
2

Panorama of evaluative studies:
Mortality indicators, data sources and methods of estimation *

Claire Fournier
Département de démographie, Université de Montréal, Canada

There is no doubt about the contribution of demography within the field of programme evaluation. Until recently, demographers were involved primarily in evaluative surveys for checking the efficiency of family planning programmes implemented in many developing countries since the 1970s.

For a number of years the evaluation of health programmes has been the domain of the medical community, its administrators and its scientists. They have concerned themselves mostly with the evaluation of resources invested, of the quality of the services and of the effects of the programmes on morbidity.

During the 1980s, the debate opposing the medical and social sciences regarding health and survival research shifted and the current context favours an interdisciplinary approach. At the same time the substantial contribution of demography in the field of health planning and more specifically the evaluation of health programmes is becoming recognized. The literature on the evaluation of health programmes, especially evaluation of the effects of programmes on mortality, gives a clear sign of this new trend.

As we have shown recently (Fournier, 1992, 1994), the evaluation of health programmes is generally based on two aspects familiar to demography: 1) adequate informational support; 2) appropriate theoretical and empirical support to back intervention or programmes. It should be noted that the first aspect has always been important in literature on evaluations, to the detriment of the second (Domato, 1994). Thanks to their expertise in methods for gathering data, for measuring and analyzing data and for demographic statistics and other results derived from demographic analysis, for example, health indicators, demographers are able to provide a response, at least a partial one, to the first requirement. As for the second requirement, demographic studies have greatly contributed to improved knowledge about health and survivorship factors.

* Translated from French by Paul Belle.
This article deals mainly with the first type of support which no doubt represents the most important contribution of demographers in the evaluation of health programmes because demography is behind several pertinent indicators in the field (Chevalier et al., 1995; Dhillon, 1989; Péron and Strohmenger, 1985; Pineault and Daveluy, 1986). The different evaluation strategies presented here are not exhaustive. The relevant measures and indicators, the different evaluation methods and data sources providing information for evaluative studies are reviewed and so are the methodological problems linked to their use. The contexts of developed countries and developing countries are taken into consideration.

The evaluation of health programmes: different strategies

There are a great number of different evaluation strategies mentioned in publications which target one or more components of a programme and which can be carried out during the programme or retrospectively (Domato, 1994; Patton, 1982; WHO, 1973). Clemenhagen and Champagne (1986) built a conceptual framework presenting different evaluation strategies which they put into two main categories: 1) administrative evaluation; 2) evaluative research. Inspired from this conceptual framework the model presented in Figure 1 uses the strategies presented by these authors adding to it the dimension of the required informational, empirical and theoretical support for their realization and provided by demography.

Figure 1. Demography’s contribution to the evaluation of health programmes.
Administrative evaluation is part of a descriptive approach and its objective is to look at the different components of a programme one by one. According to whether it concerns identifying the priority health problems of a given population, allocated resources, the services offered and the impact of the programme, the evaluation is referred to respectively as an evaluation of needs, an evaluation of structure, an evaluation of process, an evaluation of effects. In each case it consists broadly of assessing a component of a programme by applying a criteria (a norm or a standard) established empirically or by experts or by measuring the difference between what has been planned in the programme (the criteria) and what has been achieved after a given period. This type of evaluation is based on the assumption that there is a causal link between the structural and process elements and the effects, without verifying it. An example of an administrative evaluation is the "Evaluation of the strategy for Health for All by the Year 2000" by WHO.

On the other hand evaluative research is intended to be analytical. Rather than considering the components of a programme separately, it analyses rigorously the links between them, whether they concern adequacy or causality, by using a scientific method in the latter case. The results of this analysis enable an assessment of the programme or the components being looked at. Clemenhagen and Champagne (1986) identified as a part of this category strategic analysis, programme analysis, economic analysis and analysis of effects (see Figure 1). Strategies for evaluative research based on informational support where demography plays a role are examined briefly.

Strategic analysis is an attempt to assess the appropriateness of objectives in addressing health priorities and the global health problems of a population. To do so the evaluator must be able to have access to as much information as possible concerning the population in addition to information about the health problems which affect it (Champagne et al., 1987; Daveluy and Pineault, 1986; Dever and Champagne, 1984).

