

EVIDENCE-BASED PLANNING:
THE PHILOSOPHY AND METHODS OF
SENTINEL COMMUNITY SURVEILLANCE

Neil Andersson

CIETinternational
and
The Economic Development Institute of the World Bank
1996

CONTENTS

Foreword	5
Preface	7
Part 1. Evidence-based Planning: Measuring Impact, Coverage and Costs	9
Results-oriented Management	10
Routine collection	
The blunderbuss problem	
Improved planning	
Stimulating wider participation	
A Common Language	12
Impact	
Planning for results	
Coverage	
Linking coverage and impact: a management tool	
Behind the coverage indicators: quality	
Costs	
Sustainability	16
Indicators: words in a common language	
Summaries that link impact, coverage and costs	
Sentinel community surveillance: concentration of measurement resources	
Dynamic registration	
No Blueprint Exists	20
The intersectoral nexus and community voice	
Part 2. Benchmarking and Impact Assessment	23
Multiple-indicator Surveys	23
An alternative for some countries	
Purposive and random samples	
Sample size and costs of fieldwork	
Safeguards for population inference	
Behind the Indicators: Actionable Evidence	27
Statistical handling of SCS data	
Capacities and the learning curve	
Interpretation of status indicators	
Sub-national schemes	
Representative data and community participation	
Evidence-based Planning <i>as</i> Development	
Annex 1: Representative samples	33
Annex 2: Factors excluded by stratification	34
Annex 3: Feasible impact evaluation	34
Annex 4: Suggestions for interpretation and presentation of indicators	35

Part 3. Evidence-based Planning and Local Governance	37
Methods	38
Analysis: The Platform for Communication and Mobilisation	41
Project Impact	42
Improvement of knowledge	
Community voice	
Knowledge transfer and a growing evaluation culture	
Consciousness about costs	
Promoters and local health care capacity	
Micro-projects	47
Conclusion and Challenges	47
Annex 1: Description of the municipalities	50
Part 4. Meso-analysis: Quantifying Qualitative Data from Communities and Services	51
Of Macro, Micro and Meso	51
Complementing Impact Assessment with Qualitative Data	52
Cross-design Synthesis	53
Statistical Management	54
Mexico: linking household and service cost data	
Nicaragua: determinants of school dropout	
Liberia: vaccination efficacy and grouping by sites	
Site Selection and Links with Other Surveys	57
Commensurable Methods and Meta-analysis	58
Annex 1: Sentinel community surveillance: a typical cycle	59
Endnotes	61

FOREWORD

A core objective of the Governance and Public Sector Management (GPSM) programme at the World Bank is to facilitate partnerships to expand government's capacity to improve service delivery to the public. The model developed by the GPSM team hinges on collaboration with client countries, national and international NGOs, and donors. Civil society is a crucial partner, its inclusion in the service delivery dialogue having a central function in accountability.

The GPSM programme reaches to emphasise (i) *changing mindsets*, or the way problems are seen and goals defined; and (ii) *developing leadership*, the capacity to motivate people to achieve a common goal. The GPSM team recognises that local, regional and central governments, the private sector, NGOs and citizen organisations have a key role in this.

Improving service delivery means increasing cost effectiveness, coverage and impact of services. When defining service quality, the client's point of view is fundamental, reinforcing an orientation to results where a public action is justified if it produces a publicly valued impact. This mindset for policy-makers and public managers is the enabling environment for more efficient, results-oriented public sector policy making and management.

An emphasis on integrity strengthens this mindset. Services should be designed and delivered in a transparent way, with baselines of service delivery openly established,

targets feasible but not unreasonably low and progress against the baseline fairly measured. This supports the goal of efficiency and results-orientation in the public sector while encouraging governments to be accountable.

Enhancing the sustainability of effective service delivery requires more attention to regular and meaningful *evaluation, monitoring and adjustment*, the subject of this paper. In partnership with experts in client countries, the GPSM team and CIET are developing institution- and community-based information gathering and evaluation systems, such as Service Delivery Surveys and programmes that use both quantitative and qualitative data to assist in improving service delivery. Guided by Service Delivery Survey results, a government can introduce and promote evidence-based planning and management, adjusting and refining the service delivery framework through policy reformulation, institutional development and budget reallocation.

The methods presented in this paper offer governments a service delivery baseline. Perhaps a more important use is as a means to increase the involvement of civil society in the process.

As stakeholders interact, improving service delivery requires devolution of responsibility, competitive supply arrangements, or private sector provision. If this leads to redefinition of the role of the state, it may be necessary for the government to manage contracts as effectively as a

private sector firm. Moreover, the government is seen to have responsibility for the impact of services it no longer provides, at least for the short- to medium-term. Clearly, the government must be able to evaluate, to monitor and to adjust the coverage and impact of services, even if it no longer directly delivers, but manages and regulates.

Improving public sector management is an ongoing process, one in which the elements leading to improved, country-specific public sector management are uncovered, developed and sustained. Fundamental to this learning process are several methods: piloting of new approaches, identification of good practise and dissemination of lessons learned and other successes achieved. This paper provides the focus for one of GPSM*s pilot programmes, the Service Delivery Survey programme, that attempts to adapt an existing methodology (Sentinel Community Surveillance) to the pursuit of the GPSM model*s goals of creating a results orientation, enhancing transparency and accountability and ensuring sustainability through capacity building at all levels while leading to improved service delivery.

Service Delivery Surveys were designed by Dr Neil Andersson to meet the needs of clients in several regions; in an iterative process, lessons learned in one pilot have been used to adjust the methodology in another.

In addition to working with the Service Delivery Survey tools, the GPSM team responds to client demands in developing central government leadership networks, facilitating state and municipal level workshops and sponsoring initiatives through other stakeholder groups within civil society. The result of this series of interventions is the development of a service delivery function that brings the benefits closer to citizens, the customers of the public sector.

Petter Langseth
Washington, DC, May 1996

PREFACE

This collection of essays provides the methodological and philosophical backdrop to the community-based research and mobilisation methods known as Sentinel Community Surveillance (SCS), known in the World Bank as Service Delivery Surveys (SDS).

Many original ideas behind the approach to evidence-based planning were developed by the author while undertaking a comparative study of information "systems" of eight countries in 1982 for the Strengthening Health Services Division at the World Health Organisation (WHO), Geneva. The study focused on how these systems used evidence relating to effective and affordable coverage.

