Background paper for the Technical Consultation on Effective Coverage of Health Systems

27–29 August 2001

Rio de Janeiro, Brazil

World Health Organization

Geneva

2001

Department of Health Service Provision
World Health Organization
1211 Geneva 27
Switzerland

Contents

Executive summary

The health system performance assessment framework provided in the *World health report 2000* clearly distinguishes between the final and intermediate goals of the health system, and acknowledges the instrumental role of the intermediate goals in achieving the final ones.

At regional consultations held since the 107th session of the Executive Board, it has been suggested to incorporate the measurement of *effective coverage* of a selected group of interventions into health system performance assessment. The measurement of *effective coverage*, as an intermediate goal, is expected to link health system performance measurement more directly to managerial practices and decision-making processes at local, regional and national levels.

This paper serves as a background for the technical consultation on *effective coverage*. The first section, after providing a rationale for the World Health Organization's renewed focus on coverage and reviewing its past work in this area, discusses the conceptual framework of *effective coverage*. The paper argues that the concept of *effective coverage* encompasses more traditional concepts of access, utilization and effectiveness. The term effective coverage is defined in this paper as the proportion of persons who have received effective interventions.

The use of information on coverage has rarely been considered in the planning or management of health services. In order to make the measurement of *effective coverage* a useful management tool at different administrative levels, five domains of coverage measurement are proposed: availability coverage, accessibility coverage, acceptability coverage, contact coverage and effective coverage. The analysis of these domains helps in understanding constraints in different areas of health service delivery functions that prevent attainment of a desired level of *effective coverage*.

The second section, based on concrete examples from different health programmes, describes coverage indicators and measurement instruments with regard to their validity, reliability and comparability.

WHO's previous focus on vertical programmes has resulted in a multiplicity of separate instruments that measure disease-specific coverage. The analysis of different instruments and data sources to obtain coverage figures shows that they vary significantly in terms of their validity, reliability and comparability. There are some generic biases that are common across different measurement instruments and cause noise in coverage figures. Some of the biases are specific to the measurement method or the intervention itself.

The analysis of the experience accumulated in the measurement of coverage permits the following observations:

- The definition of the effective unit of intervention is crucial to constructing the indicator of *effective coverage*. For long-term intervention programmes or lifelong interventions, the unit of an effective intervention is hard to determine precisely. In this case, *effective coverage* should measure whether a patient is receiving the effective intervention at a point in time or not. If this measurement is not possible, then another domain of coverage (contact, acceptability, accessibility, availability) can be used instead. The definition of the case and the effective unit of an intervention are important factors affecting comparability of results.
- When the true occurrence of the intervention is difficult to measure, due mainly to recall bias, inaccurate administrative records or death of a patient, available figures must be adjusted. Different adjustment methods are identified, such as comparison of the occurrence of the event in different population subgroups; adjustment by specificity and sensitivity of

the data collection tool; direct measurement by an objective reference test in a subsample of a study population; adjustment for age, sex, literacy rate or other individual characteristics, etc. Such measures can increase the validity of the numerator.

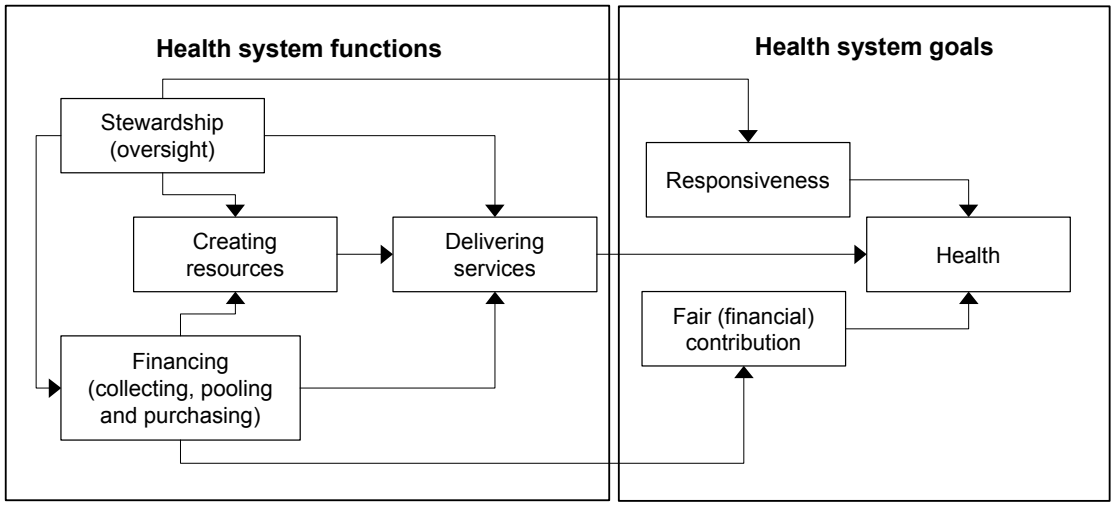
- Reliability in the measurement of the magnitude of coverage can be improved by standardizing the measurement procedure.
- When precise and direct measurement of the magnitude of needs (prevalence and incidence of disease, or prevalence risk factors) is difficult or impossible, estimated prevalence or incidence can be used. If there are both direct and indirect methods for identifying prevalence and incidence, the indirect method can be used first. The results can then be adjusted by direct methods applied to a representative subsample of the study population. This would improve the validity of the denominator.
- In order to ensure comparability of the coverage figures, it is important to have the same definition of the case and the effective intervention; the measurement tool should measure the event among different populations in the same way; and the study population across different areas should be similar. If the last two conditions cannot be met, then calibration methods must be used in order to make the results comparable.

1. Effective coverage in theory: linking the measurement of effective coverage to health systems performance

1.1 Introduction and rationale for WHO's work on effective coverage

The *World health report 2000* has provided a new framework (Fig. 1) for the assessment of health system performance and initiated a process of more rigorous scientific work and analysis on this subject. By clarifying health systems goals and linking them to health systems functions, the framework is intended to help Member States measure their own performance, understand the factors that contribute to it, and respond to the needs and expectations of people (51). A better understanding of health systems goals, functions and their determinants is crucial for designing effective health policies.

Figure 1. Health system functions and goals



The *World health report* made a clear distinction between the final and intermediate goals of health systems, and acknowledged an instrumental role of the latter for achieving the former[1](51). Regional technical consultations on health systems performance assessment held since the 107th session of the Executive Board (EB) have indicated the need to expand performance measurement to the intermediate goals, in order to more directly link health system performance measurement to managerial practices and decision-making processes at local, regional and national levels (52;53).

Over the years many intermediate goals have been suggested such as access, effectiveness, appropriateness, efficiency, acceptability, continuity, etc. Some of these goals are still applied to the performance measurement and management in a number of countries (16;19;29).

Measurement of the attainment of all possible intermediate goals seems quite a difficult task. It will be more efficient to select one or two intermediate goals that would be feasible to routinely monitor and would capture as many dimensions included in other instrumental goals as possible. Such an intermediate goal will naturally complement and fit into the existing WHO framework of health system performance assessment.

1 Final health system goals: health, fair financing and responsiveness (see Fig. 1)

At the regional consultations on health system performance assessment, it was suggested that effective coverage of a selected group of interventions could be a good intermediate goal that would meet the aforementioned criteria. It was felt that effective coverage captures many dimensions of health system functions, and can be used to trace the degree to which the health system carries out critical activities that have an impact on people's health (52;53).

Responding to the suggestions and recommendations of the Member States and WHO regional offices, WHO headquarters plans to incorporate the measurement of effective coverage of critical health interventions into the health system performance work. The information on effective coverage would be obtained and used, not only at the national levels, but – even more importantly – at the subnational levels. A number of technical challenges must be overcome in order to obtain valid, reliable and comparable data on coverage, capture the performance of nongovernmental and private providers, and devise validation instruments for improving the use of routinely reported service data for the assessment of coverage.

1.2 Historical review of WHO's work on coverage

The concept of coverage as a domain of health system indicators emerged in the late 1960s to the early 1970s. It enjoyed particular attention in the early Health For All era. Later, the focus apparently shifted towards access and utilization. Probably this was politically more acceptable for policy-makers, because access and utilization measures are concerned mostly with resources and can be more easily used for justifying investments (37;48). In the mid- and late 1990s, rising concern with equity resulted in the shift of attention towards measuring inequalities in health and health care.

One of the first large-scale efforts that contributed to the evolution of the concept of coverage was "WHO/International Collaborative Study of Medical Care Utilization", initiated in 1964. The study was carried out in 12 study areas in 7 countries, for more than 10 years. The study attempted to answer three broad sets of questions:

- To what extent are differences in the organization of health care and its resources reflected in differences in the use of health services?
- What behavioural and social factors influence patterns of use in different regions?
- What methodological problems must be overcome before valid statements can be made regarding patterns of use in different regions? (46)

The need for improvement in the coverage and content of health services and for greater equity were among the key concerns of WHO that warranted the International Collaborative Study (46). The study stressed the need for population-based data on health services in addition to institution-based data, which tended to capture only those people who had used services (46) (45). By attempting to link service data with the population-based data and understand people's needs, the International Collaborative Study was a step towards describing coverage of population by health services.

For the estimation of needs, the following domains were used: social dysfunction, perceived morbidity, special psychobiological dysfunction, perceived dental morbidity and perceived visual morbidity. On the provision side, the study measured the availability of health system resources. In terms of service utilization, the study measured use of hospital care, medicines, physician visits (outpatient), dental services and visual services (45).

The study described the use of services in relation to perceived needs, but did not measure the coverage of specific health interventions as such. In this study, the perceived needs were used as an independent variable to explain patterns of utilization, rather than as a denominator of an indicator that would measure how well health systems address population needs.

Together with the Alma-Ata Conference, measurement of coverage, among other health system goals, became an important item on the initial agenda for monitoring progress towards Health for All. The Director-General of WHO and the Executive Director of the United Nations Children's Fund (UNICEF), in their joint report to the International Conference on Primary Health Care in Alma-Ata, stressed the need to evaluate the coverage, accessibility, efficiency and effectiveness of primary care (44).

This increased focus on coverage, accessibility, and effectiveness reflected in WHO's project, "Measurement of Coverage, Effectiveness and Efficiency of Different Patterns of Health Care". The project was initiated in 1978. Its overall purpose was to contribute to the "attainment of health for all through total coverage with effective health care". The project had three main objectives:

1. To collect, analyse and disseminate national experiences on the existing approaches, methods and results in the field of measurement of health care coverage, effectiveness, and efficiency.
2. To contribute to the development of a quantitative basis on which countries may:
   2.1 do a policy assessment through measuring their progress towards the goal of health for all and in implementing primary health care;
   2.2 compare, appraise and select alternative patterns of health care in terms of their relative potential for coverage, effectiveness and efficiency.
3. To develop methodological guidelines that will be of use to the countries for the measurement and assessment of the three criteria (47).

The project carried out assessment studies in several countries, including Burma (now Myanmar), Costa Rica, Egypt, Finland, Malaysia, Nigeria, Republic of Korea and Senegal. The studies focused on the following areas: maternal and child health (MCH), health education, nutrition, water supply and basic sanitation, immunization, control of local endemic disease, treatment of common ailments, and supply of drugs (47).

For measuring coverage, 37 indicators were developed. These indicators, together with a long list of effectiveness indicators, would make the measurement of achievement of health systems goals quite a challenge.

The project contributed to the development of a methodological and conceptual basis for the global process of "Monitoring Progress Towards Health for All by the Year 2000". The efforts to develop an exhaustive set of indicators continued. The indicators for monitoring progress towards Health for All were grouped in four categories:

- Health policy indicators
- Social and economic indicators related to health
- Indicators of the provision of health care
- Health status indicators (49).

Each of these categories included several subcategories. Several approaches were suggested for the measurement of coverage:

- Separation of different levels of the health system, for example primary care as distinct from referral levels.
- Separation of different functions – maternal and child health, for example, as distinct from general curative care.
- Breakdown of coverage into availability, accessibility and utilization (49).

Strategies for Monitoring Progress Towards Health for All were updated several times. Gradually the term "primary" acquired a variety of technical and political connotations, which resulted in a multiplicity of meanings and loss of clarity of the concept of coverage. The primary health care movement was concentrating almost exclusively on people's presumed needs and paid little attention to people's demands for health care. As a result, health policies became too supply-oriented (51).

In the 1990s, interest in the concept of coverage started to diminish. The increasing awareness of poverty in the world and decreasing public financing of health care systems in developing countries determined the shift of attention from Health for All to cost-effective allocation of resources and mobilisation of public and private money for health sectors. In the late 1990s, the rising concern with equity shifted interest to the measurement of health inequalities among different population groups. The measurement of coverage and effectiveness of health services gradually lost the attention it had previously been accorded.

