living standards over time; (iii) describing the incidence and coverage of government programmes; and (iv) measuring the impact of policies and programmes on household behaviour and welfare (Grosh and Glewwe 2000). The first surveys were implemented in Côte d'Ivoire and Peru. Other early surveys followed a similar format, although considerable variation has been introduced over time.
The household questionnaire forms the heart of LSMS survey. It typically includes a health module that provides information on (i) health related behaviour; (ii) the utilisation of health services; (ii) health expenditures; (iv) insurance status; (v) access to health services. The level of detail of the health section has however varied across surveys. Complementary data are typically collected through community and price questionnaires. In fact, over half of the LSMS surveys conducted before 1997 included community and price questionnaires. Community questionnaires are administered separately to community “informants”, and collect information on infrastructure, employment opportunities, availability of credit, and public services, including schools and health facilities.45

In some LSMS surveys, detailed service provider (health facility or school) data have been collected. In the case of the health sector, facility surveys were implemented in Côte d’Ivoire (1987), Jamaica (1989), and Viet Nam (1998).46 The facility surveys have been included to provide complementary data primarily on prices of health care and medicines, and health care quality. Health facilities have rarely been the object of analysis in research based on LSMS surveys. Quality data may include both structure (staffing, equipment, drugs, etc.) and process (diagnosis, treatment, attentiveness and staff attitude) dimensions of health care quality. The facility surveys have in some cases included private health care providers (Jamaica).

Experiences with LSMS Health Facility Surveys

Jamaica. A survey of health care facilities was carried out in September 1990. The data were meant to complement the expanded health module of the 1989 LSMS survey. All public health facilities and a sample of private providers were surveyed. Data collection was based on four separate health services questionnaires: public primary, private primary, public secondary/tertiary and private secondary/tertiary. At primary level, questions related to: catchment area, facility characteristics, patient services, immunization offered, personnel, beds, transportation, drug supply and equipment, family planning services, and maternal health services.47 There were slight differences in the questionnaires administered to public and private facilities.48 The survey instruments for secondary/tertiary level included more detailed questions on facility characteristics, personnel, and equipment. In the household survey, detailed data on illness episodes and care seeking patterns, including the name of facilities visited, were collected (The World Bank 2001).

Cote d’Ivoire. The survey instruments for Cote d’Ivoire preceded the Jamaica survey, and are much more limited. The facility questionnaire includes many of the same sections as the Jamaica survey, but each section is less comprehensive. In addition, no specific data on family

---

45 In the area of health, surveys have covered question concerning the type of health care providers that are available in the community, costs of treatments and medicines, travel times to the respective providers, means and costs of transport, and public health services and programme, including immunisation programmes and information campaigns. In some countries, community level data have been sufficiently detailed to permit analysis of the relationship between health care infrastructure and health seeking behaviour and outcomes (e.g. LSMS surveys in Ghana (1987/88, 1991/92, 1998/99) Guyana (1992/93), Pakistan (1991), and Tanzania (1991, 1993)).

46 School surveys were included in Jamaica (1990) and Viet Nam (1997/98).

47 In respect of maternity care, the questionnaire sought to assess the process of care. This was done by reviewing a range of services/activities with the appropriate staff member, and asking him/her whether the respective service was included in a standard pre-natal visit and for what group of women.

48 For private facilities, questions concerning laboratory services were more detailed. Also, questions concerning staff wages, patient flows, revenues, and expenditures were included.
planning and maternity care services were collected, and no attempt was made to capture process dimensions of health care quality.

**Viet Nam.** The 1998 Viet Nam LSMS survey included school and health facility surveys. The health facility survey was limited in scope and detail relative to both the Jamaica and Cote d'Ivoire surveys. It collected information on distance to different enumeration areas, staffing and areas of operation, number of beds, service range, equipment and drug availability, and cost of services and drugs.

