WORLD HEALTH ORGANIZATION
Regional Office for the Eastern Mediterranean
ORGANISATION MONDIALE DE LA SANTE
Bureau régional de la Méditerranée orientale

منظمة الصحة العالمية
المكتب الإقليمي لشرق المتوسط

REGIONAL COMMITTEE FOR THE EASTERN MEDITERRANEAN

Forty-sixth Session

Agenda item 9(c)

EM/RC46/8
July 1999

Original: Arabic

TECHNICAL PAPER

FORECASTING IN COMMUNICABLE DISEASES

CONTENTS

page

Executive Summary

1. Introduction 1

2. Approaches to forecasting in communicable diseases 1

3. Uses and limitations of forecasting 2

4. Preprequisites for forecasting 2

5. Steps for effective forecasting 4

6. Examples of applications 7

7. Conclusions and recommendations 10

Executive Summary

Communicable diseases are still the leading cause of morbidity and mortality at the global level. New ones are emerging and some of the old ones are resurging, while large-scale epidemics are still being reported from all corners of the world. A comprehensive and feasible strategic plan is needed to address the problem of communicable diseases. Future health scenarios of communicable diseases that can predict likely, probable or even merely possible trends (i.e. forecasting) can play an important role in developing such a strategic plan.

Forecasting is the process by which future events can be predicted. There are unlimited ways to develop a forecast. Some essential features in communicable disease epidemiology make it necessary to use quantitative projection techniques for proper forecasting. In principle, forecasting should not be regarded as providing the answer but rather as a tool to increase understanding and highlight important processes.

Several items need to be well defined before developing a meaningful forecast in communicable diseases. These items include the availability of historical data and the accuracy of such data, proper understanding of the causes of changes in patterns in the past and determination of the factors that might effect changes in the pattern of the communicable diseases in question in the future. There are several steps in the forecasting process which make it effective and these should be undertaken systematically. They begin with defining the problem and then investigating the variables of interest and the reliability of data. This is followed by analysing the data and then selecting the appropriate forecasting method(s) and applying it to determine the final forecast.

Forecasting has been used to predict epidemics to project incidence and mortality of specific diseases, to select the most cost–effective intervention strategy and to design control programmes.

Forecasting in communicable diseases necessitates good epidemiological practice, a functional surveillance system and statistical and mathematical expertise. It is recommended that these capacities are strengthened and forecasting is introduced and used more efficiently in the Eastern Mediterranean Region.

1. INTRODUCTION

Despite tremendous progress in public health practice in the 20th century, communicable diseases are still the leading cause of morbidity and mortality on the global level. Out of every three deaths annually, one is due to communicable diseases. New diseases are emerging and some of the old ones are resurging. Large-scale epidemics of specific communicable diseases are still being reported from all corners of the world, reviving memories of an almost forgotten past. Many of these epidemics were not foreseen or predicted. The problem is felt much more in developing countries, particularly with the concurrent rising trends in the occurrence of chronic noncommunicable diseases.

In view of this situation and taking into consideration the vast scientific developments in this century, in addition to the availability of proven tools of intervention, WHO considers that the problem of communicable diseases should be addressed within the context of a comprehensive and feasible strategic plan. Future health scenarios that can predict likely, probable or even merely possible patterns for communicable diseases can play an important role in developing such a strategic plan.

Forecasting is the process by which future events can be predicted. Prediction of the future is based on an assumed set of circumstances. Forecasting is defined[1] as a method of estimating what may happen in the future that relies on extrapolation of existing trends. Another term used by epidemiologists in this regard is scenario building, which is a method of predicting the future that relies on a series of assumptions about alternative possibilities.

2. APPROACHES TO FORECASTING IN COMMUNICABLE DISEASES

There are many ways to develop a forecast. Some use relatively simple methodology to extrapolate or project the historical pattern, but the more formal statistical methods of extrapolation usually make use of a mathematical forecasting model.

There are certain essential features in the epidemiology of communicable diseases that make it necessary to use mathematical models for forecasting. The most important of these is the fact that the risk of getting an infectious disease is linked to the overall risk in the population. An individual who is infected with a pathogen (e.g. influenza virus) increases the risk of others contracting the infection. For noncommunicable diseases there is no such connection; an individual who develops heart disease does not increase the risk of other individuals in the population with regard to heart disease. A consequence of this is the fact that protection of an individual from infection reduces the risk of infection for other individuals and any intervention preventing infection in a proportion of individuals has immediate benefits for other individuals at whom the intervention was not targeted or to whom it was not given. A simple deduction of the above fact is that communicable diseases require mathematical tools that are non-linear since biological populations are controlled by processes that are non-linear. An example of this lies in the simple observation that reducing the density of a disease vector by half does not lead to reduction

---

[1] International Epidemiological Association, *A dictionary of epidemiology*, ed. J.M. Last, New York, Oxford University Press, 1995.

of the disease incidence by half. It is sometimes possible to stop transmission of a disease by reaching a level of immunity of 90% or 95% in the population and hence the chance of an infected person spreading infection to a susceptible becomes remote. Another important deduction is that the mathematical tools need to be modified according to the prevailing situation within communities or populations. Hence forecasting of what will happen in a community exposed to a communicable disease will vary according to the prevailing important parameters specific to the disease in question. For example, the forecast of what will happen as a result of a case of cholera in a community depends on a number of factors, most important of which is sanitation.

3. USES AND LIMITATIONS OF FORECASTING

Health services planning and disease prevention programmes require some idea of the future burden of disease in order to guide the health policy process. Decisions and choices need to be made today in order to cope with the expected disease burden in the future; forecasting provides the information support for this process. This is of particular importance in preventing unusual occurrence of communicable diseases (epidemics) or at least in being prepared for it so that it can be controlled efficiently. However, the results of forecasting models should be regarded with a certain degree of caution, particularly because of the wide confidence intervals linked to the projections made.