Despite the efforts described above, the measurement of coverage has not become common practice in health system assessment. Tanahashi suggests three major reasons for this:

"First, there has been some confusion about the concept of coverage, which has been measured and interpreted differently for various purposes and occasions. Secondly, the measurement of coverage invariably requires an assessment of "population", with which few service personnel have been concerned. Thirdly, the use of information on coverage has rarely been considered in the planning or management of health services" (37).

1.3 Conceptual framework for understanding and measuring coverage

1.3.1 Definition

As mentioned earlier, there is some confusion about the term "coverage", and significant variation exists in the contexts in which it is used. The term that will be used in this document hereafter and during the technical consultation is e*ffective coverage*. Effective coverage can be defined as:

The proportion of the population in need of an intervention who have received an effective[2] intervention

This definition applies both to personal health services and nonpersonal public health interventions.

The proposed definition of effective coverage is a provisional one and is based on the definitions found in different sources (15;18;22;23;34;37;39;47;48;56).

It should be noted here that the term "coverage" is often used also in the health insurance context, where it means the availability and degree of protection to the population or individuals from the financial risk associated with illnesses (15;39).[3]

2 Effective services are defined here as a minimum amount of inputs and processes that are expected to produce desired health effects if used by individuals or applied to the population at large. A more detailed explanation is provided in Section 3.2 - Effectiveness).

Very often in the literature, indicators labelled with the term "coverage" represent the ratio of resources or activities over population (for example, 1 primary health care worker per 1,000 population, or 3 consultations per capita, etc.). Such an interpretation of coverage is misleading, and conceals the distributional aspect of health service delivery: an average number of PHC visits per capita can be affected by a small number of frequent users of services, while a larger population of the poor, presumably in greater need, might not be getting necessary care.

Therefore it is important to note that the numerator of the coverage ratio should indicate the number of population units (individuals, houses, villages) receiving services, and the denominator should refer to the population that would need the type of services indicated in the numerator. Where true coverage has not been measured often, it is understandable that availability or accessibility of resources per unit of population size be used as proxies for coverage. From the political point of view, it may be preferred to emphasize such achievements, since they are the immediate, visible and durable results of decisions made by health policy-makers (50).

As mentioned earlier, coverage is an instrumental goal that contributes to the achievement of the final goals of health systems (51). To reflect this relationship between health and coverage, the use of the term *effective coverage,* as proposed in the working definition, seems more appropriate.

When discussing the concept of coverage, the following question comes inevitably: What is the unit of effective intervention that is expected to produce health gain in the population if a desirable level of coverage is attained? To be able to answer this question, we must include the concept of effectiveness, among other concepts discussed later, in the definition of coverage.

For example, if a policy-maker is interested in the reduction of child morbidity from diphtheria, he or she should be concerned with the level of coverage of children under age one by three doses of effective DPT vaccine[4] administered at the correct intervals. Only if the vaccination is conducted properly will it be effective[5] (28;40;55). Another example could be antenatal care coverage. In this case, an effective unit of coverage should be determined at the start: it could be either three or four visits to an antenatal care facility. Furthermore, the number of visits alone

---

3 In the context of health service provision, the use of the term "coverage" is not homogeneous. In order to avoid possible confusion in the interpretation of the concept, it might be worthwhile to mention a few other definitions as well.

For example, in the Glossary of health care terminology (14) the definition focuses on the aspect of the availability of health services to the population, and coverage is classified as physical (geographical coverage) and functional coverage. The latter can be passive or active.

The geographical/physical coverage is the ratio between the number of health facilities per administrative or geographical unit and the size in population and area of the unit. This is a physical coverage and describes the "zones of responsibility" of health institutions (14).

The functional coverage characterizes the catchment areas of health facilities, and is defined in terms of (a) the utilization of health facilities by the population (passive coverage) and (b) "the penetration into the population by mobile elements of the health institution, or the radius along which population is contacted (active coverage)". Functional coverage is a practical or operational form of coverage that corresponds to the "zones of effective activity" of health facilities (14).

4 Effective DPT vaccination is that which is administered according to WHO recommended schedule: the first dose at the age of 6 weeks, and the remaining two doses spaced at minimum 4-week intervals.

5 It should be noted that effectiveness here is assumed rather than objectively ascertained. It is expected that correct vaccination will result in a sufficient level of immunity, but to check whether the expected effect has been obtained, a seropositivity test must be administered.

will not ensure desired health effects for a future mother if she does not receive appropriate and good-quality care during the visits (24;25).

Thus, the introduction of the term *effective coverage* signifies the importance of understanding and measuring coverage in a comprehensive way.

1.3.2 Conceptual elements of *effective coverage*

Various interpretations of coverage observed in the literature reveal quite blurred boundaries between the concepts of coverage, access and utilization (10;11;17;34;38;45). Sometimes the concept of coverage is equated to access, and sometimes to utilization. The question arises: Are these different concepts, or are they the same? Probably neither. It might be more appropriate to think of access, utilization and effectiveness as determinants of effective coverage.

Access

There are several frameworks proposed for understanding access (1-3;8;30;31). Some authors define access as availability of services whenever and wherever patients need them (9). Others adopt more general definitions, such as "fit" between the clients and the health system (30).

Some theoretical models of access view it as a function of predisposing, enabling and health system factors. Predisposing factors include variables that describe the propensity of individuals to use services. Enabling factors include the means available to individuals for the use of services. Health systems factors include resources, procedures and institutions through which health services are delivered (1;45).

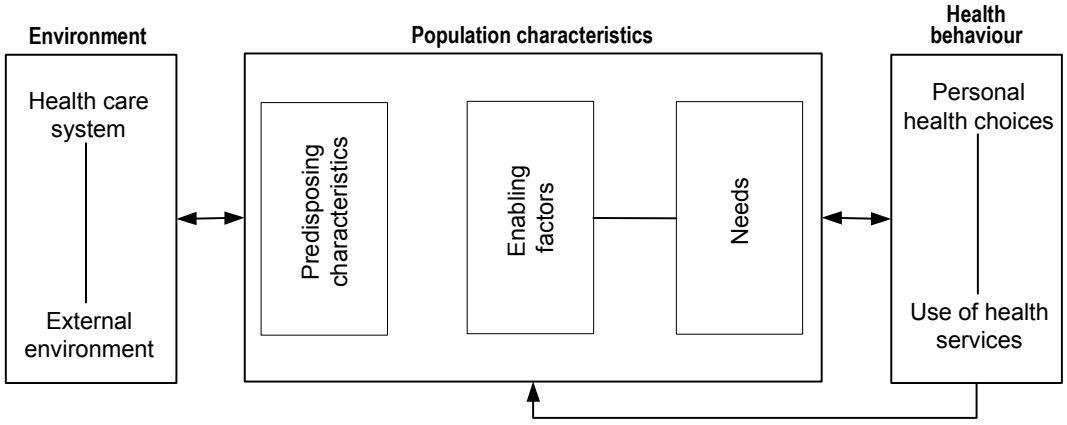
A more practical model of access is proposed by Penchansky and Thomas (30). Their model seems to be related to the model described above, but it lends itself more easily to measurement. This model proposes four dimensions of access:

- Availability: the relationship of the volume and type of existing services to the volume of clients and their needs
- Accessibility: the relationship between the location of supply and the location of clients
- Affordability: the relationship of the size of the required financial contribution for obtaining services and clients' ability to pay
- Acceptability: the relationship between clients' attitudes, expectations, sociocultural characteristics and actual characteristics of existing providers.

Utilization

Most of the conceptual frameworks that analyse health service utilization are based on the model developed by Andersen and Aday (1;2;31). The basic elements of this model have been discussed above. This model is represented in Figure 2**.**

Figure 2. The model of access to health care



Source: Phillips K.A., Morrison K.R., et al. 1998

It is important to notice that health service utilization is determined not only by access (population and environment characteristics) but also by health behaviour. An interesting model of health behaviour analysis is proposed by Rosenstock. The model identifies four determinants of health behaviour: perceived susceptibility, perceived seriousness, perceived benefits of and barriers to taking action, and a trigger to action (32).

Acceptance of personal susceptibility varies significantly among individuals. At one extreme, a person may reject any possibility of his contracting a given condition. In a more moderate position, a person may admit the statistical probability of the occurrence of an adverse condition or accident, but the reality of this probability seems very small to him. Finally, a person may express excess fear of being in real danger of contracting the condition.

Perceived seriousness refers to convictions concerning the possible outcome of a disease in terms of death, disability, effect on job, family life, social relations, etc. For instance, a person may not believe that a sexually transmitted infection is a serious health problem, but nevertheless seek care because of the tension that this condition might cause within his family.

Even if someone accepts susceptibility to disease and believes in the seriousness of its outcomes and social implications, there might be several options for action among which to choose. The selection of a particular course of action will depend on the beliefs regarding the relative benefits (or effectiveness) of known alternatives and barriers to taking different actions. For instance, one may refer to a traditional healer instead of going to a health facility.

The understanding of susceptibility and seriousness provides the force for action, and the perception of benefits and barriers directs the user to the preferred path of action. Even the combination of all these factors may not result in an overt action without a trigger, however. Such a trigger might be internal (perception of bodily states) or external (interpersonal interaction, health education, etc.)

Another important factor that affects utilization of services is provider performance. Small area variation studies show significant variation in health care utilization patterns due to provider-

related factors (27;42) As with the environmental factors[6] mentioned above (Fig 1), provider-related variables are measures of the context within which utilization occurs, yet their influence remains to be fully explored (31). Provider-related variables of utilization may vary from individual characteristics (gender, training) to the organizational culture and behaviour of a health institution. It has also been shown that much of the variation in utilization may be due to physician decision-making (20;43).

Effectiveness

One of the general definitions of effectiveness is the extent to which health intervention delivers its intended outcome or results in a desired process, in response to needs (5). Effectiveness of health interventions can be considered as a function of different variables, including efficacy, inputs (amount and quality of resources), quality assurance mechanisms (process of service delivery, provider performance), patient compliance and health behaviour and external factors (environmental, biological, social, etc.). This breakdown is similar to various models suggested for the analysis of quality and effectiveness of care (7).

Effectiveness of health interventions can be measured through impact evaluation studies that use plausibility and probability inferences to determine if the impacts observed are due to the interventions, and what is the statistical probability that the intervention had an effect (12;54).

It is very important to notice that we are not using effectiveness in the context described above in relation to the concept of *effective coverage*. *Effective coverage* does not measure the effectiveness and impact of intervention. It measures the health system's efforts (performance of health service delivery function) in terms of providing the population with a set of services that are believed to be effective if individuals use them.

Measurement of the true effectiveness of those health interventions is beyond the measurement of *effective coverage*, as it will be captured by the measurement of health and its distribution. A concept of effectiveness similar to that which is used here in relation to *effective coverage* is referred to as *technical quality* elsewhere in the literature (57).

Another concept captured by *effective coverage* is compliance with treatment procedures and protocols. This is particularly important in relation to chronic diseases, when a time-limited unit of effective intervention cannot be defined. Measurement of *effective coverage* of population with such interventions should entail measurement of the proportion of the population who fully comply with a recommended treatment regimen. For example, in the case of measurement of TB/DOTS programme, we should be concerned with the proportion of TB patients who have completed the entire course of treatment[7].

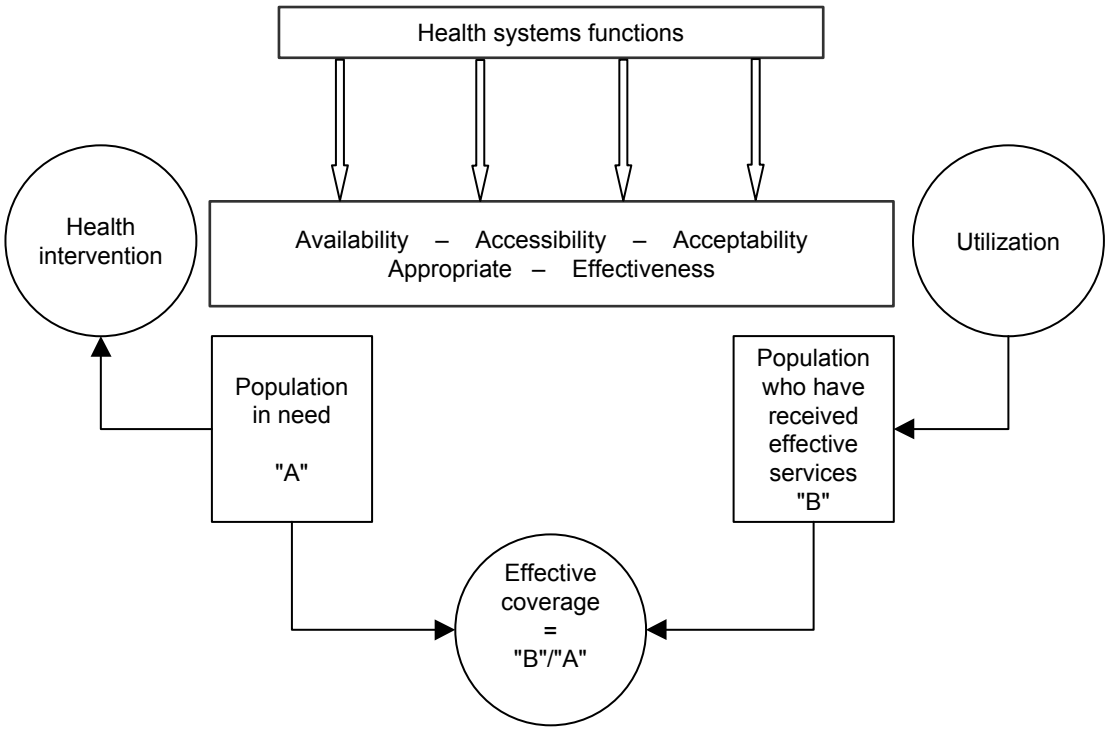
Even more difficult is to incorporate in the concept of *effective coverage* the compliance with treatment for those chronic diseases that require lifetime management, such as diabetes mellitus or hypertension. In contrast to TB, these diseases cannot be cured. Measurement of *effective coverage* in this case should be concerned with measurement of the proportion of affected people who regularly take required medicines or receive necessary procedures.