**Research and Publications**


**Situation Analysis**

Situation analysis (SA) was introduced as a tool for programme evaluation by the Population Council in 1989. In general, the SA can be described as “a comprehensive and standardized approach for systematically assessing both the readiness of family planning/reproductive health programs to delivery services and the quality of care received by clients.” (Miller, et al. 1998). The development of the tool was stimulated by indications from DHS surveys that service delivery weaknesses were important in explaining low contraceptive prevalence rates in many countries.

The first SA was carried out in Kenya. Subsequently, SAs have been carried out extensively, including in 11 African countries. The SAs are based on a representative sample of service delivery units within a geographic area of interest. The approach includes structured interviews with managers and facility staff; inventory review and direct observation of clinic facilities and the availability of equipment and consumables; review of service statistics for 12 months; nonparticipant direct observation of family planning client-provider interaction; and interviews with clients of both family planning and MCH services. The general approach has been modified in some cases to address a broader set of concerns.

Situation analyses have provided clear evidence of the poor state of service delivery in many countries. Documented problems include poor contraceptive stocks, lack of basic infrastructure and equipment, poor adherence to diagnostic and treatment protocols. Through the implementation of follow-up studies, it has been possible to measure changes over time, and to assess the impact of policies aimed at improving service delivery. Data have been used, inter
alia, in the design of family planning programs, training initiatives, and in the formulation of sector strategies.

Background publications


RAND Surveys

RAND has supported the design and implementation of Family Life Surveys (FLS) in developing countries since the 1970s. Currently available country surveys include Malaysia (1976-77, 1988-89), Indonesia (1993, 1997, 1998, 2000), Guatemala (1995), and Bangladesh (1996). Here, the Indonesia and Bangladesh surveys are discussed in further detail.

The Indonesia Family Life Survey

The Indonesia Family Life Survey (IFLS) is an ongoing, multi-topic longitudinal survey. It aims to provide data for the measurement and analysis of a range of individual- and household-level behaviours and outcomes. The survey has collected data at individual and household level, including indicators of economic well-being, education, migration, labour market outcomes, fertility and contraceptive use, health status, use of health care and health insurance, intra-household relationships, and participation in community activities. In addition, community level data are collected. These include detailed surveys of service providers (schools and health care providers) in the selected communities. The first wave of the survey (IFSL1) was conducted in 1993/94, covering approximately 7000 households. The IFLS2 and IFLS2+ were conducted in 1997 and 1998, and a further wave (IFLS3) was planned for 2000.49 Re-interview rates of over 90 percent were achieved in the second and later waves of the survey.

In respect of the health facility survey, visits to local health care providers were carried out, and staff representatives were interviewed about the staffing, operation, and usage of their facilities. The surveys covered (i) government health centres and sub-centres; (ii) private clinics and doctors, midwives, nurses, and paramedics; and, (iv) community health posts. For each community, up to 3 government health centres, 6 private clinics, doctors, etc., and 2 community health posts were surveyed. Health care providers were selected on the basis of information provided by household respondents about where they normally seek care.

Different questionnaires were used for the respective providers to reflect differences in organisation and scope of services. In general, the questionnaires collected data on availability and prices of services, lab tests, and drugs; availability of equipment and supplies; direct

49 The IFLS2+ covered a sub-sample (25 percent) of households and was designed to assess the effects of Indonesia’s economic crisis.
observations about the facility’s cleanliness and other features that might influence its attractiveness to patients. In addition, five hypothetical patient scenarios or vignettes were presented to the relevant health worker to assess the respondents’ knowledge of process in patient care.\textsuperscript{50}

\textit{Matlab Health and Socio-Economic Survey (MHSS)}

The MHSS was implemented in 1996, in Matlab, a rural region in Bangladesh. The general focus of the survey was on issues relating to health and well-being for rural adults and the elderly, including the effects on health status and health care utilisation of socio-economic characteristics; health status, social and kin network characteristics and resource flows; community services and infrastructure. The survey collected(i) individual- and household-level data, a specialised outmigrant survey (sample of individuals who had left the households of the primary sample since 1982), and a community provider survey.

