Validation of indicators for health policy research
Ronney B. Panerai

The information used to determine the subjects of health policy research takes the form of indicators obtained by data manipulation. Validation is the process whereby the ability of health indicators to measure what they are supposed to measure is determined. The author discusses ways of overcoming various limitations affecting the development of health indicators.

Health policy research has been defined by WHO’s Advisory Committee on Health Research as the process of scientific investigation leading to the formulation of strategies, priorities and plans for health development. It can be represented as a continuous loop of inquiry dependent on information collection, analysis and interpretation (Fig. 1). Information is needed not only to facilitate the selection of major health problems but also to allow the identification of alternatives, resources and interventions, and, above all, to monitor change. An analysis of health targets that have been selected for achievement by 2001 (1) indicates a substantial demand for information to be used in planning, implementation, monitoring and the evaluation of progress.

The selection of health indicators and determinants should be based on careful evaluation and validation. This cannot sensibly be accomplished by classic methods involving either prospective or retrospective studies, nor by trial and error. The development of health information systems requires a huge investment of resources, beyond the means of many developing countries, and a long-term commitment to data collection, analysis, dissemination and utilization. Research on alternative methods of assessment and validation of health indicators is essential if rational selection is to be achieved. A particularly difficult problem is presented by the setting of priorities for the promotion of global health development.

What is an indicator?
Health status cannot be simply or directly measured because the concepts of health and disease are complex and multidimensional. There is a similar difficulty with the determinants of health and disease: it is not easy to quantify nutritional, behavioural and environmental factors. Information on conditions of health and their determinants can, at best, be characterized by indicator variables rather than by the classic measurements of physical or biological science.
A rigorous definition of the term “health indicator” is not necessary and can, indeed, be counterproductive. Definitions have ranged from the relatively flexible (2) to the rigid (3). Rigid definitions may eliminate many useful variables and data.

It is generally accepted that indicators are obtained by manipulating data. Whether information can be derived directly from data or from an indicator depends on the circumstances, and for this reason no strict distinction between the terms “information”, “data” and “indicator” is required. It is also worth noting that the term “health indicators” frequently covers indicators of the determinants of health, health care facilities and other resources.

**Approaches to validation**

The literature on the construction and selection of health indicators suggests that their development and validation should be a classic linear multistage process.

Starting with raw data, there is a relatively straightforward stage of construction of indicators based on ratios, indices and so on (2). An indicator may be represented by a single datum, such as the number of houses in a community, or by more complex indices, such as life expectancy or health status indicators. Concerns are increasingly being raised about the information value of composite indicators (4).

The next stage is the classification of indicators according to their properties and attributes. Terms such as “validity”, “reliability”, “specificity” and “sensitivity” are in use, but there is a proliferation of other concepts and attributes which are less clear (2,5). “Validity” expresses the ability of an indicator to measure what it is supposed to measure; “reliability” refers to whether similar results are obtained by different people in similar circumstances; “specificity” is concerned with the indicator’s immunity to changes in what it is not supposed to
measure; and "sensitivity" refers to the indicator's ability to detect changes in what it is supposed to measure (2).

A more coherent classification of indicators would be important for the next stage, corresponding to their selection based on cost and other criteria. The process of selection has been largely influenced by the natural availability of data. With regard to the selection of indicators on the basis of their resource implications, investigators may perceive information either to be free, that is readily available, or, at the opposite extreme, to have an infinite cost (thus being impossible to gather). Further research is necessary in order to provide a more realistic picture of the resource implications of different indicators. Selection based on statistical criteria, such as information redundancy, has been demonstrated by several investigators using multivariate techniques (6).

After selection, indicators should be used and, finally, validated. Because information on the results and experience of the utilization stage is scanty, the use of indicators in policy-making and decision-making either does not happen or fails to provide enough evidence to support the validation process. Ideally, the results of the latter stage would be fed back to the selection stage, forming a continuous cycle of improvement by trial and error. Although this linear process may appear to be the most natural and rational, it has the following flaws.

■ Data collection and the use of indicators are resource-intensive and time-consuming activities, and consequently it is unrealistic to refine the selection of indicators by trial and error.

■ In a majority of cases the formulation of health indicators is influenced by the availability of data. Research or health problems tend to be selected on the basis of what can, rather than what should, be measured.

■ The objectives and expectations associated with the use of health indicators have evolved considerably. In early applications the typical objective was to characterize the health conditions in particular populations. However, as shown in Fig. 1, health policy research also needs information on the determinants of disease patterns and the potential impact of different interventions. This creates a demand for indicators of causation mechanisms, preventive actions, health care supply and performance, and so forth. Unfortunately, many authors have opted for an increasing level of aggregation of single indicators (3). There is a limit to the amount of information that can be aggregated in a single number, and a new approach is clearly needed to cope with the increased demand for information related to health policy research.

The dearth of information on the use of health indicators calls for research aimed at synthesizing the experience gained and lessons learnt in this area, particularly in applications dealing with policy-making.