6 Environmental factors include characteristics of the health care delivery system, external environment and community (31)

7 As discussed earlier, the question of whether the full course of treatment was effective is beyond the scope of measurement of *effective coverage*.

The relationship between the conceptual elements of *effective coverage* (access, utilization, effectiveness) discussed above is represented in Figure 3. In this framework the process of health service delivery starts with understanding the real needs of a population through objective methods of assessment. This should inform health policies and result in designing effective interventions (or sets of interventions) to address the needs. In order to attain a high level of *effective coverage* and thereby maximize the probability of achieving a significant health gain, the interventions should be available, accessible, acceptable, affordable and effective. These characteristics of health interventions depend on the four functions of health systems: resource generation, service provision, finance and stewardship. *Effective coverage* in this diagram is the ratio of population "B" to population "A".

Figure 3. Conceptual framework of effective coverage



1.3.3 Measurement of coverage

In the previous section we discussed the conceptual framework of *effective coverage* as an intermediate goal of health system performance. This was necessary for understanding why e*ffective coverage* is a good intermediate goal, and how it comprises more traditional concepts such as access, utilization and effectiveness (in a very specific context). This conceptual construct may not be sufficient for translating the measurement of coverage into a management tool, however.

In order to make the measurement of coverage more operational, we propose five domains of coverage measure, based on the conceptual framework described above and similar models found in the literature (37;57):

- Availability coverage
- Accessibility coverage
- Acceptability coverage
- Contact coverage
- Effective coverage.

These five domains represent the determinants of the conceptual elements of *effective coverage* discussed in the previous section. If a district health care manager is not satisfied with the level of *effective coverage* of the population of the district with essential health services, he or she should be concerned with the causes of the poor performance and should try to understand the factors that prevent the achievement of a desirable level of *effective coverage.* Measurement of different domains of coverage may help in such an analysis.

Availability coverage

Availability coverage shows what resources are available in what amount for delivering an intervention. This might include: number of health facilities, number of personnel, availability of technology (drugs, equipment, etc.). In other words, availability coverage relates the capacity of a health system to the size of the target population (22;37;56;57).

Two ways of measuring *availability coverage* are suggested:

- The proportion of people for whom sufficient resources and technologies have been made available
- The ratio of resources to the total population in need[8]
- The proportion of facilities that offer specific resources, drugs, technologies, etc.[9]

Information on the availability of resources is very important for informing resource-generation policies. There are different potential sources of information for determining *availability coverage.* For example, surveys of a sample of facilities can provide more detailed and reliable information on the availability of key inputs. Population surveys can provide information about the availability of resources from the consumer point of view.

Accessibility coverage

Accessibility coverage measures how accessible resources are for the population. The resources might be available but inconveniently located, therefore hindering physical access. The distance from a health care provider seems to be a very strong factor of accessibility (13;35) (6). Another factor of accessibility related to distance and transportation facilities is time. The travel time to a health facility and the waiting time to see a health professional seem to be well associated with consumers' perception of accessibility of services(26).

There are different opinions about which one is a better measure of accessibility – distance or time(4;21;35;36). The WHO collaborative study on health care utilization suggests using travel time, rather than distance, for the assessment of physical access. The argument for it is that the geography and transportation facilities in different countries vary so much that the comparative measurement of distance to health facilities would be very difficult (45;46).

However, even using time for the measurement of *accessibility coverage* will pose a comparability problem, because the value of time (opportunity cost of time) will be different for

8 It was suggested earlier that in the coverage measure, both numerator and denominator should be the number of people or population units. While this holds true for *accessibility, acceptability*, *contact* and *effective coverage,* for *availability coverage* sometimes using the amount or resources in the numerator might be more practical, especially in terms of data collection.

9 As mentioned in the note 6 above, for *availability coverage* in certain situations "population/population" type measurement could be replaced by a different ratio that might be a more valid measure of *availability coverage.*

different people, and its strength as an access barrier will vary. Therefore for the measurement of *accessibility coverage* by time, calibration methods will be required to ensure comparability not only among countries, which probably is of less interest to a district health manager, but also among different population groups within the country (employment categories, social and family factors, etc.).

Acceptability coverage

Acceptability coverage measures the proportion of people for whom services are acceptable. Even if resources are available and accessible, they may not be used if they are not acceptable to the population.

Acceptability includes affordability, in the first place, as well as non-pecuniary factors such as cultural acceptability, beliefs, religion, gender, type of facility, neighbourhood of facility, etc. (30). Information about *acceptability coverage* is very important for policy-makers in order to better understand some of the predisposing and enabling factors (see page Section 3.2 on access) that affect the use of services.

Because of differences in the nature of the factors affecting *acceptability coverage,* it will be useful to record during the survey the consumers' reasons for low acceptability. To facilitate analysis, the factors of acceptability should be grouped according to their shared characteristics.

Acceptability coverage could be very useful in the assessment of public health programmes, often of a preventive nature. Uptake of the programme by the population may depend largely on their cultural and religious characteristics.

Contact coverage

Contact coverage measures the proportion of the population who have had contact with a health service provider. For those health interventions that require a one-time action, such as the treatment of some soil-transmitted helminth infections that require only one tablet of antihelminthic drug per year, *contact coverage* would be the same as *effective coverage*.

For other interventions, effectiveness of which requires several contacts with a health care provider, *contact coverage* would measure the first contact only. A similar argument may apply to preventive services.

The determinants of *contact coverage* were explained in detail when we talked about utilization. A question may arise: What is the additional value of measuring *contact coverage* in addition to accessibility and acceptability coverage? This question assumes that acceptability of accessible services directly translates to contact. However, as discussed earlier, there are a number of factors, most of them related to the personal health behaviour and decision process, that interplay between acceptability and contact.

Effective coverage

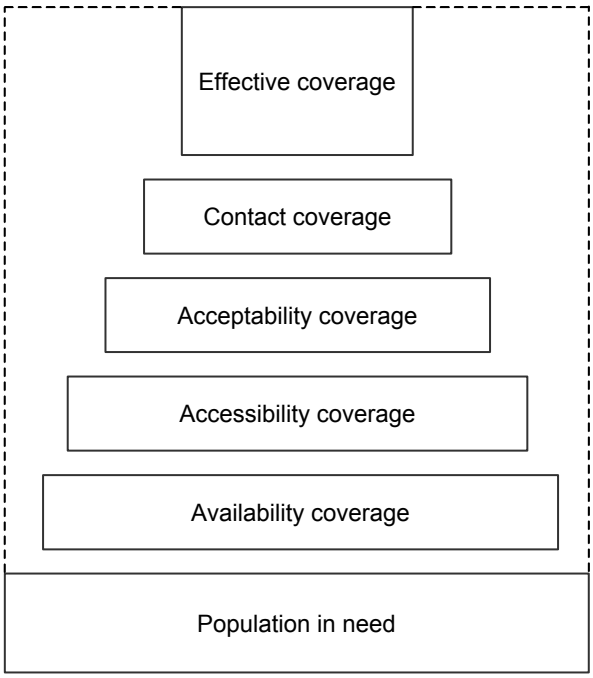
Effective coverage has been discussed above extensively, and a definition was also provided. In the hierarchy of domains of the coverage measures, *effective coverage* is the highest domain. It is itself an intermediate goal.

As mentioned earlier, the key in the measurement of *effective coverage* is to determine what constitutes the effective intervention. In other words, there should be a common agreement about the criteria for counting the occurrence of an intervention as an effective intervention, and how much variance from the established criteria would be acceptable.

For instance, vaccination with DPT vaccine is considered to be effective if all three doses are administered at the correct intervals: the first dose at 6 weeks from birth, and the other two doses spaced at minimum 4-week intervals. This schedule varies from place to place, however. How much variance should we accept? Less than three doses, or doses spaced too closely, or the first dose administered much earlier than 6 weeks, probably should not be counted as effective[10]. But should we still count the vaccination as effective if administered at 6 weeks, 2nd month, and 4th month? These kinds of decisions should be made for each intervention subject to evaluation, and the decisions should be based on scientific evidence from effectiveness studies.

To visualize the domains of effective coverage, we propose a *coverage pyramid* **(Figure 4**). In this pyramid the base represents the population in need. Every additional level represents the domain of coverage and shows how well the process of delivery of a health intervention is carried out. The space outside the *coverage pyramid* represents the size of the population in need who have been left out at each level due to either insufficient resources, poor physical access to services, cultural unacceptability, or insufficient knowledge and motivation to contact a health care provider, comply to the regimen, etc.

Figure 4. Coverage pyramid



(It should be noticed that the relationship between the different domains of coverage are not always hierarchical as presented in the pyramid. This is discussed further in Section 1.3.4.)

(Adapted from (37;57))

A health care manager should try to turn the pyramid into a rectangle, by making policy and management decisions that would ensure the maximum attainable coverage of the population in need with effective services.

10 Probably it is hardly possible to respect the recommended intervals strictly, because of different factors not depending on the performance of immunization programme. For example, if a child is sick with the flue at the time when he has to be vaccinated, the vaccination has to be postponed.

1.3.4 Linking the measurement of health system coverage to management practices

One of the limitations of the previous efforts to measure health system coverage (Section 1.2) was a poor linkage between the measurement exercise and the management practice. It was not quite clear how coverage figures could be interpreted so as to help design effective remedial actions and strategies to improve health system performance. The focus of WHO's renewed attention on health system coverage should be its practical applicability to the management of health system at subnational and national levels.

Having discussed the conceptual framework of coverage, a question arises naturally: How can this framework be used and for what?

The framework of the measurement of coverage presented earlier can be used for personal curative or preventive health services, as well as nonpersonal public health interventions. But depending on the type of an intervention, different approaches must be taken to identify and obtain valid, reliable and comparable numerator and denominator figures[11]. It is obvious that sometimes external scientific evidence has to be applied to the survey instrument in order to reduce the noise in the numerator and denominator. This can be done by adjusting them with a positive predictive value of the instrument.

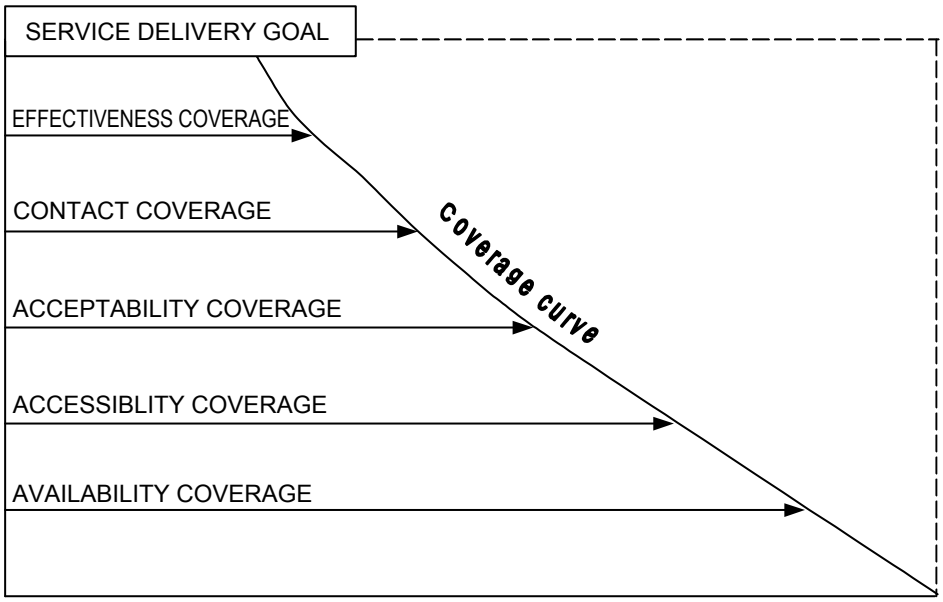
Coverage can be measured for a specific health intervention individually, in order to evaluate a vertical health programme. However, we suggest using *effective coverage* as a measure of the performance of the subnational health service delivery system. For this, interventions must be selected that serve as good tracers for assessing the health service delivery function. The selection of interventions can be based on different criteria, such as ability of the intervention to produce a significant health gain in the population, availability of valid and reliable tools for collecting unbiased comparable figures, etc. Much discussion is required on the basis and criteria for the selection of tracer interventions[12].