The economic analysis is concerned with studying the link between the resources invested and the services, which is a productivity analysis, or the link between resources and the effects of the programme, which is an efficiency analysis. There are different types of efficiency analysis, which include cost-effectiveness. The costs are expressed in monetary terms and the effects in the form of health indicators. This type of economic analysis can be carried out before implementation of the programme (ex ante) or after (ex post). Its objective is to decide whether investing resources in a programme is justified in relation to the effects expected or observed. It also allows a comparison between programmes so as to choose the most advantageous.

In the analysis of the effects, verifying whether or not there is a causal link between the programme and the effects observed is the basis for the evaluation. The following question is formulated: "Is the effect Y attributable to cause X ?" To answer this question, a scientific approach with the use of research designs allowing a maximum degree of control whilst maintaining a satisfactory degree of validity, is the most desirable. Different designs can be considered, depending on the information available: 1) quasi-experimental design with or without a control group; 2) time series with measurements before and after implementation of the programme; 3) case-control studies, etc.
Informational support and evaluation strategies

Figure 1 presents the evaluation strategies for which demographic data is useful or even necessary. One may distinguish information of a general nature characterizing the population targeted by the programme and information for measuring the state of health of the population and its changes. In this section, the emphasis will be on information requirements for evaluation strategies focused on the effects of a health intervention programme.

Demographic characteristics of the population

A minimum amount of information about the population, its demographic characteristics, its age structure, its projected growth and migration flows if possible, is the basis for health planning and for evaluation. This information is necessary to identify priority health problems among all the health problems affecting a given population. This stage corresponds to the first component of a health programme and is part of the assessment of needs and the strategic analysis. For example, estimations of migration flows of a population subjected to an epidemic is perfectly relevant. In the case of the HIV infection, several studies on this epidemic take into account the migration factor (Lalou and Piché, 1994).

Information on the state of the population and its composition is generally made available through census data. In developing countries these data often require adjustments. Demographic projections are based on fertility and mortality assumptions applied to a given population, and a migration assumption can also be integrated.

Health indicators, data sources and methods of estimation

Health indicators are indispensable for health planning. They enable decisions as to health priorities. In addition their aptitude at revealing changes makes them useful tools for monitoring health improvements, for example the monitoring of certain health indicators by WHO, and for evaluating health programmes and interventions.

There are several health indicators based on mortality and they present great potential for evaluating health programmes. However, these indicators must fulfill certain requirements. Different authors have defined the criteria for choosing indicators. A certain consensus is illustrated in the literature on the subject. The criteria are as follows: 1) validity: the ability of an indicator to supply the "... entire range of values necessary to evaluate the state and the development of the phenomenon for which it was chosen." (Péron and Strohmenger, 1985, p.16); 2) reliability: the ability of an indicator to give stable results in different times and locations; 3) sensitivity: the ability
of an indicator to react to changes of the phenomenon it represents, especially short-term variations; 4) specificity: the ability of an indicator to vary only because of modifications affecting the phenomenon under observation or because of interventions; 5) intelligibility: the indicator's property of being comprehensible for non-specialists such as decision-makers, donor organizations and practitioners (Hill, 1989; Jenkinson, 1995; Mbacké, 1988; Péron and Strohmenger, 1985; Pineault and Daveluy, 1986; Van Lerberghe, 1987; Yach et al., 1990). Of course, all the indicators do not correspond perfectly to these criteria. These criteria are guidelines for making a choice based on the information available and the objectives defined within the framework of the evaluation.

**The case of developed countries**

Compared to life expectancy at birth or life expectancy in good health\(^1\), mortality rates are usually more sensitive to variations in the phenomenon they represent (Hansluwka, 1987). They are based on data from civil registries (death and still-birth certificates, birth certificates) and the census, enabling a direct evaluation of the rates. Among them, the infant mortality rate (IMR) is no doubt the most well known. However, with decreased mortality in all developed countries the IMR has fallen considerably, making this indicator less relevant (Hansluwka, 1987; Pineault and Daveluy, 1986). In the context of developed countries, it is preferable to look at perinatal mortality rates or early neo-natal mortality rates where deaths at an early age are concentrated. What is more, these indicators are more sensitive and specific to perinatal health care given the fact that they are determined to a great extent by such health care (Péron and Strohmenger, 1985).