The provider survey covered seven types of health care providers, from government health complexes, family welfare centres, traditional practitioners, and community health workers. Separate instruments were used for each type of provider, although there were similarities in content. Clinical vignettes were administered to all providers.

\textit{Research and Publications}


\textit{MEASURE Evaluation}

MEASURE Evaluation is a project under the MEASURE Program.\textsuperscript{51} The main purpose of the project is to develop and apply methods for monitoring and evaluation in the areas of family planning; maternal health; sexually transmitted diseases, especially HIV/AIDS; nutrition; and infectious disease. MEASURE \textit{Evaluation} works in collaboration with programs in

\textsuperscript{50} The vignettes covered: provision of IUDs, provision of oral contraceptives, prenatal care, treating a child with vomiting and diarrhoea, and treating and adult with a respiratory illness.

\textsuperscript{51} The MEASURE (Monitoring and Evaluation to Assess and Use Results) Program is funded by USAID since 1997. MEASURE includes five projects, which offer technical services in data collection, analysis, dissemination, and use. Details can be found on http://www.measureprogram.org/.
developing countries, USAID and other international donor agencies to (i) improve performance monitoring systems for tracking results; (ii) identify appropriate indicators, test their measurement, and establish needed data systems; and, (iii) evaluate interventions and their cost-effectiveness. As a component of this general work programme, MEASURE Evaluation project has provided technical support in the design, implementation and analysis of a range of health facility surveys. Examples include Uganda and Paraguay.

The Uganda DISH survey

The DISH project, funded by USAID, is one of the largest reproductive health programmes in Uganda. It operates in 12 of the country’s 45 districts, and covers approximately 30 percent of the population. Activities under the DISH project include training, support of community reproductive health workers, capacity building in health management information systems, capacity building in financial management, and IEC activities. As a component of ongoing monitoring and evaluation work, a series of surveys have been implemented to measure changes in reproductive, maternal and child health knowledge and behaviour. The DISH surveys comprise interim surveys between the 1995 and 2000 Uganda DHSs. Two survey rounds have been carried out, in 1997 and 1999. Each survey collects both household and facility data. The household component is based on reduced DHS-style questionnaires. The 1999 survey collected information from 1766 women, 1057 men and 478 facilities (including health units, drug stores and pharmacies), and information is reported on IEC, family planning, maternal health, child health and nutrition, STDs and HIV/AIDS, and health facilities (Kantende, et al. 1999). At facility level, survey instruments included facility interview questionnaire and a facility inventory questionnaire. In addition, separate inventory questionnaires were designed for pharmacies and drug shops. Through these instruments, information is collected on general facility characteristics (including services offered), DISH training, availability of supplies and IEC material.


Health Care Decentralization in Paraguay

The objectives of the study was to identify and quantify the effect of transferring management control for basic health services provision from central to municipal government in Paraguay on costs, efficiency, basic health service quality, and patterns of health service use and equity. The main units of analysis were public health facilities, their clients, municipality populations. The study uses pre- and post-decentralisation design, with a control group, and focused on family planning, maternal health, and infant and child health. The first round in the survey was carried out in 1998. The survey covered 124 public and 19 private facilities and 1261 exit interviews. In addition, with a view to obtain information on socio-economic characteristics, health outcomes, and health care seeking behaviour, a general household survey was also

---

52 This has included the establishment of a Facility Survey Working Group to determine the use and procedures for critical aspects of data collection, monitoring, evaluation and dissemination.
implemented. This was carried out as a component of the 1998 maternal and infant health survey.