This framework is also quite useful for a step-down analysis of the health service delivery process. For example, Tanahashi suggests using the analysis of the slope of the coverage curve (Fig. 5) to identify bottlenecks in health service delivery (37). The domains of the coverage measure are considered as tracers of the bottlenecks in the service delivery process. The smaller slope of the curve between the domains would mean that a higher proportion of people have been left out of the system before they got closer to *effective coverage*. Tanahashi suggests tracing critical bottlenecks in service delivery by locating small slope sections.

11 The different coverage indicators, data sources and measurement instruments will be discussed further in the following sections and scrutinised during the working group sessions.

12 This topic will be discussed during the working group sessions.

Figure 5. Coverage curve



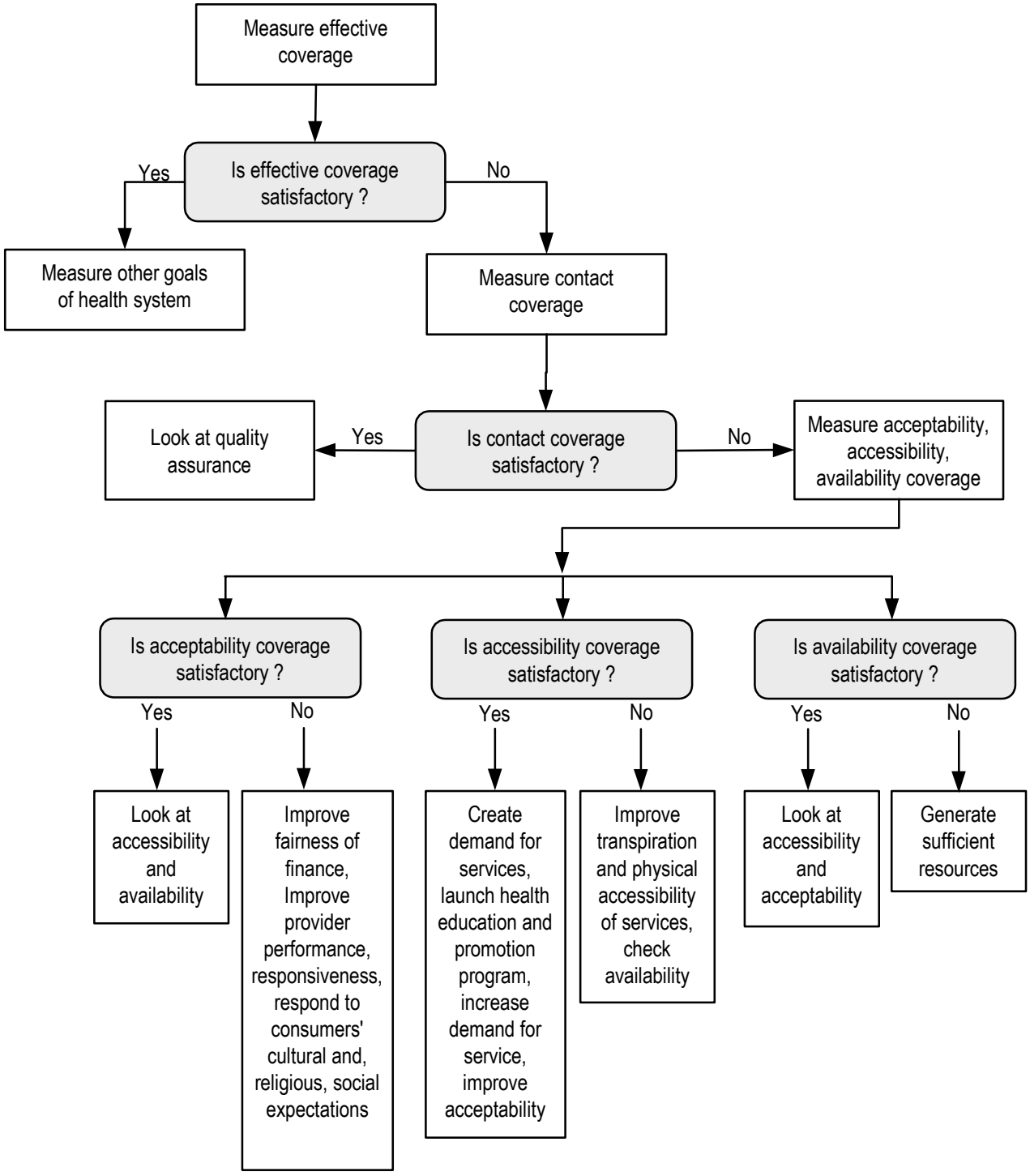
Source: Tanahashi, T, 1978

Based on this analysis, a step-down approach can be applied to the evaluation of the health service delivery function at the district level (Fig. 6**)**. The evaluation should start with the measurement of *effective coverage.* If the level of *effective coverage* is satisfactory, the evaluation process does not have to go further. If the level of *effective coverage* is low, the manager should look at *contact coverage*. If *contact coverage* is satisfactory, the manager must determine the factors that prevent consumers from receiving effective services. The question to be answered is: Why does *contact coverage* not translate into *effective coverage*?

If *contact coverage* is also low, then *acceptability* and *accessibility* coverage should be measured. If the physical accessibility and acceptability of services are high, the manager must look for the problem among those factors that affect individual personal behaviour (for example, through social marketing research methods), in order to find out why acceptable and physically accessible services are not used. If *accessibility coverage* is low, the manager should check whether resources are available, and take necessary action based on the findings.

This is, of course, a hypothetical algorithm of step-down analysis of the system. It should be acknowledged that in the real world the relationship between the different domains of the coverage measure is not strictly straightforward and hierarchical. There many parallel links as well, which depend, among other things, on the type of intervention. The usefulness of this framework is in its ability to serve as a checklist for health policy-makers and managers in the analysis of the health service delivery function through coverage indicators.

Figure 6. Decision-oriented action tool



1.3.5 Key issues

Some of the key issues that ensue from the discussion of the conceptual framework of effective coverage are the following:

- How to select interventions to measure health system coverage
- How to define the effective unit of intervention that is expected to produce a desirable outcome
- How to measure the prevalence of needs in a population so as to result in unbiased figures
- How to ascertain through a survey questionnaire whether a respondent has received effective services
- How to control and reduce stochastic and deterministic biases in data collection in order to improve validity, reliability and comparability of the measurement instrument and the data obtained
- How to interpret different coverage domains and coverage figures.

The following sections will discuss some of these questions based on concrete examples. More in-depth discussion of specific technical issues will take place during the technical consultation.

References

1. Aday LA, Andersen R. A framework for the study of access to medical care. *HEALTH SERVICES RESEARCH*, 1974; 9(3):208-220.
2. Andersen RM, McCutcheon A, Aday LA, Chiu GY, Bell R. Exploring dimensions of access to medical care. *HEALTH SERVICES RESEARCH*, 1983; 18(1):49-74.
3. Beck RG. Economic Class and access to physician services under public medical care insurance. *INTERNATIONAL JOURNAL OF HEALTH SERVICES*, 1973; 3(341).
4. Berk ML, Bernstein AB, Taylor AK. The use and availability of medical care in health manpower shortage areas. *INQUIRY*, 1983; 20(4):369-380.
5. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *SOCIAL SCIENCE & MEDICINE*, 2000; 51(11):1611-1625.
6. De Winter ER. Are we ignoring population density in health planning? The issues of availabity and accessibility. *HEALTH POLICY AND PLANNING*, 1992; 7(2):191-196.
7. Donabedian A. Models for organizing the delivery of personal health services and criteria for evaluating them. *MILBANK MEMORIAL FUND QUARTERLY - HEALTH & SOCIETY*, 1972; 50(103).
8. Donabedian A. *Aspects of Medical Care Administration,* Cambridge, MA, Harvard University Press, 1973.
9. Freeborn DK. Evaluation of the performance of ambulatory care systems: research requirements and opportunities. *MEDICAL CARE*, 1973; 11(68).
10. Goddard M, Smith P. *Equity of access to health care,* York, University of York, Center for Health Economics, 1998.
11. Goddeeris JH, Hogan AJ. *Improving access to health care. What can the states do?,* Kalamazoo, Mich., W.E. Institute for Employment Research, 1992.
12. Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY*, 1999; 28:10-18.
13. Haynes R, Bentham G, Lovett A, Gale S. Effects of distances to hospital and GP surgery on hospital inpatient episodes, controlling for needs and provision. *SOCIAL SCIENCE & MEDICINE*, 1999; 49(3):425-433.
14. Hogarth J. *Glossary of Health Care Terminology,* Copenhagen, WHO/EURO, 1975.
15. Hogarth J. *Glossary of Health Care Terminology,* Copenhagen, WHO/EURO, 1975.
16. Hurtado MP, Swift EK, and Corrigan JP. Envisoning the National Health care Quality Report.2001, Washington, D.C.
17. Institute of Medicine (U.S.) Committee on Monitoring Access to Personal Health Care Services IoM. *Access to health care in America,* Washington, D.C., National Academy Press, 1993.
18. *A guide to conducting health services coverage studies.* Intitute of National Planning Egypt Health Services Research Group, 1981.
19. *Performance measurement and performance management in OECD health systems*. 2001 Labour market and social policy - occasional papers No 47) *(DEELSA/ELSA/WD(2000)8)*.
20. Katz BP, Freud DA, Heck RS et al. Demographic Variation in the Rate of Knee Replacement: A mluti-Year Analysis. *HEALTH SERVICES RESEARCH*, 1996; 31(2):126-140.
21. Kloos H, Chama T, Abemo D, Tsadik KG, Belay S. Utilization of pharmacies and pharmaceutical drugs in Addis Ababa, Ethiopia. *SOCIAL SCIENCE & MEDICINE*, 1986; 22(6):653-672.
22. *Strengthening PHC Services in Africa: An Operation Research Agenda.* Knippenberg R, Ofosu-Amaah S, and Parker D, 1990(Draft).
23. Knippenberg R. The Bamako Initiative: Experiences in Primary Health Care from Benin and Guinea. *CHILDREN IN TROPICS*, 1990; 184/185.
24. Kwast BE. Quality of care in reproductive health programmes: concepts, assessments, barriers and improvements--an overview. *MIDWIFERY*, 1998; 14(2):66-73.
25. Lever P. Management by targets: is coverage an adequate measure for health care? *HEALTH POLICY*, 1994; 14(3):356-359.
26. Melnyk KA. Barriers to care: operationalizing the variable. *NURSING RESEARCH*, 1990; 39(2):108-112.
27. MIller R, Luft H. Managed Care Plan Performance Since 1980: A Literature Analysis. *JOURNAL OF AMERICAN MEDICAL ASSOCIATION*, 1994; 271:1512-1519.
28. Monteith RS, Warren CW, Stanziola E, L'opez UR, Oberle MW. Use of maternal and child health services and immunization coverage in Panama and Guatemala. *BULLETIN OF PAN AMERICAN HEALTH ORGANIZATION*, 1987; 21(1):1-15.
29. *The OECD list of social indicators.* Paris, OECD, 1982 OECD social indicator development programme).