An early lesson of the WHO study was that these are not really systems at all. They are more often disjointed activities that make available certain data that theoretically could be used in planning. In reality, they seldom are. It was also found that not all evidence needed for planning can be gleaned from routine sources, for example, the reason people might choose not to take advantage of an available service suggesting that existing sources of information should be complemented by survey data of some kind.

These ideas relating to evidence-based planning were first implemented in UNICEF-supported projects in Honduras and Nicaragua. The Central American projects evolved a pragmatic cross-design synthesis of methods, including analysis of existing data, household surveys, rapid anthropological procedures and institutional reviews. The innovation of SCS is its cross-design of methods. With the household survey serving as the methodological cornerstone, SCS allows fine-tuning of other quantitative and qualitative data gathering processes to produce reliable, actionable data.

While developed in the health sector, SCS has been refined and used during the last decade in a wide range of development concerns: education, food security, environment, land mines, economic structural adjustment, urban transport, justice and corruption in the police. This flexibility across sectors, considering the very real differences across sectors and countries, is another advantage of SCS.

The first essay was originally published by UNICEF's Central American Area Office in 1985. With the challenge of selecting *where* to measure (the sentinel sites), it is easy to lose sight of the questions really driving the scheme: *what* to measure impact, coverage and costs -- and, more important, *how data are used*. Specifically, we are concerned with how the evidence supports decisions in the context of results-oriented management, strengthening processes that will result in improved outcomes and impact.

The second essay addresses benchmarking and the technical concerns of sample survey methods for monitoring target-driven development. The essay *Benchmarking and Impact Assessment* attempts to marry the quite different objectives of advocacy-driven ("top-down") development with practices that might

promote customer-focused development. In advocacy-driven development, national level indicators feed into the advocacy machinery to produce more resources that then may be channeled into local development. This section outlines some drawbacks in this approach.

Local application of SCS methods and their use in planning are described in the third essay, *Micro-regional Planning*. The essay illustrates the fusion of standard epidemiological techniques with social dynamism, the communication of data and capacity and consciousness building. Based on a project to develop an evidence-based local planning system in Guerrero State, Mexico, with support from the International Development and Research Centre (IDRC), Canada, this essay has strong implications for strengthening governance and public sector management through increasing community participation in planning.

The fourth essay reflects one of the substantive methodological advances in this field, notably the use of quantitative and qualitative evidence at the interface between communities and services. *Meso-analysis* describes the fusion of well-recognised Rapid Assessment Procedures (RAPs) with institutional reviews of local services and modified cluster survey techniques. The innovation here is the application of all these methods, well-developed in themselves, in a panel of coterminous communities, thus permitting linking of methods and follow-up over time.

Neil Andersson
New York, April 1996

EVIDENCE-BASED PLANNING: MEASURING IMPACT, COVERAGE and COSTS SUMMARY

Most countries produce at least some data of direct relevance to planning and management. These are of variable quality and seldom well-integrated into the planning process. For the most part, they are used only to monitor specific activities within health and education services. In many countries, even vital statistics are rarely usable, death registrations in general being worse than birth registrations. Scarcity of resources has resulted in an increasing awareness, on the part of national authorities and international development agencies, of the need for planning to be based on evidence.

Evidence-based planning is guided by four principles:

Evidence contributes to development

This can be done by using the evidence in advocacy and results-oriented management (ROM) and in the dynamic or potentially interactive dimension of measurement. In the process that produces evidence, a balance must be reached between different uses (advocacy, ROM and mobilisation), since the needs of each can at times pull in different directions.

No more and no less evidence should be collected than is required to improve development and certainly no more than can be used rapidly by those who

analyse the data. This implies the need for training to develop capacities not only in data gathering but also in analysis.

A common framework is needed for sectoral and intersectoral measurement

A common language can facilitate communication between levels in a single sector (local, district, regional and national) and tie the measurement process more closely to planning. *Impact* implies changes in status; *coverage* is the proportion of the people who benefit from a particular activity out of all who need it; *costs* include those incurred by providers and by the people themselves.

Evidence-based planning is sustainable

It should not add to the work of already overburdened service workers; it should be tailored to the resource constraints of the country. This means obtaining necessary evidence for planning that is otherwise not available from existing sources in an economic and repeatable way. Selection of indicators is one mechanism to limit the amount of data taken into a planning process. Other options are to limit the geographic area of collection (sentinel sites) or to select specific times during which data are collected (dynamic registration).

No blueprint exists

No simple formula exists that can be applied in every situation. The system should evolve from the needs and the resources of each country. An appropriate place is the analysis of existing data within the framework of effective, affordable coverage; any missing data can be obtained later during a series of data collection cycles.

In initial phases, the statistical rigour of the measurement must be balanced with capacity building. Repeatable cycles of data collection and analysis are preferable to “one-off” exercises for which special skills need to be imported. Nationally established priorities could be worked through in a spirit of audit, starting with those issues upon which action can be taken most easily and with greatest impact, rather than relying on a “blunderbuss” (scattershot) approach that measures too many aspects at the same time.

Acknowledgement

This essay was part of the original proposal of Sentinel Community Surveillance commissioned by UNICEF and published in September 1985. It does not represent an official UNICEF or World Bank position. A second version was published by CIETinternational in March of 1995.

To plan is to choose

Planning involves, among other things, setting priorities. That is the easy part: to set a strategy and to outline tactics within it. Much more difficult is the task of assessing the success of the strategy in terms of its component tactics.

Measurement terminology, even before getting into the practicalities of different methods, can seem muddled. *Monitoring* usually implies a continuous process, rather like a movie camera recording a stream of images; *evaluation* usually implies an element of judgment, a cognitive component. Within these broad definitions, it is probably useful to remember that we are concerned with measurement, determining the dimensions of development and, as far as possible, what we can do to improve it.

The framework outlined here is an attempt to systematise an international pool of experience in a way that can be used, immediately and in the longer term, to develop a durable and useful information system. The four main themes are not intended to be final prescriptions; rather they are categories for organizing thoughts about the practicalities of using community-based evidence in planning.

The first principle for evidence-based planning is that data should contribute to development and not simply seek to reflect the state of development. This implies the collection of no more and no less data than are necessary for the planning process. To the extent possible, the collection of evidence should play a dynamic role in development.