Mortality rates according to cause of death are also interesting indicators that must, however, be used with precaution because they can be influenced by the age structure of the population. But the solution is to use a rate per cause and per age.

Whereas the quality of mortality data sources is generally good, information about cause of death can introduce errors. The cause of death indicated in the death certificate is the initial cause. This may not reflect reality when, for example, the death is due to multiple causes (Chevalier et al., 1995; Pineault and Daveluy, 1986). Chevalier et al. (1995) add another warning. The geographical and temporal comparability of data on the causes of death may be limited because of several problems: 1) improved medical diagnoses over time; 2) variations in the methods used for declaring and classifying deaths. The ICD-10, International Classification of Diseases is the one currently in use.

Within the analysis of the effects of a programme aimed at reducing deaths due to a given cause, the fact of using the cause-specific mortality rate rather than the global mortality rate from all causes adds specificity to that indicator. This will make it possible to check more easily the causal link between the programme and the effects observed, especially if one is able to prove that mortality due to other causes has not changed.

---

\(^1\) This indicator is an indice based on the life table and data on morbidity and disabilities from hospital statistics and surveys.
However, as Ewbank (1984) points out, one would have to be sure that in such a case there is no synergy between certain diseases.

One indicator used in the economic analysis of health programmes is the number of potential years of life saved (PYLS). Its counterpart, the number of potential years of life lost (PYLL) has been developed to measure premature mortality, and notably premature mortality due to one particular cause of death (Chevalier et al., 1995; Pineault and Daveluy, 1986; Péron and Strohmenger, 1985). In both cases, the calculation uses data on deaths (civil registry) and on the population (census), and depending on the method of calculation the life table may be necessary (Chevalier et al., 1995; Péron and Strohmenger, 1985). For instance, deaths between the ages of 1 and 75 years are taken into account, this last figure corresponds to the life expectancy of the population, for both sexes.

The objective of the economic analysis is to estimate the PYLS by one or several different health programmes, *ex ante* or *ex post*, with the aim of: 1) determining whether a programme is worthwhile from the point of view of gains in the number of years of life, taking into account the sums expended, or 2) choosing the programme which ensures the greatest gain in the number of years of life. To calculate this indicator, the simplest method is to multiply the number of deaths due to a cause occurring between the first birthday and the 75th birthday by the difference between 75 and the average age, for each of the age groups. The next step involves making the sum of the results obtained. However, this method tends to overestimate the gains in years of life. A method based on the comparison between the life table observed and a hypothetical life table in which the cause of death would no longer be effective represents a better choice (Péron and Strohmenger, 1985).

By the way it is calculated this indicator gives greater weight to causes of deaths occurring at an early age. There is, therefore, a value judgement at the heart of this indicator, which results in a prejudice in favour of interventions affecting young people (Chevalier et al., 1995; Péron and Strohmenger, 1985; Pineault and Daveluy, 1986). However, this bias can be got around by introducing a weighting factor in the calculation which will make it possible to confer more weight on the years still to be lived in the close future compared to the years to be lived in a more remote future. This notion of current value is one used in economics (D'Souza, 1989; Ewbank, 1984).

Also critical of this indicator, Preston (1993) adds that the method of calculation used is not appropriate because it supposes that the health programme will be completely efficient from the first year of its implementation, therefore the gains in years of life will be almost immediate. According to this author, the effects of a health programme are in general felt gradually and should also benefit future generations. He therefore believes that the calculation should be established on the basis of demographic projections.

Other indicators concern potential years of life saved but include information on morbidity and quality. One of the indicators most often used is the quality-adjusted-life-years (QALY) (Green and Barker, 1988; Pineault and Daveluy, 1986). Although this indicator has been very often used for the economic evaluation of health programmes it has also been criticized because of the subjective approach regarding the appreciation of health and quality of life.
Finally, given that equity in health is a recognized objective, it is important to check whether health programmes make it possible to reduce inequality in chances of survival. Measuring this phenomenon poses a number of major problems. There is, however, one relevant indicator for the differential mortality between the sexes: the excess male mortality index which relates the rate (or quotient) of male mortality to the rate (or quotient) of female mortality according to age or age group, with female mortality expressed on the base 100. However, when interpreting this indicator one should take into account the role played by biological factors. The existence of excess male mortality during the intrauterine period and the perinatal period (neo-natal component) is well documented. The fact that maternity exposes women to life-threatening risks which men are not subject to should also be taken into account. However, the literature indicates that exogenous factors (cultural, social, behavioral) have an impact on the mortality differences according to sex and therefore that it is possible to reduce the existing differences (Péron and Strohmenger, 1985).