30. Penchansky R., Thomas J.W. The concept of access: definition and relationship to consumer satisfaction. *MEDICAL CARE*, 1981; 19(2):127-140.
31. Phillips K.A., Morrison K.R., Andersen R., Aday L.A. Understanding the context of healthcare utilization: assessing environmental and provider-related variables in the behavioral model of utilization . *HEALTH SERVICES RESEARCH*, 1998; 33(3 Pt 1):571-596.
32. Rosenstock IM. Why people use health services. In: White KL, Frenk J, Ordonez C, Paganini JM, Starfield B, editors. *Health Services Research: An Anthology*. Washington, D.C., Pan American Health Organization, 1992:366-382.
33. Ryland S, Raggers H. Childhood Mortality and Treatment Patterns. *DEMPGRAPHIC HEALTH SURVEYS, COMPARATIVE STUDIES NO*, 1998; 27.
34. Slee V, Slee D. *Health Care Terms,* Saint Paul, MN, Tringa Press, 1996.
35. Stock R. Distance and the utilization of health facilities in rural Nigeria. *SOCIAL SCIENCE & MEDICINE*, 1983; 17(9):563-570.
36. Stoller EP. Patterns of physician utilization by the elderly: a multivariate analysis. *MEDICAL CARE*, 1982; 20(11):1080-1089.
37. Tanahashi T. Health services coverage and its evaluation. *BULLETIN OF THE WORLD HEALTH ORGANIZATION*, 1978; 56(2):295-303.
38. Taylor GD, Aday LA, Andersen R. A social indicator of access to medical care. *JOURNAL OF HEALTH & SOCIAL BEHAVIOR*, 1975; 16(1):39-49.
39. Timmreck T. *Dictionary of health services management,* Owings Mills, MD, National Health Publishing, 1982.
40. Tonglet R, Soron'gane M, Lembo M et al. Evaluation of immunization coverage at local level. *WORLD HEALTH FORUM*, 1993; 14(3):275-281.
41. Victora CG, Bryce J, Fontaine O, Monasch R. Reducing deaths from diarrhoea through oral rehydration therapy. *BULLETIN OF THE WORLD HEALTH ORGANIZATION*, 2000; 78(10):1246-1255.
42. Wennberg JE. Future Directions for Samll Area Variations in Health Services Utilization in Quebec, Canada. *SOCIAL SCIENCE AND MEDICINE*, 1993; 16(7):811-824.
43. Wennberg JE, Barnes WA, Zubkoff M. Professional uncertainty and the problem of supplier induced demand. *SOCIAL SCIENCE & MEDICINE*, 1992; 16(7):811-824.
44. "Primary Health Care" A joint Report by the Director-General of WHO and the Executive Director of UNICEF. In: Geneva-New York, 1978.
45. WHO. *Health care - an international study report of the World Health Organization. International Collaborative Study of Medical Care Utilization,* London, Oxford University Press, 1976.
46. *Health services concepts and information for national planning and management experiences based on the WHO/International Collaborative Study of Medical Care Utilization.* Geneva, WHO, 1977 (White KL, Anderson DO, Kalimo E, Kleczkowski BM, Purola Tapani, and Vukmanovich C).
47. *Measurement of Coverage, Effectiveness and Efficiency of Different Patterns of Health Care.* WHO, 1978(SHS/78/1.
48. *Background paper on the measurement of coverage, effectiveness and efficiency of different patterns of health care.* WHO, 1979(SHS/79.4).
49. *Developing indicators for monitoring progress towards health for all by the year 2000.* 1981 (Health for All).
50. *National assessments of health care coverage and of its effectivenss and efficiency.* WHO, 1983(SHS/83.7).
51. WHO. The World Health Report 2000.2000, Geneva
52. Background Paper. Regional Consultation on Health System Performance Assessment. In:Harare, WHO/AFRO, 2001.
53. Critical Issues in health System Performance Assessment. Regional Consultation on Health System Performance Assessment. In: Washington D.C: 2001.
54. *Design of multi-country evaluation of IMCI programmes.* WHO/Department of Child and Adolescent Health, 2001(FCH/CAH/00.33;
55. *Immunization Services Assessment Guidelines.* WHO/EIP, 2000.
56. World Bank. *Better health in Africa,* Washington, D.C., The World Bank, 1994.
57. World Bank. Technical Note No. 5: *Assessing health sector performance*, 2001.

2. Effective coverage in practice: indicators and measurement instruments

2.1 Background

The purpose of reviewing coverage indicators in this work is to better understand how coverage is measured within different personal and nonpersonal interventions and how well these indicators reflect the definition of effective coverage previously proposed. The objective of this analysis is to discuss the major issues that must be addressed in measuring effective coverage, based on practical examples from different health programmes.

The interventions for assessing health system performance in terms of service delivery are selected according to their ability to serve as tracers for the measurement of health system coverage. Although the intention is to look at a variety of interventions within the broad spectrum of care, the selection of interventions does not necessarily fall on cost-effectiveness criteria or any other criteria expressed in health gains or reduced illness per unit of input.

Based on the definition of effective coverage in section 1, programme-specific indicators for effective coverage should follow the same model:

- A numerator that represents those who receive the minimum amount of process and input required for an intervention to be effective (for example, three doses of diphtheria/pertussis/tetanus combined antigens)
- A denominator that can indicate a disease or risk factor prevalence; in some cases, it can also be represented by an incidence rate (tuberculosis).

For example, in the case of DTP immunization, the numerator is the number of children who have received three doses of DTP combined antigen according to a recommended schedule, which actually varies from country to country; the denominator is either the birth cohort or the population of infants surviving until 12 months of age, depending on the data collection method used.

Given the definition of effective coverage, three important sets of issues arise for any given intervention:

- What the indicators truly measure (validity)
- Interpretation of results
- The methods of measurement.

2.1.1 Issues related to the interpretation of what the indicator truly measures

Construction of an indicator for effective coverage requires a clear understanding of what one wants to measure. The decision about what to measure should be sensitive to the health problem. If the effective unit of an intervention for a particular health problem is a discrete, time-limited (or even one-time) event, then the indicator of effective coverage should measure the occurrence of this unit. If the effective intervention for a particular health problem is a long-term process, successful completion of which is difficult to measure by a cross-sectional survey in a general population, then some pragmatic modification can be applied to the definition of effective coverage indicator to facilitate measurement. For example, in the case of tuberculosis the ideal indicator for effective coverage would be the proportion of people with TB who completed the full course of treatment. However, since the precise estimation of tuberculosis prevalence or incidence is difficult, we can measure the proportion of the patients with a definite diagnosis who get a full course of treatment.

2.1.2 The interpretation of results

The purpose of this analysis is to discuss how effective coverage and other domains of coverage measurement can be used to assess health system delivery. The essential analytical step is to propose or select an indicator that best captures effective coverage of an intervention and to determine whether other coverage domains can be captured as well.

Coverage indicators as defined in this paper estimate effective coverage but teach very little about the four other domains of coverage. A low result in an indicator of effective coverage would tell the observer that the minimum amount of process and input of an intervention is insufficient to effectively reach – or cover – everyone in need, but does not tell anything about which domain of coverage is the most affected.

The combination of a low effective coverage and a high drop-out rate shows that there is a problem in translating contact into effective coverage. It does not tell anything about any other domain. If contact coverage is high, one can assume that the services are available, that the client has accepted it and has access to it. A low effective coverage and a low drop-out rate would show that contact has not been sufficiently made, calling for in-depth analysis of availability, accessibility or acceptability.

The case of immunization illustrates this concept very well. DPT3 is effective coverage, as defined above. DTP2 and DTP1 show contact. If the figures for the three are close to each other, the issue lies more in the availability of vaccine and immunization services, acceptability of immunization or accessibility to vaccine and services.

An important issue in designing a global health system coverage survey is to focus analytical capacity on capturing these differences on a global scale. Local determinants of various domains of coverage can be left to the interpretation of locally performed community-based surveys.

For example, in sub-Saharan Africa a low result in DTP3 can be driven by a low perception of risk on the part of the mother, given that the morbidity or mortality of these vaccine-preventable diseases is low in any case. This behavioural determinant of coverage (utilization) may play a major role in the measurement of the acceptability coverage domain that only local household surveys can capture, by comparing these results to a high measles immunization rate – since measles is a much more lethal infection and is perceived as such.

2.1.3 The methods of measurement

This part of the analysis focuses on measurement issues, in particular for the problems linked with the collection and analysis of numerator and denominator data for effective coverage indicators. Both the numerator and the denominator in effective coverage can be affected by different deterministic and stochastic biases inherent in the measurement tools. These biases should be controlled or adjusted-for in order to achieve a desired level of validity, reliability and comparability.

On the numerator side, issues of validity arise with the method used to confirm that the effective unit (amount of process and input) of an intervention has occurred or has been delivered to the one who needs it. The confirmation method can be invalid if the question used for collecting this information is ambiguous or biased.

For example, assessing whether a child has been truly immunized requires evidence documented by an immunization card or verified by the mother. In the first case, accepting an undated stamp on a card as evidence of a DTP dose may overestimate the valid immunization coverage. In the second case, rejecting a mother's recollection may

underestimate results (recall bias). Corrective methods would consist of adjusting for the mother's recall as well as the probability of valid immunization among cardholders.

Sometimes it is difficult to confirm true occurrence of interventions by asking questions. In BCG vaccination, the mother's recall bias can be adjusted by looking at the objective sign of occurrence of intervention. In this case, this would be to check the mother's recall of BCG injection against the presence of a scar.

Denominator errors of validity for coverage indicators are those that are well known when it comes to measuring prevalence or incidence rates. High blood pressure prevalence can be measured by means of population-based surveys. Then the validity of the questions attempting to measure the prevalence of high blood pressure can be adjusted by an objective reference test: measurement of BP in a representative subsample of the surveyed population. Knowing the positive predictive value (PPV) of the reference test would help to estimate with a certain degree of confidence the true prevalence of high blood pressure in the population.

Other difficulties arise with the effect of high case fatality rate or the inability for an event to be detected by asking questions. In some cases, events are difficult to recall and recognize, as with tuberculosis. Here there are very few methods for cost-effectively confirming a sample of positive answers. Effective coverage of tuberculosis personal interventions will always be inaccurate, because the denominator – the prevalence of TB – is difficult to measure.

Reliability errors affect both numerator and denominator and result from inconsistency in the procedures and questions used as a measurement instrument. BP results are very sensitive to the procedure for measuring BP.

Similarly, comparability errors reflect individual variation of the subject of a measurement event and therefore require calibration for systematic differences. At the population level, comparability can be affected by underlined population characteristics. For instance, two population groups with different prevalence of obesity will show different average BP making them non-comparable for BP intervention. The aforementioned issues must be addressed in designing coverage indicators and relevant measurement instruments. The following steps serve as a checklist in this process:

1. Definition of the intervention
2. Selection of coverage indicators and construction of an indicator for effective coverage
3. Definition of the numerator (amount of input and process) and the denominator (population in need) of one coverage indicator used for that intervention
4. Analysis of the numerator's validity: confirmation method, data collection issues, adjustment, recall bias
5. Analysis of the denominator's validity: issues in prevalence, data collection issues, adjustment, recall bias
6. Analysis of reliability of the numerators and denominators: consistency in questions (variation in questions)
7. Analysis of the comparability of the numerators and denominators: calibration for interpretation of a question (variation in respondents)
8. Interpretation of low or high result of effective coverage, interpretation of drop-out rate and conclusion on coverage measurement domains.

2.2 Immunization

Immunization coverage indicators are straightforward and easy to understand within the effective coverage framework proposed in this paper. The interest here is to pick up on these indicators and make an in-depth analysis of the validity, reliability, and comparability issues of their measurement instruments.

2.2.1 Coverage indicators

"Coverage in immunization" measures the proportion (rate) of the population that has received immunization service (regardless of whether this service results in a biologically effective immune status). In the case of "Coverage in childhood immunization" the coverage measures the proportion of the 1-year-old population who received full immunization according to recommended schedule. "Coverage" with the first dose of a particular antigen corresponds to contact coverage. Complete vaccination with all required doses will represent effective coverage (three doses of DTP).

DTP3 is one of the six immunization coverage indicators used in the Expanded Programme on Immunization (EPI), which are the following:

- BCG coverage
- POV3 coverage (polio)
- DTP3 coverage (diphtheria, total tetanus, pertussis)
- MCV coverage (measles)
- TT2plus coverage (neonatal tetanus)
- YFV (yellow fever vaccine coverage).

Each indicator has a definition and operational context or purpose. For example, DTP3, MCV and POV3 have different target populations than TT2plus and YFV; BCG is subject to a variety of national health policies. The validity of POV3 coverage can be compromised by the interference of eradication (NID) campaigns in the routine reporting process. MCV coverage is also less preferable as a single indicator than DTP3 because it is administered at different schedules, usually later than other antigens, and requires only one contact with care providers. TT2 coverage has a validity problem, as it measures not only protection of a child from neonatal tetanus, but also quality of antenatal care, which does not make it very specific to effective coverage of immunization.

Based on these considerations and on consultations with EPI, we propose to use DTP3 as the best proxy for measuring immunization coverage as a whole. Measurement of coverage by all three doses of DTP vaccines allows the analysis of both the contact and the effective coverage.

2.2.2 Measurement issues

Two main mechanisms exist that measure immunization coverage worldwide. These are country-generated statistics (service data), and general community-based household surveys.

- There are three widely administered surveys: the USAID's Demographic and Health Surveys (DHS), the Expanded Programme on Immunization Cluster Survey (EPI Cluster) and UNICEF's Multiple Indicator Surveys (MICS).
- Service data are based on records of health providers and health service facilities. Typically countries are supposed to report annually to the World Health Organization and UNICEF their immunization coverage estimates based on these service statistics.

Service data carry a significant amount of of noise in the numerator. For example, vaccination can be recorded at a facility that does not correspond to the civil registration area of the patient. This would result in inappropriate matching of numerator and denominator.

Another common issue with service data is the inconsistent use of either birth cohort or surviving infant in the denominator among countries.

The EPI Cluster Survey, which was designed in the 1970s, is an affordable but less precise instrument. A wide confidence interval makes it imprecise for detecting small changes in immunization coverage and therefore compromises its validity. This is particularly true for the measurement of coverage today, when the figures at global level are higher than before.

DHS is considered as the gold standard in the measurement of the immunization. It considers all children born in the last five years and allows both period and cohort analysis. In comparison, the EPI Cluster Survey considers only the 0-1-year cohort of living children who were born during the year preceding the survey.