The second principle is the adoption of a common language for

communication between local, district, regional and national levels and among the different sectors. This implies collection of evidence on impact, coverage and costs that easily could be interrelated to initiate, at each level, discussions about the allocation of resources.

The third principle for evidence-based planning is that its establishment and maintenance should be affordable. It should evolve alongside growing national and local capacities.

The fourth principle, derived from the first three, is that there is no blueprint for all countries, or for one country at different times of its history. What is required is a cross-design synthesis of core methods built upon *and building* national capacities.

Most would agree that there is little benefit in collecting information unless doing so contributes to development. Obvious as this point may be, it is almost a defining characteristic of information systems in many developing countries that they enjoy considerable independence from decision-making processes. Measurement of projects supported by external agencies frequently has even less to do with the national planning process.

Routine collection

There are good and familiar arguments in favour of comprehensive routine data collection schemes. First, monitoring is a continuous process and we need to keep an eye on progress in the implementation of programmes and on the changing seasonal demand for services. Secondly, it may be useful to have the information

later, when we do have the capabilities to act upon them.

The concern is how much information is collected without being put to practical use, and how much information is not collected despite being needed.

Another problem is the quality of the data. Perpetual collection without conspicuous use, in the sense of visible change stemming from the data, seems to have a soporific effect on the collectors. It is difficult, day after day, to make ticks on charts with quite the same enthusiasm for accuracy as the day before, particularly when there is no visible link between those ticks and changing health patterns. It is even more difficult when there is a lot of other work to do that may have some visible impact.

The “blunderbuss” problem

When resources are limited, it may be easy to forget that analytical capacities also tend to be lacking. Quite specific capabilities are required to make good use of evidence. The sparseness of these capabilities in many countries has led to an approach that can be compared to a blunderbuss: a barrel tightly packed with assorted items that spread out widely and with a loud report when discharged to hit everything in its path including, if fortune smiles, the subject of original interest. In a blunderbuss approach, evidence is gathered about a wide variety of activities and indicators to be presented in long and comprehensive (and frequently undigested) annual reports.

The blunderbuss approach carries yet another disadvantage: even if the capacities exist to interpret all these data soon after collection, it is never possible

to act upon all the recommendations at once, correcting the shortcomings and the misallocation of resources of previous years. But if all the recommendations cannot be acted upon, nor all (or even most) of the shortcomings made good, nor the misallocation rectified, it makes little sense to use scarce resources perpetually to gather and analyse a huge collection of data every year.

Improved planning

There are several ways that evidence can contribute to development. The first is advocacy, holding up evidence about the situation and what will work to improve it. Another way is through better informed results-oriented planning and management, at the Ministry level, or at regional, district, or local levels.

People at each level of the development process need to know what the problems are and what solutions exist to deal with them, how the solutions work out in practice and what the real costs of each option will be. These decisions frequently are made on the basis of narrow experience, hopeful optimism, or plain arbitrariness. But even when one or a combination of these methods results in a

An administration that has no mechanism for systematic and regular appraisal of its achievements is likely to make less than optimal use of its resources. It will have to face the consequences of this in time.

correct decision, the process is at best centralised and removed from those who have to act on the decisions. This runs directly counter to international experience: unless ordinary people take an

increasingly determining hand in their own development, we cannot expect much improvement.

Beyond the usual plea for an intersectoral approach, desirable as that is, planners need to be able to sort out the different components that contribute to impact. Unless we are able to separate out the effects of the various contributions, planning will be little more than guesswork and the allocation of resources will be little more than a continued act of faith. A development information system should focus on all those activities, including intersectoral ones, that are necessary to improve development.

Stimulating wider participation

Measurement can contribute to development not only through advocacy and better planning and management, but also by what I would call its dynamism, its active dimension. When citizens are involved in measurement activities, their broader participation in the planning process may be sparked. Interaction around development-related data as they are gathered can support local initiatives.

My eye-opener in this respect came from a project I was invited to evaluate in 1977, an experimental information system developed over a ten-year period for a population of approximately 100,000 people. Three full-time data collectors were responsible for certain components of health in 500 households. Each household was visited once a month by the three collectors who enquired about their particular area of responsibility (MCH, first aid, sanitation). Every health event, from births and deaths

to the common cold, including every use of available health services, was recorded on prescribed forms. For each death there was a description and a validation process to confirm the likely cause of death. Several layers of supervisors did random checks on the accuracy of the recording and all reports were relayed promptly to the district hospital where they were aggregated and analysed. This system had been running for ten years but improvements in health status (as measured by reduced infant mortality, child mortality and so on) were no greater than in the rest of the country over the same period. People did not use the new latrines, they were slow in coming forward for immunisations and numerous health education drives seemed to have no impact.

It was difficult to understand how this apparently perfect information system had failed. One day, a supervisor explained that the collectors felt the ticks they made on their charts were the "umbilical cords tying their pockets with the bank". They had no incentive to discuss symptoms with other members of the community, whose health knowledge and health status consequently remained static. Since this potential advantage of a catalytic interaction between data collector and community was lost, there was no difference between the project area and the rest of the country. The problems were roughly the same, although much better documented in the project area and they were tackled in roughly the same way. The issue is not only to collect data, but to stimulate processes of development.

A contrasting lesson came from a project I visited in 1982, serving a population of similar size. Here the management team was aware of the need for regular information on various aspects of health status and resources, but saw this data collection process as a possible starting-point for development actions. It was also seen as a framework for continuing education of the village health workers. The team developed, through trial and error, a series of simple one-page questionnaires that could be completed in ten minutes at a weekly meeting between the illiterate village workers and their supervisor. Each week the questionnaire enquired about the number of births and deaths, including age at death; this was supplemented with a few additional questions that varied weekly about the causes of death or about a disease particularly prevalent at that time of the year.

For example, when village health workers heard of a death from neonatal tetanus, they were taught to ask questions about how the cord had been cut and

What they need is a framework for taking decisions about their circumstances: what type of interventions are needed to bring about a change, and what type of impact they can reasonably expect of the available resources. Not least, they need visible assurance that they do get their share of available resources.

whether the mother had received tetanus antitoxoid during pregnancy. This, in turn, was the format of feeding information back to the supervisor who in turn would pass it on to the management.