At facility level, three types of questionnaires were administered: (i) short inventory questionnaire; (ii) interview with facility director or other staff member, and (iii) direct observation of availability of supplies, equipment, infrastructure, etc. In addition, a self-administered time sheet was completed by staff members. The exit poll collected information on socio-economic characteristics, reason for visit, services received, travel time, payments, perceptions, etc. More detailed data were collected from randomly selected households, covering both socioeconomic and demographic characteristics, health, and health care seeking behaviour. A second round of data collection is planned, but has been delayed due to slower than expected implementation of decentralisation.


**WHO Multi-Country Evaluation of IMCI**

There is a considerable history of facility surveys in the WHO. Originally, health facility instruments focused on issues relating to child mortality and morbidity, and survey instruments were often disease specific (i.e. designed to evaluate specific programmes or projects), e.g. ARI (acute respiratory infection), diarrhoea, etc. The principal interest lay in assessing quality (typically process focused assessment, though also measurement of input/structure – e.g. drug availability, staffing, etc.), and effect of quality on household behaviour and health outcomes. Hence facility surveys were typically accompanied by exit polls and/or household surveys (coverage surveys).

Recently, integrated surveys have been designed and implemented. An important element of this work is carried out in the context of multi-country evaluations (MCE) of the Integrated Management of Childhood Illnesses (IMCI), and is implemented by different institutions (John Hopkins, CDC, etc.), under coordination by WHO. Integrated instruments for costs and quality have been developed and implemented (or are being implemented) in Bangladesh, Tanzania, and Uganda. The purpose of the MCEs is to (i) document the effects of IMCI interventions on health workers performance, health systems and family behaviours; (ii) determine whether and to what extent, the IMCI strategy as a whole has a measurable impact on health outcomes (reducing under five morbidity and mortality); (iii) describe the cost of IMCI implementation at national, district and health facility levels; (iv) increase the sustainability of IMCI and other child health strategies by providing a basis for the improvement of implementation; and (v) support planning and advocacy for childhood interventions by ministries of health in developing countries and national and international partners in development. Worldwide there are 30 countries at different stages of implementation of IMCI among which Uganda, Peru, Bangladesh and Tanzania will participate in the MCE.

53 The Integrated Management of Childhood Illnesses (IMCI) Strategy was developed by WHO and UNICEF to address five leading causes of childhood mortality, namely: malaria, pneumonia, diarrhoea, measles and malnutrition. The three main components addressed by the strategy are: improved case management, improved health systems and improved family and community practices.
Public Expenditure Tracking Surveys

In recent years, the public expenditure tracking surveys (PETS) have been implemented in a range of countries in response to perceived public expenditure management and service delivery problems. These surveys have varied considerably both in terms of scope and focus, to the extent that the use of a single term in reference to the variegated survey activities may be misleading. In general, it is possible to distinguish two types of surveys under the umbrella of public expenditure tracking: tracing studies and quantitative service delivery surveys (QSDSs). Tracing studies have been used to identify the location and extent of impediments in financial flows to sub-national levels of administration or service delivery units. Government resources, which are typically earmarked for particular uses in the budget (votes, line items), flow, upon release, within a pre-defined legal, regulatory, and institutional framework, passing through the layers of the government and via banking system down to districts or facilities. Tracing studies track the flow of resources through these institutional strata in order to determine how much of the original resource reaches each new level, and how long the resource takes to get there. In contrast, QSDSs are more akin to facility surveys, although they have often tried to capture different aspects of the relationship between service delivery units and the administrative and logistical system “upstream”. The focus on efficiency and quality in service delivery is in recognition of the fact that even if budget outturns reflect allocations, ultimate impact in terms of development objectives also requires that financial resources are efficiently transformed into appropriate public services.