MICS uses the same criteria as DHS in terms of design, but does not capture the children who were immunized but died between ages 1 and 5.

All the survey methods share the same issues, which are related to the method of confirmation of vaccination (card, card and history, or history only) and the criteria for valid vaccination (existence of documented proof of valid schedule).

Depending on the method of validation used in a particular survey, immunization coverage figures may be overreported or underreported. Crude immunization figures must be adjusted for recall bias as well as probability of the vaccination's being wrongly recorded on the card. These biases and the relevant corrective methods are described in the Annex.

Major biases are summarized for each measurement method in the following table.

Table 1. Comparative assessment of errors in measurement instruments for immunization coverage (DTP3)

Method	Stochastic errors	Systematic (deterministic) errors
Routine health report	• Missing data (underreporting) • Data entry mistakes	• Over-reporting of >1-year-old children • Rounding up • Incentive-driven over-reporting • Under-reporting (no reporting compliance) from private sector • Definition of population cohort (birth vs. survival) • Aggregation of non-routine data under officially reported routine data
Cluster sampling survey	• Sampling	• Individual clusters cannot be compared due to the statistics underlining the method • No accurate household list • Definition of DTP full coverage status (card, card+history, crude/valid, 12-23 months, under 12months) • Confidence interval too wide • Not accurate for high true population immunization rate • Recall bias • Cheating, incentives (GAVI) • Randomization not always respected in practice • Zero-dose definition
DHS **MICS**	• Sampling	• Questionnaire length, fatigue response factor • Definition of DTP full coverage status (card, card+history, crude/valid, 12-23 months under 12months)

2.2.3 Main issues

•Issues inherent in service reporting (incentive, missing data, numerator/denominator matching)
•Inconsistency in denominators (surviving infants vs. birth cohort)
•Inconsistency in method of validation
•Recall bias
•Capturing private sector coverage
•Cross-country variation of immunization schedules.

2.3 Tuberculosis control

Tuberculosis control programmes have a long history of associating treatment or diagnosis outcomes (the intervention provided) to incidence data (the population in need of that intervention). The purpose is to assess trends and performance in order for the health system to detect persons with TB and offer them the best possible case management. The interest of discussing TB control lies in the diversity of both sets of outcome items for which data are available: patients who are cured or not; those who complete the full course of treatment or not; those who transfer; and those who die from causes other than TB. This therefore allows for many interpretations about assessing coverage of TB control activities.

2.3.1 Coverage indicators in TB control

TB control programmes uses a number of coverage indicators related to case detection, case management and the implementation of DOTS programmes at national level. The following table summarizes some of them.

Table 1. Tuberculosis control programme indicators

Coverage	*At national level* : Number and % of districts with DOTS strategy / Population living in areas covered by DOTS strategy	Surrogate for population coverage by DOTS strategy
	At district : Number of health facilities with DOTS services / total health facilities	Particularly public facilities, but may include NGOs and/or private sector, if needed
Detection	*At health facility and district levels:*	
	Number of symptomatic patients detected with cough >2 –3 weeks / total outpatients	Needed for resources planning
	Number of smear-positive cases detected / number of symptomatic patients detected with cough >2–3 weeks	**Needed for resources planning and is surrogate for TB frequency in community**
	At region / state / national level:	Under the same conditions, the trend over time can indicate the trend of TB
	Number of total new cases detected / number of total estimated new cases	To be compared with WHO target : 70%
Treatment outcome	*At all levels:*	
	Cohort analysis for smear-positive new cases:	Should be done on all patients registered in specified period (i.e. on quarterly basis)
	% cure, % treatment completion, % failure, % death, % default,	To be compared with WHO target: 85% treatment success rate
	% transfer out	Should be done for each category
	% treatment success = % cure + % treatment completion	Early surrogate for treatment success
	Cohort analysis of re-treatment of smear-positive cases by category: relapses, failures, treatment interruption and chronic / MDR cases	
	Sputum conversion rate: % of TB patients who are smear-negative at 2nd-3rd month of anti-TB treatment	

At the national level DOTS coverage measures the proportion of health administrative units implementing the strategy. At subnational level, programme coverage focuses more on the

coverage of case detection among outpatients presenting pulmonary symptoms (long-lasting cough), giving an indication of the local capacity to detect cases in a sensitive and specific way.

Another type of coverage indicator used in TB control programmes is the percentage of health facilities with TB-control capacities in terms of case detection, availability of drugs and human resources for diagnosis and case follow-up.

Coverage by TB health facilities measures the proportion of facilities satisfying the criteria of TB-control capacities. The denominator captures facilities of all levels, because TB control should not be broken down by care provider level. The health facility coverage indicator measures availability coverage but does not teach about people's behaviour towards TB.

TB control programmes also pay attention to accessibility coverage for both treatment and diagnosis.

As previously mentioned, according to our definition, effective coverage of TB programmes could be represented by the proportion of people with TB who complete the full course of treatment. The denominator can be represented by either the annual number of new smear-positive cases or the estimated incidence of TB in the population. In the latter case the incidence is estimated and adjusted by the notification rate. The methodology for such estimation is described in the *World TB report 2001*. In the case of Africa, for example, estimated incidence rates for 1997 of 19 countries have been adjusted according to trends in the notification rate. The resulting estimates for 1999 are closely correlated with the estimated prevalence of HIV in adults. The corresponding linear regression has been used as a calibration curve to estimate the TB incidence rate for 41 African countries.

2.3.2 The measurement issues

Both numerators and denominators are reported at service level on Tuberculosis Treatment Cards. No survey methods are used in TB programmes to check the validity of data on treatment outcome and diagnosis capacity. There have been examples only of surveys used to assess TB prevalence (i.e. in China and India). These surveys have shown some relationships between incidence and prevalence, in particular that prevalence approaches incidence when TB control activities are effective. The use of survey methods to monitor TB intervention outcomes is irrelevant for the following reasons:

- TB treatment is not a one-time event, unlike immunization. The outcome occurs only after a full course of treatment that can vary in duration.
- TB is not randomly distributed among households, which poses challenges for sampling.
- The cards are not kept at home, which excludes the possibility of measuring case management outcomes at household level.

  Basically, TB data are generated by the notification system for diagnosis and case follow-up for the treatment course. Biases occur when patients leave public providers and pursue treatment within the private sector. These patients have been registered but not followed up, so the type of therapeutic outcome cannot be entered properly in the numerator. In some instances cases are not even registered, because they are diagnosed at private provider facilities.

2.3.3 Main issues

- Multiple definitions of treatment outcomes make the selection of a numerator for effective coverage difficult:

  - Cases effectively covered are those who are cured (smear-negative), plus those who completed treatment (clinically cured, with full therapy compliance), plus those who died during the treatment course from causes other than TB.
  - This last category is taken into account for effective coverage because death from non-TB causes does not preclude positive or negative response or compliance with the treatment.

- The difficulty of measuring incidence of TB, which could be the most inclusive denominator for quantifying people in need
- The irregular reporting of case registration and treatment outcomes from private facilities.

References

1. Dye et al. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. *Journal of the American Medical Association,* 1999, 282(7):677–86.
2. Dye et al. Prospect for worldwide tuberculosis control under the WHO DOTS strategy. *The Lancet,* 1998, 352:1886–91.
3. Maher D, Nunn P. Evaluation and determinants of outcome of tuberculosis treatment. *Bulletin of the World Health Organization,* 1998, 76(3):307–8.
4. Raviglione MC et al. Assessment of worldwide tuberculosis control. *The Lancet,* 1997, 350:624–9.
5. World Health Organization. *Global tuberculosis control: WHO report 2001*. *WHO/CDS, 2001.*
6. World Health Organization. *WHO Tuberculosis Programme: framework for effective tuberculosis control*. Geneva, World Health Organization, 1994 (unpublished document WHO/TB/94.179, 1994).

2.4 HIV/AIDS

There is a dearth of publications that directly address the issue of coverage measurement, particularly when compared to documentation on programme indicators. Major issues involved in the definition, selection and application of indicators for HIV programmes are well documented. They include what each indicator is used for, how scientifically robust it is, how understandable it is, data sources for its computation (accessible from), justification for selection, interpretation of changes in composite indicators and the need for disaggregation. What are not dealt with at great length are indicators for estimating the coverage of these programmes. The one exception is the area of prevention of mother-to-child transmission (MCT) of HIV.

2.4.1 Coverage indicators for MCT

UNICEF, UNAIDS and WHO have carried out joint work on effective coverage for prevention of mother-to-child transmission of HIV (PMTCT). In keeping with Knippenberg and Tanahashi, coverage is viewed in terms of five domains (availability, accessibility, utilization, continuity and quality). Accessibility is examined in both physical and financial terms, while utilization is limited to initial use. With regard to the model of coverage developed in the background document, utilization in the MCT model corresponds to *contact coverage*, while quality is most closely related to *acceptability*.

The MCT package is composed of four groups of interventions: primary HIV prevention and voluntary counselling and testing (VCT), obstetrical care, postpartum child care, and mother and family care.

13 Table 1. Coverage domains and tracers of MCT

Coverage domain	Components/tracers of MCT
Availability	Primary prevention, VCT, ARV, long-term support
Access	(nothing suggested here)
Utilization / *Contact coverage*	Family planning, ANC, VCT, infant feeding, ARV
Continuity	Family planning, VCT, ARV, safe delivery
Quality / *acceptability coverage*	VCT, infant feeding, ARV, long-term support

Source: UNICEF/UNAIDS/WHO. *Local monitoring and evaluation of the integrated prevention of mother to child HIV transmission in low-income countries*. Draft (03/2000).

2.4.2 Measurement methods

Table 1 also shows data sources for numerator and denominator values of each indicator. These are all routinely collected data. In general, data for input and output indicators can be collated centrally from regular health system reporting systems. Data for outcomes and impact indicators, on the other hand, must be collected through health-facility or population-based surveys, such as:

- Censuses
- Demographic and health surveys
- Vital and other data registries
- Health service documents (hospital registers – delivery, theatre book, discharge register, etc.)
- Structured observations
- Service exit interviews.

The cost and difficulty of data collection tend to increase as indicators shift from input through output and effect to impact. This bodes well for coverage indicators, which focus on input and process, such as those for MCT.

However, there are also problems with the exact number in the target population (denominator). As interventions are targeted at subpopulations, difficulties arise in the denominator of the coverage indicator. This is illustrated by a study in Zimbabwe, which examined coverage of a community-based home care programme for AIDS patients in the city of Mutare. Five different models were used to obtain the denominator.

13 A matrix, with the five coverage domains on the left and interventions on the top, is provided for the first three groups of interventions (Tables 3, 4 and 5 of the UNICEF/UNAIDS/WHO document). Tracers of MCT interventions are identified for four of the five coverage domains (see above), and a minimum set of effective coverage indicators is suggested (Table 6 of the document). The indicators are given in Table 1 above, with an additional column for coverage domains. Specific indicators dealing with *access* are not suggested.

Table 2. Effective coverage indicators for MCT

Coverage domain	Indicator	Numerator of definition	Numerator data source	Denominator definition	Denominator data source
Availability	Availability of condoms	# condoms available in related services for the period considered	Stock forms	# clients attending related services for the period considered	Service registers
	Availability of STD drugs	# days with no shortage of STD drugs		# days in period considered	Calendar
	Availability of HIV test	# days with no shortage of test kits (at least 2 types of kits)			
	Availability of ARV drugs	# days without shortage of ARV drugs			
	Referral to support network	# women who delivered during the period considered referred for long-term care	MCH forms Registers	# HIV+ women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	MCH cards Registers
Utilization	Use of family planning	# women 15-49 using modern FP method in the area covered	FP services	# women 15-49 living in the area covered	Census Fertility rate
	Use of ANC	# pregnant women with at lest one ANC visit	Service registers	# pregnant women in the area covered	
	Use of HIV pre-test counselling	# women who delivered during the period considered who received pre-test counselling	MCH forms Registers	# women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	MCH cards Registers
	Use of HIV testing	# women who delivered during the period considered who were tested for HIV			
	Infant feeding counselling	# women who delivered during the period considered who received infant feeding counselling prenatally			
	Initial use of ARV drugs	# HIV+ women who delivered during the period considered who initiated ARV treatment		# HIV+ women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	MCH forms Registers

Coverage domain	Indicator	Numerator of definition	Numerator data source	Denominator definition	Denominator data source
Continuity	Use of HIV post test counselling	# women who delivered during the period considered who received post test counselling	MCH forms Registers	# women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	MCH cards Registers
	Family planning counselling	# women who delivered during the period considered who received FP counselling before delivery			
	Continuity of ARV	# HIV+ women who delivered during the period considered who received ARV during at least 2 weeks before delivery		# HIV+ women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	MCH forms Registers
	ARV during labour	# HIV+ women who delivered during the period considered who received intra-partum dose of ARV			
	Exclusive BF at 3 months for HIV+	# HIV+ women who delivered during the period considered who exclusively BF up to 3 months			
	Use of FP by HIV+	# HIV+ women who delivered during the period considered who use a modern FP method after delivery			
Quality	Use of Iron/folic acid	# women who delivered during the period considered who received iron and folates	MCH forms Stock forms	# women with at least one contact with MCH (either pre, per or post partum) whose date of delivery corresponds to the period considered	MCH cards Registers
	Partners tested for HIV	# women who delivered during the period considered whose partners were tested			
	Episiotomy rate	# HIV+ women who delivered during the period considered who delivered without episiotomy			
	Exclusive BF at 3 months for HIV-	# HIV- and untested women who delivered during the period considered who exclusively BF up to 3 months	MCH cards	# HIV- and untested women with at least one contact with MCH (either pre, per or post partum) whose date of delivery falls within the period considered	

Source: UNICEF/UNAIDS/WHO. *Local monitoring and evaluation of the integrated prevention of mother to child HIV transmission in low-income countries*. Draft (03/2000).