By limiting the information collected and discussed in this way to two

of the most important problems each week, the village health workers and later the community from whom they obtained their information could begin to think about what caused and what prevented diseases. This educational interaction provided the starting-point for programmes to combat each of the diseases, with education modules designed by the district management team to coincide with the peak season of the diseases in question. Hence no work was undertaken unnecessarily in recording health events that were not relevant to the action programme.

The second project was very much less expensive than the first if for no other reason than that the existing health workers collected the information: no special collectors were employed. In terms of impact, moreover, it was infinitely more successful. Of course it meant more work for the management team, who had to design the education modules and collate the forms sent for analysis. To offset this additional workload, the data collection forms were designed so that a certain amount of the analysis could be done at the point of collection. Summary Reports are discussed in more detail below.

A COMMON LANGUAGE

If the first principle for national level measurement is to provide reliable answers for planning and management in the quickest and least expensive way, then the second principle must be the adoption of a common framework or language and the systematic sharing of information between levels.

Each level of planning and management and indeed each interest group at any one level, has its own information requirement.

People at the periphery of the national system who are concerned with a small part of it, as in a village or an urban neighbourhood, can assess their health circumstances directly. For them, perception is not a great problem. They know, for example, whether children attend school, or what resources are available on the spot, or the reasons they do not use the locally available facilities.

Regional or national planners, who are more removed from the community, may have difficulty in perceiving what is actually happening in each locality though they might have a good overview. They need, at frequent intervals, a summary of facts about development at the local level. These are the "eyes" without which planning decisions are taken blindly.

There inevitably will be mixed demands of any information system. National and intermediate level planners and managers probably believe their needs are more pressing than those of local level managers or the communities. The conflict cuts across sectors as well. Even in the Ministry of Health, the people responsible for immunisations probably believe their needs should take precedence over programmes against diarrhoea, or education on breast-feeding. Since national planners and managers tend to have louder voices than people at the regional or local level, information systems are frequently geared for information to flow from the periphery to the centre and the competing subsectoral

interests (e.g., immunisations, health education) fracture even this one-way flow of information.

The information system needs to address these challenges in finding a way to sustain integrated measurement. I would suggest three linked concepts as the "language" for communication between levels in a single sector and between sectors:

Impact is the change of status (for example, death, disease, disability, discomfort, loss of income or, on the positive side, literacy, education, food security, employment, well being) that can be attributed to a particular intervention.

Coverage refers to the proportion of people who obtain a particular service (health, education, water) out of all those who need it.

Costs include time, personnel, cash, supplies, transport and all other elements required to achieve or to take advantage of a given service or programme.

Impact

Those concerned with the allocation of resources, at whatever level, need to know how many people will be spared from loss of employment or income, death, disease or disability, or who will benefit from development as the result of a particular activity. This can be called the effect or impact of the activity. Its measurement is the foundation of evidence-based planning. There may be side effects—the unwanted outcomes—that also should be measured.

For planning purposes it is crucial not to confuse effectiveness—the potential impact of an activity—with its actual effect

(impact) in reality. Obviously, a measure should not be introduced unless it carries the potential for impact, but the most pressing recurrent need is for planners to assess actual impact. For example, if we measure the change in the number of deaths from pulmonary complications among children immunised against measles, we may be measuring effectiveness and not effect (impact). The children who were immunised may be in better health (they were not ill on the day of the campaign, or they were well enough to travel to the clinic for their immunisation, or they come from families able to take them to the clinic) than those who did not receive the immunisation.

We cannot assume that the effect of immunisation would be the same among those currently excluded. Planners need to know the effect immunisation has for the whole childhood population, not simply among those who are privileged to receive benefit of the development services. Equity is at stake as much as accuracy.

It is also worth clarifying the distinction between different types of result: output, outcome and impact. In considering the results of education, output might refer to the number of pupils going through a particular education. Outcome might be their performance, for example, on a standardised national examination. Impact would refer to the social contribution they make, through employment, participation in governance and so forth.

The choice of priorities in using resources to achieve a certain impact will depend partly on the frequency we are considering. Maximising impact lies at the

heart of evidence-based planning.

Reduced deaths from a rare cause will be less of a priority than reduced deaths from a common cause. Addressing a common problem in agricultural extension is likely to have a different impact than attention to a rare problem. In order to obtain some idea about frequency in the entire population, evidence is needed about the whole population, not only those who are covered by one or another development activity.

Planning for results

For individuals, whose common interests lead to community participation in planning and implementation, a development effect means a reduction in *risk*: lower risk of dying from a preventable cause, lower risk of dropout from school, or lower risk from crop failure. Each of these things is perceived by individuals as *risk*, so risk is a fruitful focus for discussions. In discussions regarding the allocation of development resources, particularly those at a more local level, the concept of risk probably has more intuitive and personal meaning than do rates or absolute numbers.

High infant mortality rates are important not only as an "indicator" of bad social conditions, but also because they mean a risk for individual babies.

The language of measurement, therefore, not only should reflect the parameter of true interest (effect rather than effectiveness), but also should be accessible conceptually to ordinary people in addressing changes in risk rather than absolute numbers.

Changes in risk (or impact) can be assessed by counting the numbers directly, or by recording reductions in rates of people spared from death, of children spared from dropout, of people suffering assaults on public transport. Direct number counts are appropriate where the number of individuals involved is very small in relation to the population (deaths due to assaults on public transport, to cite one instance, may be in the range of one to ten per 100,000).

The problem is that the size of the population is not always known. In this case, it is possible to use a third measure of impact, proportional rates. Whereas a direct rate is the number of deaths or disease in, say, every 1000 people, a proportional rate is the number of deaths or disease in relation to another cause of death or disease. This type of calculation, called "numerator statistics", makes it possible to estimate the changes in risk or impact attributable to a particular health or development activity without knowing the exact population size.

Coverage

The term coverage refers quite specifically to the proportion of people who benefit from a particular service out of all those who need it. For example, one definition of antenatal care coverage is the percentage of all deliveries attended by trained personnel; coverage with basic education refers to the proportion of all children who attend school.