Need for change

A problem-oriented approach is required to overcome the limitations affecting the development of indicators and information for health policy research (Fig. 2).

■ Instead of exaggerated concerns about indicators currently in use and their expected performance, the initial focus should be on the definition of the
problem to be tackled or the research question to be answered.

- It is not usually sufficient to formulate an initial problem or research question. Subsequent stages of analysis have to be anticipated regarding the use of results for policy-making. If the global assessment of health levels is being analysed it is important to establish what use is to be made of the information with a view to planning the level of indicator disaggregation, and so on. If this exercise indicates a need to investigate cause-and-effect relationships it becomes necessary to hypothesize and to draw up lists of the variables one wants to measure. *Conceptual modelling* tends to generate a large number of different indicators rather than highly aggregated, single-number indicators. Furthermore, much of the information generated is not in the form of input or output indicators per se, but is derived from the structure of the relationships or from their intensity.

**The limitations of indicator production**

Given a list of variables required for the analysis of a selected problem, the next step is to find the most appropriate indica-

> **Instead of exaggerated concerns about indicators currently in use and their expected performance, the initial focus should be on the definition of the problem to be tackled or the research question to be answered.**
infinite resources existed, it would be possible to obtain any data with perfect accuracy.

In reality, of course, resources are limited and there is a trade-off between availability, accuracy and cost. For example, reasonably accurate information on morbidity might be available for only a small sample of a population. Measurement theory distinguishes between accuracy or bias and precision or scatter. Reproducibility also has to be considered. Clearly, absolute accuracy and absolute precision are unattainable because all measurements are susceptible to systemic and random errors. However, accuracy is significantly compromised by ignorance of which targets to use for different data and indicators. As a consequence there is a tendency for overspecification and an accompanying waste of resources. This is particularly critical in developing countries, where

**Accuracy is significantly compromised by ignorance of which targets to use for different data and indicators.**

large amounts of data remain unused because inaccuracy is suspected. Research and simulation studies are required to establish the tolerance or robustness of health policy research in regard to data errors and inaccuracies.

**An alternative: indicator construction**

By and large the main problem connected with the supply of data is their non-availability for many variables or for particular regions or countries. Distortions commonly result, among them the exclusion of countries or regions, the use of inappropriate indicators, and the adaptation of the objectives of studies to what can be measured instead of to what should be measured.

**Knowledge-based methods,** combining artificial intelligence technology with expert opinion techniques, give a new perspective on data availability. These methods have enormous potential in the field of health policy research. Their feasibility should be assessed in pilot projects. In many instances, knowledge-based methods could provide more relevant information on variables that are not easily measurable, for instance the quality of community primary care services. Furthermore, these methods would be comparatively affordable by most developing countries.

Conceptual modelling might create a demand for multidimensional variables, such as environmental risk and sexual behaviour, which cannot be measured directly and require a subgroup of indicators for their characterization. A typical example is the organization of information into health domains, components and subcomponents. This hierarchical structure provides an efficient approach to the construction of more elaborate indicators with the support of classical multivariate methods or the more recent knowledge-based methods.

The significance of multicomponent indicators is that they can synthesize information to describe an entire subsector. There is some potential in this approach for overcoming the problem of data availability and uniformity in cross-national studies, since it is possible to use raw data from different regions to construct comparable sectoral indicators.
Validation

Examined in the light of the problem-oriented approach, the properties, attributes and characteristics proposed for the selection and validation of health indicators decrease in relevance, except for the issues of availability, accuracy and cost. The main reason for the reduced applicability of such concepts as those of validity, sensitivity and specificity is that the properties of an indicator generally depend on its application. Thus it is meaningless to inquire about the validity of “number of dwellings with running water” or the specificity of “infant mortality” unless a purpose is specified. The properties and performance of “infant mortality” differ between its use as an output indicator of “health in the population aged 0–5 years” and its use as an input indicator of “health services supply”. Consequently, the focus of evaluation shifts from individual indicators to the net contribution of a group of indicators as suggested by the conceptual modelling stage. In this case a hierarchical structure can provide an evaluative framework with the following levels.

- **Effectiveness** is the final contribution of health policy research to the achievement of gains in equity and health development. As such it involves the performance of conceptual modelling, policy-making, resource allocation, implementation, monitoring and other elements.

- **Intervention value** is the contribution of indicators and conceptual modelling to the decision to change policy.

- **Diagnostic value** is the contribution of indicators and conceptual modelling to the identification or classification of new health patterns or scenarios of relevance to health policy research.

These definitions show that it is impossible to separate indicators from the conceptual modelling stage. There is also a strong link with health policy research which gives a purpose for the utilization of these indicators and determines the extent to which they can prove their usefulness with regard to changing health conditions or their determinants. It is inconceivable that prospective studies could be set up to assess the above parameters for all indicators and their combinations. The most practical approach is to attempt to estimate impacts using expert opinion and knowledge-based methods.

Advances are unlikely to be made towards more refined measurement of health conditions without a major contribution of expert knowledge (7). Research into the application of newly available knowledge-based engineering methods therefore deserves particular attention.

References