•programme goal (target set for the period)

•excess mortality statistics (all attributed to AIDS)

•registry data combined with a survey carried out as part of the study

•national census figures, seroprevalence rates and estimates of home care needs from the industry

•death rates in the surveyed population, extrapolated to the entire target area using census figures.

With a fixed numerator (the number of clients served), coverage figures ranged from a low of 2.5% to a high of 23%.

On the other hand, where both the numerator and denominator are known with a high degree of accuracy, coverage estimates can be tightly bound. Routine neonatal metabolic screening in children born to women known to be infected with HIV-1 in inner London gives figures of 96.4% (94.6% in infants of African origin and 100% for whites). Similar figures are reported for the general population in the same area.

Interestingly, in addition to their intrinsic value for measurement of coverage, coverage indicators may be used for measuring effectiveness. Due to difficulties inherent in the methodologies used, it is impractical to measure effectiveness directly. One model measures individual effectiveness through a case control study, and estimates community effectiveness from this value using a coverage indicator.

2.4.3 Coverage indicators for other HIV/AIDS interventions

Indicators for monitoring and evaluation of programmes measure attainment in one or more programme domains: input, process, output, effects, outcome and impact. The domains of coverage, on the other hand, are concerned primarily with issues of input and process: the effort made by the health system to provide health services for the population. The benefit the user derives from the services offered depends on other factors, for instance proximate determinants, which are more readily influenced by the user.

It would appear, therefore, that indicators for programme evaluation that measure aspects of input and process are suitable to measuring coverage as well, while those that measure other programme domains are not suitable as indicators of coverage. HIV/AIDS programmes are usually evaluated using 14 sets of indicators. Of these, six sets deal with inputs and process, and could be proposed for measurement of coverage (policy, condom availability and quality, voluntary counselling and testing, mother-to-child transmission, blood safety/nosocomial transmission, care and support), while three out of four indicators within one set (STI care and prevention) may be suitable, and none of the remaining seven sets (stigma and discrimination, knowledge, sexual negotiation and attitudes, sexual behaviour, young people's sexual behaviour, injection drug use, health and social impact) appear to be appropriate for this purpose.

2.4.4 Main issues

- One of the biggest problems is that the set of coverage indicators proposed for MCT (the only component of HIV/AIDS for which coverage indicators have been suggested) is still untested. Actual measurement is likely to be fraught with difficulties. Quality of counselling, for example, plays a crucial role in VCT. However, measuring the quality of counselling is particularly difficult, due to the need for confidentiality. Mystery patients and exit interviews of clients have been suggested as possibilities.

- Recall bias is likely to become more significant as the recall period increases. This raises questions with periods of 12 months, as suggested for HIV/AIDS interventions. In addition to the usual problems of reliability and validity of survey data, surveys of sexual behaviour pose additional problems. Other sources of errors are non-participation bias, subjectivity of the answer categories and social desirability of the answers.
- Collecting data on, for example, the number of HIV-positive mothers who exclusively breastfeed for up to three months is extremely difficult, and may be virtually impossible in many settings.
- As to attitudes, the problem is that while qualitative methods are more appropriate to measuring beliefs than quantitative methods, the former are considerably weaker in establishing trends over time. Then there is the problem of how to partition behaviour into various categories of risk. Apparently women continue to underreport sexual encounters with non-regular partners in standardized surveys. In general, those most at risk tend to be those outside of conventional frameworks through which interventions are provided – young people outside of school, for example.
- Finally, in developing country environments there are also greater requirements for high quality in survey design and execution.

References

1. Boerma T, Pisani E, Schwartlander B, Mertens T, A framework for the evaluation of national AIDS porgrammes, MEASURE Evaluation January 2000
2. Cleason M, Griffin C, Johnson T, McLachlan M, Soucat A, Wagstaff A, Yazbeck A, Technical Note 5A: Assessing health sector performance, available at http://www.worldbank.org/
3. Dare OO, Cleland JG., Reliability and validity of survey data on sexual behaviour, Health Transit Rev 1994;4 Suppl:93-110
4. Drew RS, Mgombane G, Nyaruwa T and Foster G, Estimating coverage of a community-based home care programme, Tropical Doctor, 1997;27:210-14
5. Gibb DM, Faulknall W, Nokes L, Appleby S, Holland FJ, Berry T Ades AE, Coverage of routine neonatal metabolic screening in children born to women known to be infected with HIV-1, Commun Dis Rep CDR Rev 1995 Jul 21;5(8):R123-4
6. James NJ, Bignell CJ, Gillies PA, The reliability of self-reported sexual behaviour, AIDS 1991 Mar;5(3):333-6
7. Kauth MR, St Lawrence JS, Kelly JA, Reliability of retrospective assessments of sexual HIV risk behavior: a comparison of biweekly, three-month, and twelve-month self-reports, AIDS Educ Prev 1991 Fall;3(3):207-14
8. Knippenberg R, 1986, «A methodology for monitoring and evaluation of coverage with essential primary health care interventions», PH.D. dissertation, the J>ohns Hopkins University , Baltimore MD.
9. Lengeler C, Snow RW, From efficacy to effectiveness: insecticide-treated bednets in Africa, Bull World Health Organ 1996;74(3):325-32
10. Saltzman SP, Stoddard AM, McCusker J, Moon MW, Mayer KH., Reliability of self-reported sexual behavior risk factors for HIV infection in homosexual men, Public Health Rep 1987 Nov-Dec;102(6):692-7
11. Strickler H, Hoover DR, Dersimonian R, Problems in interpreting HIV sentinel seroprevalence studies, Ann Epidemiol 1995 Nov;5(6):447-54
12. Tanahashi T, Health service coverage and its evaluation, Bull World Health Organ 1978;56(2):295-303
13. UNAIDS, National AIDS programmes: a guide to monitoring and evaluation, June 2000
14. UNAIDS / WHO Working group on global HIV/AIDS and STI surveillance, Second generation surveillance for HIV: compilation of basic materials, January 2001
15. UNICEF/UNAIDS/WHO, Local Monitoring and Evaluation of the Integrated Prevention of Mother to Child HIV Transmission in Low-Income countries, draft, March 2000.

2.5 Diabetes mellitus

Chronic diseases, such as diabetes mellitus, that rely extensively on health services for lifetime disease management may provide coverage indicators that can offer valuable information about the performance of the health system in the wider context, especially in relation to sustainability of personal services. The concern with diabetes control is mostly with preventing (or postponing) late complications in persons with the disease. Persons diagnosed with diabetes need continuous and frequent contact with the health system for the rest of their lives in order to implement preventive strategies for late complications.

2.5.1 Coverage indicators

According to our definition of effective coverage, the denominator is the persons who need diabetes services. This includes the usually substantial number of undiagnosed cases in the community (about 50% worldwide) and all the individuals ever diagnosed, because the disease cannot be cured and requires lifelong control. This fact brings about challenges to the health system that are not present in most other critical interventions, aside from those for other major chronic noncommunicable diseases (for example, hypertension). Diabetes clusters with other noncommunicable diseases of public health importance (in terms of risk factors, inputs needed and treatment goals), but compared to those other diseases it has more rigorous measurement tools available.

Several coverage indicators may be used to assess the performance of diabetes control programmes. One such indicator is the *proportion of diagnosed cases in the population*. The denominator is the population with diabetes mellitus, as established with either the two-hour glucose tolerance test or fasting glucose blood test, through community-based prevalence surveys. The numerator consists of those who respond that they were previously informed by a doctor that they have diabetes.

This indicator may tell something about the health system's competence to detect individuals with diabetes. In terms of effective coverage, it may be seen as a measure of how much of the population in need of diagnosis is really covered with diagnostic services through the health system. It does not, however, provide specific information about the availability coverage (diagnostic facilities), the acceptability (public awareness and willingness to use services) or physicians' knowledge and skills in case finding.

A second indicator measured by community-based prevalence surveys is the *proportion of diagnosed persons receiving treatment*. All individuals who state that they were previously diagnosed are questioned about their treatment (insulin, oral drugs, herbal drugs). The principal treatment for diabetes in the formal health system is insulin, oral antidiabetic drugs, diet and exercise advice. In many developing countries, traditional remedies are used as well. The survey allows for measuring the ratio of cases treated with traditional remedies over cases treated with "modern" medicines. This may be useful as an acceptability coverage indicator.

Other coverage indicators are based on registered patient data and established through health facility surveys. The *frequency of consultations*, the *availability of insulin* or the *proportion of patients tested for complications* are examples of indicators relevant to diabetes control that provide information about different domains of coverage. They reflect only the registered patient population, however. A true estimate of coverage depends on the availability and reliability of figures about the prevalence and proportion of diagnosed persons in the population.

An indicator that reveals important information about how effectively the population is covered is the *degree of metabolic control in patients*. Metabolic control is directly related to incidence

and progression of the long-term complications and is therefore a key measure in diabetes control. The denominator is again the population of registered patients, but this is useful if considered as a proxy for quality, or sustainability in personal services or a proxy for tertiary prevention in diabetes.

2.5.2 Measurement methods

Community-based surveys

The most common community-based method used is the cross-sectional field survey. WHO is guiding countries in applying a standardized methodology for implementing the survey. Although the primary objective of the survey is to obtain prevalence figures, coverage of diabetes diagnosis and treatment is assessed as well. Cluster-, multistage- or simple random sampling are used. Implementation is not in individual households, but in a central place in the community. Prospective participants must be prepared thoroughly in order to obtain a good response rate.

The tool used to detect diabetes cases is the oral glucose-tolerance test. Whole blood or plasma glucose value is measured two hours after a 75 gram oral glucose load. Individuals who claim to have diabetes are exempted from this test if they are currently taking insulin or oral hypoglycaemic drugs, or if the fasting blood glucose test, which is offered when there is doubt, is indeed diagnostic for diabetes. The two-hour glucose tolerance test is sensitive and highly specific and therefore considered to be a good tool to obtain valid prevalence numerator data and valid denominator data for the diagnostic coverage indicator.

For logistic or economic reasons, the fasting plasma glucose sample is sometimes used instead of the two-hour glucose tolerance test. In the USA this is the recommended test, while for surveys WHO recommends the two-hour glucose tolerance test. The fasting test may result in false positives when individuals are included who have not fasted.

If these two different tests are used across countries, comparability of prevalence will be compromised and consequently there will no longer be certainty about the comparability of the denominators of the diagnostic coverage indicator.

Health facility surveys

These surveys provide information of direct importance to health facility managers and planners, but do not provide information about how much of the total population in need of services is covered by the health system. The denominator is the registered patient population. The definition of criteria for the diagnosis of diabetes is standardized and largely applied all over the world, providing few problems in terms of validity of the denominator.

With regard to the interpretation of the figure, if the coverage of registered patients is low (contact coverage, treatment coverage), it can be assumed that coverage in the total population in need is low as well . If the coverage is high, not much can be said about the coverage of the population in need, unless reliable figures exist from population-based surveys.

The *proportion of patients with a satisfactory degree of metabolic control* is established with HbA1c estimation. Blood glucose control, averaged over a period of approximately one month, may be estimated by the measurement of glycosylated proteins, especially Haemoglobin A1c (Hb A1c) in blood. Hb A1c is therefore a good indicator of underlying effectiveness of care in the clinical setting. It may also be used to measure the success of specific tertiary prevention interventions. In the field of NCDs such a long-term indicator is probably unique.

Hb A1c estimation is relatively affordable and available. While lack of standardization of the assay between laboratories has prevented this technique from being adopted as a diagnostic test in preference to the traditional glucose tolerance test, its performance within a single health care setting is reliable and effective. Hb A1c concentration has been shown to correlate as closely with incidence of diabetic complications as do fasting and two-hour blood glucose concentrations.

2.5.3 Main issues

- Issues related to assessing the prevalence of diabetes mellitus in the population
- Issues related to validating a numerator for a lifetime intervention
- The survey questionnaire must be validated by the reference test among the subsample of the study population.