This is not the same as deliveries or children "programmed" (in the phraseology of some planners) to benefit from these services, a definition of

coverage that is still used in some countries. The difference between the two definitions raises, however unintentionally, a question of equity. In the first definition, need is defined in terms of equity: all those who benefit out of all those in need. The second definition reflects concern only for those in the catchment areas as defined by the planners.

The two definitions of coverage have to do with another divide: between "service-based" (based on number of people using the services) and "population-based" (number of people who need the services). The difference between service and population based calculations of coverage is rather the same as between impact (effect) and effectiveness. While it is useful to know how the activities work out among those who receive them, this is not enough. We need parameters that reflect the equitable use of resources. However, attainment of full coverage should not be an end in itself. It is the means to an end, a means mediated through the impact it can produce. An equitable planning process requires that coverage and impact go hand in hand.

The link between coverage and impact is also the basis for efficient management and supervision.

Linking coverage and impact: a management tool

In many countries, the management system is entirely service-based. In practical terms, this means evidence is collected only from the service (activities, attendance, etc.) and is useful for introspection regarding the functioning of that system. Very often management of

this evidence is carried out via performance targets, essentially setting numbers of activities to be undertaken by various levels of service workers or institutions. The problem is that these targets are very seldom set in terms of the needs of the "customers".

For example, in antenatal care, a "norm" may be set regarding how many antenatal contacts each woman should have. This "norm" is translated into a target for each health facility. But in honesty we cannot be sure that three antenatal visits are better than two; certainly we do not know that they add one third more benefit. Very few biological phenomena (and change in a disease pattern is a biological phenomenon) follow a linear or simple relationship with its determinants. The relationships may be exponential, of one sort or another and affected by many factors.

When coverage of services is very low and the development status very poor, it is rational to start with setting an arbitrary level (such as two antenatal contacts per pregnant woman) based on experience. As the services become more sophisticated, however, the demands for more calculated use of scarce resources become more pressing and a new level of sophistication is required of the information system. If we had data on the impact on development for each level of contact and to a certain extent these data exist, we could base the targets more firmly on knowledge.

If coverage is to be linked with impact in planning and management, it is not very helpful to know that a given

service is provided for 30 percent of the population, except of course where somewhere else in the country it is provided for 80 percent. Then the issue is one of equity. In achieving impact, we need more specific information that can be linked with impact of the service and used for planning and management. Management decisions then can be based on real evidence.

Provided there are data on a sufficient number of people, if the evidence is derived from individuals it is sometimes possible to separate out the component effects of the various interventions.

Data disaggregated to a large number of districts, areas and sites also provide the opportunity for separation of coverage with a given service into its component parts to allow attribution of impact to one or other specific component. For example, in several locations with differing reductions in child deaths due to diarrhoea, among certain areas there may be roughly the same level of sanitation; among others a similar water supply; among others a similar intensity of programmatic oral rehydration; among yet others a similar level of maternal education.

If the evidence is derived from individuals\$provided there are data on a sufficient number of people\$it is sometimes possible to separate out the component effects of the various interventions. Even if the data are not attached to individuals but are collected on a community or area basis, provided there are enough different areas with sufficient numbers of people in each and

provided the appropriate data are collected in the first place, a similar process of stratification will help clarify the relative impact of each type of activity.

Behind the coverage indicators: quality

If our evidence collection tells us that three children in four attended primary school in a given age group, we have to rely on supposition to understand the extent to which the educational contact was meaningful. If certain physical conditions are not met, we would not expect the education to have as much impact; if the attitude of the teacher is not correct, we do not expect the teaching to be taken to heart. Behind the coverage numbers lies the quality of basic education. If the quality is inadequate, the coverage will not have its desired effect. At least as importantly, if there is inadequate coverage, a high quality of the service on its own will not change education patterns. What is needed is attention to quality while expanding coverage.

There are different approaches to "quality control" in development. The first and most widespread approach seeks to define a series of checks at various stages of the activity. Perhaps the best known and most successful among these in health care is the case of vaccination, where a detailed series of checks accompany all stages from packaging and arrival of the vaccine, to transport, storage and administration. Very often the existing information system is nothing more than a record of these various checks. In the generally successful case of vaccination

this process carries the disadvantage that it is expensive to maintain. Applied to other activities, like education, the system of pre-designed checks may be arbitrary and unhelpful.

The second approach to quality, not mutually exclusive of the first, is to begin with impact. If the activity has its desired impact in terms of child survival, the argument goes, then we need only concern ourselves with whether another activity could have a greater impact or the same impact at a lower cost. If the activity does not have the desired impact, then it is time to start using resources to investigate quality, to work out what lies behind the numbers. In several countries (for example, Panama, Liberia and Somalia), mothers were asked if their children were vaccinated or not and if they had suffered measles. By relating the timing of these events, it was possible to estimate vaccine efficacy and vaccine failure rates, indicators of quality of the coverage.

In measurement of coverage, then, it is important and possible to seek clarity as to what is and what is not producing an impact. The population to which the coverage refers should be those who need it, on objective grounds. Components of coverage can be taken into account separately through fairly simple statistical manipulations like stratification. And when the coverage with a particular activity does not appear to be achieving any impact, its quality can be investigated before modifying the activity or discarding it in favour of another that does produce an impact.

Costs

Some issues in the deployment of resources are not and should not be, settled on the basis of cost alone. Equity, visible equity, is very often an important imperative in planning that transcends analysis of short term costs. It is also difficult to put a price on self-reliance in the national and local community context. Having said this, short- and long-term costs should not be overlooked.

Perhaps there is something mildly distasteful about assessing impact in terms of cash and this discourages the use of data regarding costs. However, the real weigh up is not in terms of life or development in relation to cash, but impact in relation to more (or improved) impact, life in relation to more (or improved) lives, development in relation to more (or improved) development. At no point in the equation do we attempt to place a cash value on a life. Within any development strategy, choices have to be made: this activity rather than that; these facilities rather than those. The choices would seek to balance the maximal benefit for the largest number of people, with the minimum wastage of resources. This requires an appraisal of costs of each activity in relation to its impact.

Costs incurred by the services are usually quite readily to hand, for example, how much money, time and manpower is required to launch and maintain a particular programme. In relation to immunisation programmes, there are data about shipping and storage costs, personnel time and equipment. At ministerial level there are details of budget allocations and manpower

availability. At district level, sources may include accounting or order books from service facilities. At community level the sources are even more precise: an agricultural extension worker visits the clinic for so many hours a month and spends her/his time doing specific tasks; vehicle log books and local costs are easily to hand.