As noted, PETS can in part be seen as a response to the lack of information about the systems through which budget resources are transformed into services. Indeed, the first PETS, implemented in Uganda in 1996, asked why it was that service delivery outputs in the health and education sectors had failed to respond to considerable increases in public spending. The survey was carried out in 19 of Uganda’s 39 districts. It focused on a sample of 250 government schools and 100 public health clinics, and spending data for the period 1991-95 were collected (Ablo and Reinikka-Soininen 1998; Reinikka 1999). For the education sector, the survey compared central government outlays for non-wage spending with the corresponding school income. Specifically, it focused on the capitation grant that is paid out per student enrolled as a matching government contribution against the mandated tuition fees paid by parents. The results demonstrate how weak governance and lack of accountability can hamper the process of budget implementation, and that de facto public spending may be determined by the bargaining power of individual facilities (Reinikka and Svensson 2000). The survey also demonstrated the inaccuracy of administrative data on school outputs. In the health sector records were poor or non-existing, and little quantitative data were collected. On the basis of interviews, it was however estimated that on average, almost 70 percent of medical supplies and drugs were expropriated by staff.

A follow up study was carried out in Uganda in 1999 (Ministry of Education and Sports (Republic of Uganda) 2000). This was a tracing study in the education sector, where, for a sample of 427 schools and 11 District Headquarters, four monthly releases in 1998 and 1999 were traced through the payment system. The objective of the survey was to track flow of funds, verify compliance regulations, and assess the scope and sources of delays in budget execution. The survey found considerable improvements in the amount of funding that schools received.
compared to the 1996 study. In part, these improvements can be attributed to measures to increase transparency and accountability – e.g. through publication of school and health facility in local papers and on local notice boards (Reinikka 1999). However, the study also documents considerable delays in the transfer of funds. The study found that these delays do not occur uniformly throughout the system. Central ministries, the Central Bank, and the headquarters of the government owned Uganda Commercial Bank (UCB) all transferred resources in full and without delays. At lower levels, however, transfers were held up by local branches of UCB and by districts administrators.

The problems highlighted in the Uganda studies are far from new. However, these studies provided important quantitative evidence on the scope and seriousness of problems in budget execution, and demonstrated how a new tool can be applied to analyse these issues. Other surveys have been carried out in Ghana, Honduras, Tanzania, and Uganda. These surveys have focused on leakage, “job capture” and other means of extracting private benefits from public positions may be systematically, the implicit allocation of resources at different levels of administration, and the costs and efficiency of health care delivery. For example, the study of public employment in the Honduran health sector showed that the best qualified health sector staff engaged in more shirking, ‘ghost working’, and job capture than lower-ranked employees (World Bank 2001b). In addition, a number of surveys are currently being planned or implemented. Many of these surveys are variants of school of facility surveys, and in some cases efforts are made to link the surveys with household data collection exercises.