References

1. King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. *Diabetes Care*, 16(1):157–177.
2. King H. *Diabetes and noncommunicable disease risk factor surveys, a field guide*. Geneva, World Health Organization, 1999 (unpublished document WHO/NCDNCS/99.1; available on request from Department of Management of Noncommunicable Diseases, World Health Organization, 1211 Geneva 27, Switzerland).
3. *Definition, diagnosis and classification of diabetes mellitus and its complications. Report of a WHO consultation*. Geneva, World Health Organization, 1999 (unpublished document WHO/NCD/NCS/99.2; available on request from Department of Management of Noncommunicable Diseases, World Health Organization, 1211 Geneva 27, Switzerland).

2.6 Hypertension control

The case of hypertension control provides a good illustration of the difficulties of capturing valid information by asking questions and the possibilities for improving validity.

2.6.1 Coverage indicators

There is no coverage indicator as such in HBP intervention. However, the MONICA project classifies the general population in categories of blood pressure:

•Controlled hypertension (SBT<160 and DBT<95 mmHg in those with drug treatment)

•Non-controlled hypertension (SBT>160 and/or DBT>95 mmHg in those with drug treatment)

•Observed hypertension (SBT>160 and/or DBT >95 mmHg in those not treated)

•Normotensive (SBT<160 and DBT<95 mmHg in those not treated).

•Real hypertension is the sum of the first three.

Effective coverage for hypertension control can potentially be constructed as the proportion of hypertensive patients who receive hypotensive drug therapy (for at least two weeks). For this indicator, the numerator includes the first two groups and the denominator corresponds to the number of real hypertension cases. Not much needs to be said about validity issues concerning such indicators, except that these standards (blood-pressure cut-off points) may vary. The analysis of effective coverage for hypertension control should focus on how to measure effective coverage for that intervention.

2.6.2 Measurement methods

The biggest population surveys about hypertension control are carried out by collaborating centres of the MONICA project (Multinational Monitoring of Trends and Determinants in Cardiovascular Diseases). The aim of these surveys is to assess the awareness and control of hypertension in the community.

With regard to sampling, MONICA surveys use 200 subjects in each age and sex group. The total sample size should be 1200 to1600, depending on whether the youngest age group, which is optional, is included.

Sample selection preferably follows simple random sampling or a systematic random sampling method. However, because of the organizational and logistic difficulties associated with simple random sampling, in practice multistage sampling is used. In this case it is recommended to increase the number of individuals selected.

The individuals selected in the original sample who die or move out of RU are considered to be ineligibles and no survey data are collected on them. If a person correctly selected into the sample does not meet the age criteria on the day of examination, he or she still should be considered eligible, if it was not possible to determine the exact date of survey examination at the time when the sample was selected.

A lower participation rate may be expected among the younger groups and in men as compared with women. The self-selection resulting from low participation rates may introduce systematic biases in the estimation of means and rates.

Questions aim to capture different information in relation to either the numerator or the denominator. For the numerator, questions centre on drug treatment and can take the following forms:

- "Are you taking (in the last two weeks) drugs for high blood pressure?"
- "Have you ever been told that you have high blood pressure?"
- "Do you suffer or have you ever suffered from the following diseases?"
- "Are you taking drugs for high blood pressure at present?"
- "Have you taken drugs for high blood pressure in the past two weeks?"
- "When did you take antihypertensive drugs?"
- "How frequently did you use the following medicines?"

Referring to the first question, the MONICA manual suggests coding the answer as uncertain "if the use of blood pressure lowering drugs is reported but the person in question is uncertain if these have been used during the last two weeks or s/he is not sure whether drugs were used for hypertension".

The questions on drug treatment, asking about the treatment "at present" versus "during the past two weeks", might produce slightly different results. However, studies have confirmed that this bias is insignificant and that both questions produce almost the same results. Quality assessment of the data from hypertension control surveys reported obvious discrepancies due to the variation of questions asked about the drug treatment, and coding of drug treatment status as either "yes" or "uncertain".

In some reporting units the question was: "Are you receiving medical treatment at the present time?" without specifying the type of treatment. This systematic bias may cause overestimation

of drug treatment. Asking respondents whether they are under antihypertensive treatment at present versus during the past two weeks may cause underestimation, as those individuals who were on treatment during the past two weeks but stopped the treatment a day or two before the survey would not be captured. However, the quality assessment study of the hypertension control survey considered this as insignificant bias.

Concerning the denominator, HBP prevalence is estimated by asking about awareness of high blood pressure, as in: "Have you ever been told by a doctor or other health worker that you have high blood pressure?" or "Do you have any of the following diseases or symptoms?" Asking the question about high BP awareness without specifying the source of diagnosis (doctor or other health worker) may prompt the respondent to reflect subjective feelings rather than an objective diagnosis. This might cause bias in either direction, depending on the interpretation of the question by a respondent. In some surveys, the question about high BP awareness is not asked. The awareness was considered positive if a respondent answered positively the question about the drug treatment. This results in a bias that overestimates the level of high BP awareness.

2.6.3 Main issues

- A household survey requires objective examination (reference test) of a representative subsample of the study population in order to adjust its validity by the positive predictive value of the reference test.
- Individual variations (such as sex, age, etc.) among subjects affects comparability of the findings.

References

1. Kuulasmaa K, Hense HW, Tolonen H. *Quality assessment of data on blood pressure in the WHO MONICA project*. World Health Organization, 1998, NMH.
2. Molarious A, Tuomilehto J, Kuulasmaa K. *Quality assessment of data on hypertension control in the WHO MONICA project*. Geneva, World Health Organization, 1998, NMH.
3. *MONICA manual, part III: population survey, section 1: population survey data component*. World Health Organization, 1997, NMH.

2.7 Angina pectoris

2.7.1 Coverage indicators

As in the case of hypertension control and management there are no currently established coverage indicators. According to our definition effective coverage of ischaemic heart diseases can be expressed as a proportion of the patients with angina pectoris who get appropriate treatment.

2.7.2 Measurement methods

Medical history-taking and interviewing are prone to systematic and random biases and to low reliability. Therefore, use of standardized questionnaires is recommended in preference to history-taking in order to reduce biases and increase reliability and comparability.

Even with standardized questionnaires, reliability tends to be low, especially as far as the individual is concerned. Since errors are mostly random, however, estimates in a group may be repeatable.

The Rose questionnaire for angina pectoris is a good example of a method that tends to minimize validity errors in determining the prevalence of medical conditions in surveys. The questionnaire includes nine questions on symptoms and medical history and defines a clinical case based on a predefined canvas of answers. The questions are the following:

1. Have you ever had any pain or discomfort in your chest? Yes; no
2. Do you get it when you walk uphill or hurry? Yes; no; never hurries or walks uphill
3. Do you get it when you walk at an ordinary pace on the level? Yes; no
4. What do you do if you get it while you are walking? Stop or slow down
5. If you stand still, what happens to it? Relieved; not relieved
6. How soon? 10 minutes or less; more than 10 minutes*
7. Will you show me where it was? Sternum (upper middle); sternum (lower); left anterior chest; left arm; other
8. Do you feel it anywhere else? Yes; no
9. Have you ever had severe pain across the front of your chest lasting for half an hour or more? Yes; no

Diagnostic criteria for angina pectoris are positive when Q 1 is yes, Q2 or 3 is yes, Q4 is stop or slow down, Q5 is relieved, Q6 is 10 minutes or less, and Q7 is sternum or left anterior chest and left arm. Pain of possible infarction is defined when Q9 is yes. The validity issues with regard to the Rose questionnaire are as follows:

- Epidemiological studies (reference tests) have reported different values of the sensitivity and specificity of the Rose questionnaire. Hagman *et al.* reported sensitivity of 74% and a specificity of 94%, compared with clinical examinations. Another study observed sensitivity of 67%, positive predictive value of 50% and negative predictive value of 99%. The reported sensitivity and specificity of the Rose questionnaire seem to depend on the specificity of the reference test used for validation. A coronary heart disease prevalence study conducted in Spain observed somewhat lower sensitivity and particularly specificity of the Rose questionnaire (52.9% and 52.1%, respectively). The authors explain low specificity of the Rose questionnaire by low specificity of the exercise test itself (reference test) in the general population, especially with low prevalence of coronary heart diseases.
- Sensitivity and specificity reported by other studies vary. In most of the cases, however, sensitivity varies from 25% to 83% and specificity is almost always higher than 75%.
- Sensitivity and specificity of the Rose questionnaire differ between men and women. The prevalence of Rose-positives is higher among women than among men, especially in younger age groups. This could be partly explained by the higher rate of chest pain related to non-coronary diseases among women, and higher fatality of acute coronary heart conditions among men.
- A number of studies examined gender differences in the relationship between the angina reported through the Rose questionnaire and cardiovascular risk factors and other reported symptoms. Nicholson *et al.* have shown in their study that the Rose questionnaire is a better predictor of true angina in men than in women. The study demonstrated that in younger men and women, associations between Rose angina and cardiovascular risk factors and ECG abnormalities are weaker in women, and with a larger proportion of Rose angina unrelated to CAD than in men. The relationship between increased symptom reporting and angina is stronger in women than men. The study permits us to suggest that a smaller proportion of women with angina have underlying atherosclerotic coronary pathologies than men.

2.7.3 Main issues

- The survey questionnaire must be validated by the reference test among the subsample of the study population
- Individual variation (such as sex, age, etc.) among subjects affects comparability of the questionnaire.

References

1. Bass EB, Follansbee WP, Orchard TJ. Comparison of a supplemented Rose questionnaire to exercise thallium testing in men and women. *Journal of Clinical Epidemiology*, 1989(42):385–94.
2. Cederlof R, Jonsson E, Lundman T. On the validity of mailed questionnaires in diagnosing "angina pectoris" and "bronchitis". *Archives of Environmental Health*, 1966(13):738–42.
3. Cook DG, Shaper AG, MacFarlane PW. Using the WHO (Rose) angina questionnaire in cardiovascular epidemiology. *International Journal of Epidemiology*, 1989(18**)**:607–13.
4. Cosin J, Asin E, Marrugat J, Elosua R, Aros F, de los RM *et al*. Prevalence of angina pectoris in Spain. PANES Study group. *European Journal of Epidemiology*,1999(15):323–30.
5. Erikssen J, Forfang K, Storstein O. Angina pectoris in presumably healthy middle-aged men. Validation of two questionnaire methods in making the diagnosis of angina pectoris. *European Journal of Cardiology*, 1977(6):285–98.
6. Glader EL,.Stegmayr B. Declining prevalence of angina pectoris in middle-aged men and women. A population-based study within the Northern Sweden MONICA project. Multinational monitoring of trends and cardiovascular disease. *Journal of Internal Medicine,* 1999(246):285–91.
7. Hagman M, Jonsson D, Wilhelmsen L. Prevalence of angina pectoris and myocardial infarction in a general population sample of Swedish men. *Acta Medica Scandinavica*, 1977(201):571–7.
8. Nicholson A, White IR, Macfarlane P, Brunner E, Marmot M. Rose questionnaire angina in younger men and women: gender differences in the relationship to cardiovascular risk factors and other reported symptoms. *Journal of Clinical Epidemiology*, 1999(52):337–46.

2.8 Oral rehydration therapy

Oral Rehydration Therapy (ORT) has been historically considered and used as a good proxy measure for assessing the coverage of diarrhoeal diseases management programmes. The main sources of information are (and have been) WHO's household surveys, UNICEF 's multiple indicator cluster surveys and Demographic and Health Surveys.

2.8.1 Measurement methods and main related issues

The use of ORT raises a number of issues that have implications for understanding and measuring effective coverage of that intervention. These issues are:

- Issue of longitudinal comparability, since because of expanding knowledge, four successive definitions of ORT have been proposed within a decade (see table below).

Table 1. Changes in definitions of indicators of diarrhoea managementi

Years of references	Proportion of diarrhoea cases in children aged under 5 years who were :
1981	Treated by oral rehydration salts
1988	Treated by oral rehydration: oral rehydration salt solution and/or recommended home fluids
1991	Treated by oral rehydration therapy (increase fluids intake)
1993	Treated by oral rehydration therapy (increased fluid intake) and continues feeding

- Issue of comparability of data accross different types of surveys (WHO/CDD/ARI Household Survey, UNICEF/MICS, DHS). The management considered in the numerator is currently comparable (increased fluid intake + continued feeding), but it had not always be the case.DHS considers children under-threes, while WHO and UNICEF's surveys consider under-fives.

- Issue of mother's recall bias
- Issue of case definition (dirrhoea) that is inherent to the measurement of prevalence by questioning the mother
- Issue in the interpretation of results, particularly in terms of identifying and obtaining data on factors and constraints external to the intervention for interpreting the results: socio-economic factors, environmental factors (including water supply, housing, sanitation), demographic factors, health services related factors (confounding factors and effect modifiers).

Reference

Victora CB, Bryce J, Fontaine O, Monasch R. Reducing deaths from diarrhoea through oral