A recurring problem is that, because these data are seldom used for planning, they are not kept in a form easily related to impact of specific activities. Sometimes they exist simply to discourage private use of public resources. With slight reorientation, the record can serve both purposes.

The hidden management overheads and the costs of supervision are less readily available. In this second category, perhaps the most frequently forgotten cost is the time of volunteers. Virtually every developing country has at one time or another developed a scheme based on unpaid labour of volunteers, most of the real costs of which are never formally analysed. When some of these schemes are less successful than hoped, there is little analysis as to why this occurred, sometimes leading to erroneous conclusions.

Another type of cost, even less accessible from existing data sources, is the cost of the service to the community. Some attempts have been made to measure these costs in a service-based enquiry. For example, in Uganda it was found that the transportation cost to patients was three times that of the fee for each outpatient visit.

Because of these costs to the users, one

of the prime concerns is to measure the costs from the point of view of those who do not use the available services. The methodology is simple: in each cycle of data collection and analysis (the first cycle concentrates on diarrhoea, another may concentrate on malaria or any other priority problem), people in a household survey are asked about their knowledge of existing services, their use of those services and the reasons they are not used. The idea behind this is to obtain data on costs, including cash, time and attitudes, that can be linked to data on impact and coverage of the services.

An idea of some likely costs to the community can be derived from existing data. For example, in thinly populated areas, it is possible to get some idea of likely distance people will have to travel (geographic access) by drawing circles, called isochromes, around service delivery points. In areas of high population density, the length of wait before receiving attention at the point of service will give some idea of the time costs to the community. Relating this to loss of work time provides some indication of economic access. Certain other costs are not as easily quantified, like those of education and employment, or political costs.

Cultural access, or acceptability of the service, is also difficult to investigate without specially designed studies; it may be possible, however, to investigate superficially by assessing the representation of various minority groups in a given service in relation to their proportion in the general population.

SUSTAINABILITY

A fundamental principle for evidence-based planning is that it should be *affordable*. There is no point in implementing a system that relies on continuous, unpaid contributions by people at the periphery, when those people barely have the means of daily subsistence. Likewise, there is no point collecting data when capacities to analyse it are lacking, or when there are no resources to act on the conclusions. The information system should reflect resources available for development generally and it should have the capacity to expand and become more sophisticated as services expand. These include the use of indicators, Summary Reports, sentinel sites and a "dynamic" registration system.

Several devices are available for the reduction of workload while retaining validity and usefulness of a national level system.

Indicators: words in a common language

Indicators are pointers from which inferences can be drawn regarding impact, coverage and costs. These are tools for measurement that allow comparisons.

Like a sharp knife, an indicator correctly used can be incisive and save a lot of work. Incorrectly managed, however, an indicator can cause considerable confusion. An approach to indicators would be first to establish the priorities, then to decide which indicators best reflect progress in addressing those priorities.

The temptation should be avoided to sit down with the idea of working out a "short

list" of desirable indicators. This leads to a long list that is easily biased towards what is desirable. A different matter would be to consider which two or three questions most urgently require answering in the planning process. It is realistic to examine indicators that emerge from existing data to answer these questions and then to work out what new indicators are needed to fill the gaps.

Indicators, like much else, are

Indicators only filter reality. They do not necessarily reflect reality directly, and they are not all there is to reality.

relative. They are seldom of great value except when used in comparison, for example, between different locations with different levels of coverage and different levels of impact; the comparison may be between subsectors of the population in the same location with different levels of coverage; it may be across time. The important aspect is that indicators only filter reality. They do not necessarily reflect reality directly and they are not all there is to reality. Thus, in every comparison it is necessary to consider the adequacy, the sensitivity and the specificity of the indicators.

"Sensitivity" refers to the property of an indicator to show a change where there is such a change in reality. While perinatal and maternal mortality are good indicators in their own rights, they may not be sensitive enough to indicate positive impact of improvements in perinatal care. Another indicator, increased birth weight, may well show the effect.

A second example of sensitivity comes from a supplementary feeding

programme in one developing country. It provided hot meals for school children and had the apparent cost of US\$6,900 for each case of retarded weight-gain avoided, with each meal costing the equivalent of US\$0.35. Superficially, the cost was high for small benefit as measured by retarded weight-gain among those fed, but a look behind these numbers demonstrates the limitation of the earlier analysis. It is now quite widely accepted that supplementary feeding of school children has potential effects for the family of the children, particularly preschool siblings: the weight gain of children actually receiving the food may be an insufficiently sensitive indicator of the impact of the intervention. There are also more subtle possible effects of the feeding programme, such as improved academic performance or longer-term intellectual abilities, that would not reflect in weight gain.

Indicators should be as specific as possible to the effect they measure. For example, if infant mortality is taken as an indicator of measles immunisation, it may be difficult to demonstrate an impact because so many other things (for example, health education, water supplies, sanitation) also contribute to infant mortality. This becomes important when planners are trying to work out how best to invest scarce resources to obtain maximal impact.

Summaries that link impact, coverage and costs

It is unrealistic to expect the linking of information on impact and coverage to emerge unassisted from evidence collection at each level, community,

district, region and country. For this reason, "Summary Reports" were introduced in first projects on evidence-based planning in Honduras and Nicaragua. These are simple sheets of paper that combine data on impact, coverage and costs for assimilation and use at local, regional and national levels.

The formats are designed to answer specific questions asked in the planning process; consequently, the answers are direct and concise. For example, if one seeks the educational components affecting deaths from diarrhoea, a questionnaire or (less easily) analysis of routine data could be used to derive indicators of maternal education. Number of years of education can be cross-linked with mortality or morbidity from diarrhoea (impact) and with, for example, reasons for not attending adult education classes (costs). This can be used for discussions at the community level, among teachers, or between the Ministries of Health and Education.

The logic behind this approach is to initiate a repeatable series of cycles that delineates a priority problem; lists questions to be answered in the planning process; reviews existing data in that area, noting what is not available; collects data in locations that can be followed over time for a longitudinal perspective; generates a communication strategy; and reallocates resources that address the problem in question.