The case of developing countries

The health indicators identified until now are just as valid for developing countries, as long as there is information of sufficient quality to calculate them. However, given the fact that these countries have a different health and epidemiological situation from that of developed countries, it is useful to approach the question of the usefulness of health indicators by taking their context into account.

Given the high mortality level which is generally characteristic of these countries, especially at young ages, the IMR and the child mortality rate (in certain countries) are relevant indicators which add to mortality indicators for the perinatal period. These indicators can be used in analyzing or evaluating the effects of a health programme for improving infant and child health.

In developing countries the fact that the quality of information is often questionable is a major problem. Civil registries are often lacking in coverage and in precision (Ewbank, 1988, 1984; Hill, 1989). In some cases there is no information at all (WHO, 1981). Problems linked to the identification of cause of death which are to be found even in developed countries are even greater in developing countries: 1) diagnosis errors (Graham, 1989; Unger and Killingsworth, 1986); 2) omission in declarations (Rumeau-Rouquette, 1986).

Despite these imperfections, some authors mention the value of data from civil registries (or death registries) when they are of sufficient quality and can be adjusted using various methods which we will not examine here (Fargues, 1986; Fargues and Ouaidou, 1988). What is more, this data makes it possible to identify variations due to seasonal factors (or to other cyclical phenomena which are characteristic of the epidemiology of certain diseases). This type of information is very useful when interpreting causes behind mortality changes during the evaluation of a health programme.
Given the problems related to civil registration, developing countries often use data from surveys or censuses. For developing countries surveys have been and still are today a solution to problems in measuring population phenomena, notably mortality. There are several methods for the gathering of data and the estimation of mortality, each of which has its own advantages and drawbacks (Ewbank, 1984; Palloni, 1991; Zoungrana, 1990). This is a field in which demography has a long-established expertise.

There are direct and indirect methods for measuring infant and child mortality from information gathered in censuses and surveys. Among the direct methods there is the use of the information on events having occurred during the last twelve months, births and deaths, in each household. It is a well known fact that this method is not a good way of estimating infant mortality level, as Zoungrana points out (1990). However, it can provide some indications on long-term changes. The method using birth histories, more appropriate for evaluating infant and child mortality, also provides indications about changes in the phenomena (Zoungrana, 1990). However, according to Mbacké (1988), this method does not enable an estimation of short-term changes. This problem also concerns indirect measurements based on live-born child survivorship of the "Brass" type (Garenne, 1994; Mbacké, 1988), which in any case give a very approximate indication of mortality changes with the problem of timing of estimates (Hill, 1989). Because all these measurements are based on information collected retrospectively, they are tainted by bias due to problems of memory, such as omissions and errors in declarations (Zoungrana, 1990).

According to Zoungrana (1990) another type of survey makes it possible to monitor more closely the rates, age pattern, short-term and long-term trends in infant and child mortality, since they include a prospective approach: these are the multiple-round surveys. Palloni (1991) does not share this positive opinion of this type of survey. After having reviewed the main surveys conducted in Africa and Latin America he concludes that the results are affected by different sources of distortion, notably those linked to losses affecting samples.