References


<table>
<thead>
<tr>
<th>Organisation</th>
<th>Population Council</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situation analysis</td>
<td>Countries</td>
</tr>
<tr>
<td></td>
<td>Indonesia, Matlab, Bangladesh, Cote d'Ivoire, Jamaica, Viet Nam, Paraguay, Uganda, DISH, Kenya, Tanzania, various</td>
</tr>
<tr>
<td>Broad objectives</td>
<td>Multi-topic survey to measure and analyse different individual- and household-level behaviours and outcomes.</td>
</tr>
<tr>
<td></td>
<td>Measurement and analysis of determinants of health status and human capital investments for rural adults and elderly.</td>
</tr>
<tr>
<td></td>
<td>Analyse effect of decentralisation on costs, efficiency, basic quality, and pattern of health service use.</td>
</tr>
<tr>
<td></td>
<td>M&amp;E of DISH project Assessing readiness of family planning / reproductive health programmes to deliver services and the quality of care received by clients.</td>
</tr>
<tr>
<td>Sample size (facility)</td>
<td>approx. 4000, 254, 88, 478, 388, 445 varies (100-300)</td>
</tr>
<tr>
<td>Level of care</td>
<td>Outpatient care providers at community level</td>
</tr>
<tr>
<td></td>
<td>Health complex, welfare centres and individual practitioners</td>
</tr>
<tr>
<td></td>
<td>Primary to tertiary district hospitals, health centres, and health posts</td>
</tr>
<tr>
<td></td>
<td>Health units (hosp., health centres, dispensary), drug stores, pharmacies</td>
</tr>
<tr>
<td></td>
<td>All levels (dispensaries to hospitals)</td>
</tr>
<tr>
<td></td>
<td>All levels (dispensaries to hospitals)</td>
</tr>
<tr>
<td></td>
<td>Varies</td>
</tr>
<tr>
<td>Agency types</td>
<td>Govt. (health centers, subcenters &amp; community health posts); private (clinics &amp; practitioners).</td>
</tr>
<tr>
<td></td>
<td>Govt., private, traditional practitioners</td>
</tr>
<tr>
<td></td>
<td>Public and private</td>
</tr>
<tr>
<td></td>
<td>Public, NGO, for-profit</td>
</tr>
<tr>
<td></td>
<td>Public, mission, private</td>
</tr>
<tr>
<td>Disease / intervention focus</td>
<td>Private clinics and doctors, midwives, nurses, and paramedics (kliniks, praktikum, perawat, bidan, paramedis, general)</td>
</tr>
<tr>
<td></td>
<td>Special attention to family planning and maternal health services</td>
</tr>
<tr>
<td></td>
<td>Family planning, maternal and child health, reproductive, maternal and child health</td>
</tr>
<tr>
<td></td>
<td>Child health, family planning, sexually transmitted infections &amp; HIV/AIDS maternity care</td>
</tr>
<tr>
<td></td>
<td>Family planning and reproductive health</td>
</tr>
<tr>
<td>Data</td>
<td>Household - health data xx xx xxx xx Complements 1999, Complement 1999</td>
</tr>
<tr>
<td></td>
<td>Household - complementary data xx xx xx xx (x) (x)</td>
</tr>
<tr>
<td></td>
<td>Client perception / satisfaction xxxx</td>
</tr>
<tr>
<td></td>
<td>Facility infrastr. and basic char. xx xx xx xx (x) xx</td>
</tr>
<tr>
<td></td>
<td>Management / supervision (x)</td>
</tr>
<tr>
<td></td>
<td>Service range xx xx xxx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Facility procedures (e.g. hygiene and sterilisation) (x) (x) (x) (x) xx</td>
</tr>
<tr>
<td></td>
<td>Process quality xx xx xxxx</td>
</tr>
<tr>
<td></td>
<td>Charges (consultation / medicines) xx xx xx xx (x) xx</td>
</tr>
<tr>
<td></td>
<td>Activities and outputs xx xx (recall) (x) (x) xx</td>
</tr>
<tr>
<td></td>
<td>Input/resource use and costs</td>
</tr>
<tr>
<td></td>
<td>Equipment &amp; instruments xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Drug availability xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Vaccine availability xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Staff numbers and categories xx xx (x) (x) xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Salaries xx (x) sample of staff</td>
</tr>
<tr>
<td></td>
<td>Education and training xx xx</td>
</tr>
<tr>
<td>Methods</td>
<td>Household questionnaire xx xx xxx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Separate survey</td>
</tr>
<tr>
<td></td>
<td>Exit poll xxx</td>
</tr>
<tr>
<td></td>
<td>Community questionnaire xx xx</td>
</tr>
<tr>
<td></td>
<td>Facility interview xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Facility inventory xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Direct obs. / visual inspection xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Record review xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Staff interviews vignettes xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>District management questionnaire xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Consultation observation xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Clinical vignettes xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Time and motion / staff time sheets (x) (x) xx xx xx xx xx xx xx xx xx xx</td>
</tr>
<tr>
<td></td>
<td>Monitoring of living standards; measurement and analysis of impact of policies and programmes on household behaviour and welfare.</td>
</tr>
</tbody>
</table>