The point is not that these Summary Reports contain "the truth, the whole truth and nothing but the truth". Far from it. Much more important is the obviousness to virtually anyone who looks at a Summary Report that there is much more to health and to education, than

appears in the simple cross-tabulations. They usually say so. And that is the usefulness of the Summary Report. It is a tool to get people talking, to express their ideas and experience.

Sentinel communities: concentration of measurement resources

Sentinel sites are viewed here as a representative selection of locations in different administrative regions throughout the country where, with a concentration of measurement resources, it will be possible to generate more detailed and reliable information. This information supplements the existing information system as it evolves over the years to cover a greater percentage of a country's population.

The main activity in the sentinel sites is a series of data collection and analysis cycles. Using house to house survey techniques that radiate outward from a constant point, data can be obtained that cannot ordinarily be gleaned from the routine data base: for example, attitudes towards the health service and reasons (costs) for failure to use available services. Because the data collection will be standardised across all sentinel sites, biases will be consistent. This aspect of the scheme will allow validation of data generated by the routine system.

The sentinel sites are seen as training centres, or demonstration areas, for strengthening analytical capacities. Summary Reports, combining data on impact, coverage and costs, have been developed for application in the national information system to assist decision making at the various levels. Additional

resources available in these sites in terms of expertise, personnel and equipment (mostly hand-held programmable calculators) are intended to serve a catalytic function for using information in the country's planning process generally.

Sentinel sites are not pilot areas to test special interventions. They are intended to be representative of conditions and services, different only in that they are a concentration of measurement resources. This important difference was taken into account in the initial design of sentinel community surveillance by not contacting all houses in each site. In Nicaragua and Honduras, after a rapid baseline demographic study of all houses, three of every four houses with children under the age of five years were contacted. These same households would be contacted every two or three months in future cycles, data from which could be linked to that collected in the demographic survey and the subsequent, more detailed enquiries. The remaining quarter of households were not contacted again for the remainder of the two years in an attempt to measure any "repeated contact bias", should this develop. After two years, some of the sites were to be dropped and replaced by fresh ones. After four years, the baseline study was to be repeated in all sites, new and old.

For data collected in sentinel sites, as by any mechanism in an information system, the first concern is that it should be manageable in volume and concept. The data thus should be kept simple and sparse, increasing in quantity and sophistication as the capabilities increase. The second concern is that statistical

analysis should be accessible also at regional level. Statistical analysis of data from the sentinel sites is relatively straightforward. Each site is treated as a separate population, as indeed it is. In Nicaragua and Honduras, this provides 36 levels of coverage, with 36 levels of cost and 36 different impacts of each health or development activity to be considered.

The strength of the sentinel sites approach lies in its potential links with the existing information system. First, the existing data base is examined and whatever data are lacking or unreliable can be sought in the sentinel sites. A particular advantage here is the ability to obtain data that are not normally accessible (such as why people do not use health services) or data not linked to individuals in the routine system (for example, incomes, nutrition, or maternal education).

Second, people involved in the collection of data and the design of collection-analysis cycles are part of the existing system: the sentinel sites are not managed as a parallel information system. Analytical tools and Summary Reports may be used in the rest of the system, perhaps enriched by the detail obtained in the sites. After a few years there will be substantial local expertise in the design and conducting of special studies in a variety of areas.

Third, the sites can be used to test the validity of specific components of the existing scheme, perhaps allowing the calculation of correction factors if the under- or over-estimation is consistent between the different regions. Some use eventually can be made, for example, of

death registration even before this reaches universal coverage. It should also be possible to test systematically the validity of indicators used in the planning process.

Fourth, the data on impact and coverage from a range of sites will allow derivation of norms and targets based on biological association rather than on prejudice or inspiration.

Fifth, the outcome of the sentinel site activity will feed into the planning process with the ultimate objective of contributing to development. Resources can be tracked away from activities that are not effective towards those with proven impact. It should help to assess whether the allocation of resources corresponds to a need-based distribution, that is, whether the system maximises equity.

The sentinel scheme can be used as a reference base for special projects that are ordinarily outside of the routine measurement process. For example, if an international or bilateral agency runs an integrated development project (sanitation, agriculture, water supplies, or housing) it can be required to establish an accessory sentinel site in that area for progress to be monitored within an integrated structure.

Finally, sentinel sites are not fixtures in the development landscape. Their size and number are limited only by the availability of resources. As these increase, the system can expand to complement the other components of a national system in all locations.

Dynamic registration

The term "dynamic" speaks to the interaction around the data collection

resulting from the education modules designed to go with each data collection cycle. It might also refer to the changing focus of the system from month to month, while generating a "core" of baseline data, analysed periodically in conjunction with education modules on specific topics. This sort of system decreases the amount of boring, repetitive work to be done at the regional and central levels, while requiring a fairly strong initial input in designing collection formats, data Summary Reports and education modules.

It is also possible to develop a series of more limited registers that address specific problems. In the health sector, for example, until sufficient knowledge is obtained on a condition, it may be worthwhile collecting details on each snakebite (if snakebite is a particular problem in one area). If it is a very frequent problem, the scheme could involve collecting details on every fifth snakebite to reduce the number of data units collected and thereby sustaining the quality of data.

Most countries have a list of 30 or more diseases that are notifiable by law. The additional workload that this implies for already overburdened health workers and the fact that these data are seldom if ever analysed in the context of planning, results almost inevitably in underreporting with unquantified biases from place to place. If, by way of a practical example, a decision was taken by national level planners that tuberculosis required additional attention, they could designate between one and three special times of the year for TB data collection, analysis and education. One of these would be

scheduled at the anticipated peak season (in the case of TB, winter) and perhaps another just before the peak season.

After the decision regarding the disease in question and the appropriate times of year to measure it identified, each cycle has several practical steps. First, health planners and educators prepare a series of education modules about the disease. These might focus on the respective value of different therapeutic regimes; for nurses working at rural clinics, the focus might be on improving diagnostic practices; in the mass media, the module might focus on the nature of the disease, its social context and what can be done to diagnose and treat it.

Along with these education modules, a Summary Report is prepared for data collection at various levels. The heightened consciousness about the disease at all levels results in a higher rate of diagnosis and better recording of data. These reports, on the same single page, synthesise data regarding health status (from which impact can be calculated), coverage and costs of efforts against TB.