Of all indirect methods for estimating infant and child mortality, one is deemed more promising for the evaluation because it makes it possible to monitor short-term changes in the phenomenon and to identify risk groups (Mbacké, 1988). This technique is the previous birth technique elaborated by Brass and Macrea and taken up by Aguirre and Hill. Based on information gathered following a simple question posed to women attending a health or maternity centre, regarding the survival status of their previous child, it is possible to estimate the mortality quotient $2q_0$ (Mbacké, 1988; Palloni, 1991). Of course, this method is not problem-free because among other reasons, there is selection bias at the level of the sample of women. However, it is an interesting approach.
Sources of continuous reporting such as longitudinal surveys (for example, Ngayokhème in Senegal since 1961) or the population laboratories (for example, Matlab in Bangladesh since 1963) enable direct estimations of infant and child mortality and even of adult mortality, if the size of the population being studied permits it (Palloni, 1991). The quality of these data is widely recognized (Palloni, 1991; Zoungrana, 1990). The monitoring of short-term changes is possible which makes them an almost unique source of data for analyzing the effects of the health programmes on mortality, as reported by Zoungrana (1990). However, although these surveys are particularly adapted to testing certain health interventions, they are very expensive, difficult to replicate, and their results cannot be considered as representative (D'Souza, 1989; Hill and Kaufman, 1986, see Zoungrana, 1990).

Infant and child mortality rates (or quotients) estimated using the methods and sources mentioned previously are useful to try to monitor the effects of a health programme. However, it should be pointed out that these indicators are not specific enough regarding the health interventions since other factors also affect mortality at these ages, notably socio-economic factors, cultural factors, the education of the mother, etc. In addition to infant and child mortality indicators, it would be preferable to use other indicators so as to get a good picture of the sources of influence.

In addition to infant and child mortality, the high levels of maternal mortality observed in several developing countries are a cause for concern. WHO has therefore decided to ensure the monitoring of the maternal mortality rate in the world (WHO, 1981). Increased interest in this phenomenon has been observed in recent years among health planners and demographers, the former requesting the latter to provide them with a reliable portrait of the situation, including indicators and differentials (Fargues, 1992). The maternal mortality rate is a ratio of the deaths of women during pregnancy, at childbirth and shortly after childbirth compared to women of childbearing age. There is also the maternal mortality ratio for which the number of live births is the denominator. This ratio is expressed in x for 100,000 live births.

Fargues (1992) worked at comparing different methods of estimating maternal mortality with the aim of checking their potential for adequately measuring rates, short and long-term trends, and mortality differences. Four methods are compared: 1) direct measurements based on data from continuous reporting; 2) indirect measurements based on the survivorship of the sisters; 3) estimation based on the life table; 4) measurements drawn from case-control studies. The results of his comparative analysis showed the potential of direct measurements and of estimates based on the life table to detect short and long-term variations in maternal mortality. However, the levels must be considered as approximate whichever method is used. Fargues considers case-control studies as a potentially promising method. However, not having tested this method with real data, his evaluation remains unverified.
Conclusion

Different authors have supported the idea of resorting to using different information sources and estimation methods, when they are available (Fargues, 1992; Garenne, 1994; Hill, 1989). Garenne (1994) stressed the usefulness of combining retrospective data for mortality trends prior to the intervention and continuous registration data for short and medium-term mortality trends after the beginning of the health programme.

The advantage of using other health indicators, for example, the incidence and the prevalence rates of certain diseases, has also been established as a condition for conducting evaluative studies on the effects of health programmes. Mortality measures one aspect of a population's health and it is important to use other measurements which will help complete the picture of the health situation and how it changes (MBacké, 1988; Miller and Hirschhorn, 1995). Other information is also relevant, for example, information concerning the use of health services (prenatal services, oral rehydration salts, etc.) provided by surveys. Qualitative methods can also prove very useful for checking whether interventions are favourably considered by target populations, and for which reasons (cultural or economic factors, time allocation, etc.).

The use of information from different sources can facilitate the interpretation of the results of an evaluative study on the effects of a health programme insofar as these results appear to converge. In the opposite case, of course, it may well complicate matters. However, this will simply contribute to increasing the depth of the analysis.

The evaluation of health programmes involves a combination of several disciplines and approaches. Demographers, thanks to their expertise in the collection, exploitation and management of data, in estimation techniques and in the production of demographic and health statistics, play an important role in programme evaluations. In collaboration with researchers from other disciplines and with health planners and administrators, it is up to them to continue contributing to this field of research and action to improve sanitary conditions.

References


Fournier C. (1992) La place de la démographie dans l'évaluation des programmes de santé, Texte soumis à titre d'épreuve écrite de l'examen de synthèse dans le cadre du programme de doctorat, Université de Montréal, 73 p.