Provided the Summary Reports are simple, they can be analysed rapidly (addition, subtraction, multiplication and division) at the level where they are collected before being sent to the supervisory level. A rapid feedback of how this compares with other places is thus possible. The following month a different cycle of education, data collection and analysis could be commenced. Each month or two, another cycle could be implemented, each considering between one and five diseases. Obviously when an

education/data collection/analysis cycle on TB is repeated, it will not require more than revision of the education modules. The time can be used to develop modules for collection of other data during the same time as the TB cycle, preferably limiting it to two or three diseases per cycle.

The obvious disadvantage of this dynamic registration process is that one cannot monitor all diseases at once. The main advantage is that it forces planners to decide upon priorities, to remain alert to changing priorities and it provides accurate and usable data about their priorities.

NO BLUEPRINT EXISTS

Perhaps the most difficult guiding principle in measurement is that there is no such thing as a perfect information system.

There is no easy recipe to be applied as one might get from a cookbook. As the situation in the country changes, as the cash resources fluctuate, as the human capabilities increase, as longer-term planning becomes more realistic, so the information system will have to evolve. It is necessary, therefore, to develop capacities and practices in several directions at the same time to facilitate this evolution: better use of existing data sources, streamlining of routine data collection schemes and development of capacities to carry out and to analyse special studies.

Routine collection of data, the methods traditionally viewed as "the system", cannot realistically be seated at the same time on two stools: depth and breadth of information.

Each data collection procedure (routine collection or special study) has its own strengths and weaknesses. There are several advantages in a routine system. If data are available for all periods of the year, going back over several years, it is possible to assess the effects of seasons. This information can be used to prepare for a season of high demand on services. If the scheme is reasonably comprehensive and if the data can be analysed soon after it is collected, the scheme may be used to monitor trends in the area under study. Routine record-keeping is also appropriate for the monitoring of use and flows of resources.

Routine collection of data, the methods traditionally viewed as "the system", cannot realistically be seated at the same time on two stools: depth and breadth of information. Data emerging from the existing system may be full of unquantified biases or may simply not address the questions raised in the planning process. Moreover the people who do not come in contact with the service delivery system, those on whom planners need information most of all, are excluded from the service-based routine collection of data.

Because filling in of reports is more easily checked by supervisors than the process of stimulating community participation, routine data schemes can take on a symbolism beyond their real value. In this way they can detract from other meaningful contributions of health workers to their communities.

Unless one reviews well-kept clinical notes, data on individual benefit from coverage with specific activities can

be difficult to link with other individual characteristics in the assessment of impact. Perhaps the main disadvantage, however, is that if too much data are collected by already overworked and possibly underpaid development workers who do not actually see what happens to the data, interest in the data collection will soon be lost. When this happens the quality will start to deteriorate and with it the usefulness of the entire data collection exercise.

Whether strengthening a routine system or special study capacities, a constant threat to the durability and usefulness of the scheme for planning purposes is the planner who says: "Since we are collecting data anyway, we may as well get this, that and the other too". All of us do it to one extent or another. It is easy to forget what it is to do the work, when we only have to decide what work is to be done.

The first clear step for reengineering a routine data system is to start using it in the planning process. This can be done, issue for issue, linking data on impact, coverage and costs.

Then, following the guiding principles outlined at the beginning of this essay, the system can be revised piecemeal. The burden of work can be decreased by limiting it to core indicators; limiting the areas where data are collected and reinforcing those areas with additional

The most difficult part is to persuade planners to stop routine collection of data that are not useful for planning: the temptation is always to collect more.

resources; and limiting the time period over which each type of data is collected.

These changes are only of relevance, however, if they happen with strengthened analytical capacities. Without this, an innovative data gathering process is as meaningless as its predecessor.

Intersectoral nexus and community voice

The guiding principle that a common framework -- impact, coverage, costs -- be adopted can be applied to communication with other sectors. Intersectoral collaboration, to some extent, has become a hollow incantation in the development process.

The problems with intersectoral links between ministries or between subsectors within one ministry are legion. But at the periphery of the development process, this is not really an issue.

This is why the best place to start with intersectoral links is in the household, where the links are strongest. And one

For ordinary people there are no brick walls between sectors; the gaps between health and education, or occupation and water supply, are not nearly so finite.

useful mechanism for drawing on this strength, allowing it to guide planners at regional and national levels, is measurement.

If the information system allows us to step outside the confines of the ministry, or office, or health facility and into the community, as is the case in special studies of the type done in the sentinel sites, there is a good chance of bringing this mechanism into effect. The process can be started by regional and local managers analysing existing data in

terms of impact, coverage and finally the costs. Even using conventional indicators, it may be necessary to go out into the community to collect data to make up the shortfalls or to test the validity of existing data sources. This provides the opportunity to open up the measurement process and, later, the planning process for wider participation.

It is quite feasible in a special study to obtain basic data on other sectors and to

The process of negotiation between local workers and the community members is the crucial link.

include this in the Summary Reports that relate impact, coverage and costs. It is important that the Reports are discussed at the level where the data are collected. The analysis, in the general sense of discussions around the data, can be conducted in whatever community level structures exist. This implies the issues have been made sufficiently clear and specific for brief and pointed discussions. It implies an escape from the blunderbuss type of measurement.

The second important aspect is that the process is continuous. If the people in the community feel other factors are also important, a future cycle of data collection and analysis can address these issues in a way that is apparent to community members, reinforcing the idea that their voice is significant.

Commenting upon a summary format relating child mortality, vaccination and the costs to the local health service of this activity, for example, community members may feel that the costs of their transport (or perhaps the condition of the roads

during the wet season) are not taken into account. A subsequent study cycle could focus on the intersectoral issue and can provide planners in the health sector with a valid rationale and the local details to open intersectoral discussions with planners from the other sector.

The process of negotiation between local workers and the community members is the crucial link. This can be supported from the central or regional levels by the structuring of education modules, as discussed in relation to "dynamic registration". It may help to demonstrate the feasibility and positive outcome of these discussions in other locations, perhaps using tape-recorded testimony or film, to encourage the process. Training of workers not only in data collection and the use of Summary Reports, but also in management of these discussions, can be a useful boost to confidence and to the quality of community interaction.

Finally, political will at all levels is necessary to bring about this devolution of the measurement process.