4. PREREQUISITES FOR FORECASTING

Several items need to be well defined before developing a forecast for a communicable disease. Of these, the most important are related to the basic factors determining the epidemiology of the disease in question and the past experience with that disease in a population. These items can be summarized as follows.

a) *Availability of historical data about the communicable disease and the accuracy of such data.* Sufficient historical data are needed to reveal significant patterns of changes by time, such as long-term trend, and seasonal and cyclic fluctuations. These time patterns are of great value in forecasting. For example, if unusual occurrence of a disease starts at the beginning of its season then it can be expected that a large-scale epidemic will follow. If, however, this unusual occurrence begins at or near the end of the season the increase to be expected will not be much. The season for meningococcal meningitis in Sudan begins in December and continues until May when it quickly disappears as rains begin to fall. If the increase in meningitis begins in December or January it may reach epidemic proportions, but if it starts in April it will not usually reach a serious level. The same picture can also be seen in other diseases such as malaria and cholera.

   There are usually different sources for collecting data, of which disease surveillance is the most comprehensive although not always the most reliable. Disease surveillance systems vary greatly from one country in the Eastern Mediterranean Region to another, and within the system in each country from one communicable disease to the other. In general, the surveillance systems in the Region face one or more of the following constraints:

- lack of comprehensive surveillance guidelines which outline clearly all the steps in the surveillance process;
- lack of concentration on priority diseases;
- weak laboratory support and weak environmental surveillance activities;
- inability to initiate a system that is efficient, "simple" and at the same time sufficiently "useful";
- lack of involvement of the private sector, nongovernmental organizations and governmental non-ministry of health institutions;
- shortage of trained health personnel;
- strong centralization and poor role of the periphery; and
- lack of a feedback mechanism.

In the past few years steps have been taken to upgrade and vitalize the surveillance systems in the Region, including disease prioritization, use of standardized case definitions, better reporting and appropriate data analysis. Other efforts have included the upgrading of national epidemiological skills through proper attention to training in epidemiology, both in undergraduate courses and in development of an international course on epidemiology and control of communicable diseases, support for national training activities in that area, and creation of a field epidemiology training network (TEPHINET) in collaboration with the Field Epidemiology Training Programme (FETP) supported by the Centers for Disease Control and Prevention, Atlanta.

However, there is still a need for more efforts in this regard. A good surveillance system is essential for proper action now and in the future. It is an important tool to provide the base from which to forecast.

Other sources of data include historical records, surveys and longitudinal studies. These other sources might be helpful but they cannot replace regular surveillance which is the basic tool.

Data integrity and accuracy is another prerequisite for successful forecasting. It is important to see that the data used in the forecasting process is "healthy data". This entails transparency, consistency and completeness in data reporting. Unfortunately, the ultimate benefit of providing valid data is still not well understood. A number of countries continue to mask or suppress data concerning specific diseases, giving various reasons for such action. However, whatever the reasons, falsification of data or lack of provision of data ultimately reflect negatively on the country itself. The quality of data is critical in developing good forecasts and enabling appropriate steps to be taken to plan for the future.

b) *Proper understanding of the causes of changes in the past.* Changes reflected in the trend or seasonal, cyclic or irregular changes which are detected in a communicable disease time series should be explained by identifying the epidemiological factors that might be the causes of such changes. This will entail identification of the agent-related, host and environmental factors that may be shaping the observed past pattern of the communicable disease time series. All external factors, such as reclassification, introduction of a new diagnostic method or implementation of a specific intervention technique at a specific point in time, should be identified as well. There is no doubt that sound epidemiological analytical thinking is of paramount importance in such a process.

Epidemiological skills are essential for understanding disease patterns and elucidating the determining factors of such patterns, and in this respect they are also essential for the forecasting process.

Unfortunately, there remains a deficit in the overall epidemiological capacity in the Region, particularly in the countries where the epidemiological service is deficient.

c) *Factors that might effect change in the patterns of communicable diseases in the future.* A recent document endorsed by the Regional Committee discussed in detail the changing pattern of diseases and the factors behind such changes[2]. These factors are environmental, such as climatic changes; human, such as changes in lifestyle and behaviour; and agent-related, such as genetic changes. The basic factors that might affect the dynamic transmission of communicable disease in the future are numerous, and were discussed extensively in EM/RC41/7. A proper understanding of the relevance of any of these factors and their possible role in shaping the epidemiology of communicable diseases at the regional and national levels necessitate continuous monitoring, using proper epidemiological approaches and indicators.

   In addition, external factors, particularly those related to the introduction of specific interventions and their effects on herd immunity, should be clearly outlined as they play a significant role in changing the future pattern of communicable diseases.

5. STEPS FOR EFFECTIVE FORECASTING

Forecasting is a process with a set of steps that can be reproduced if necessary. Leaving aside the mathematical details, these steps can be summarized very generally as follows.

1. *Defining the scope of forecasting*

- Determining the communicable disease to forecast. The communicable disease of interest should be a priority disease which constitutes an important health problem, whether because of its magnitude or because of its high potential for hazard.
- Determining the forecast horizon, i.e. the length of time into the future for which forecasts are to be developed. The horizon might be short-term (less than 3 months), medium-term (3 months to 2 years) or long-term (more that 2 years). The forecast horizon affects both selection of the forecast method and the accuracy of the forecast. Generally speaking the shorter the horizon the easier and more accurate is the forecast.

---

[2] *Changing patterns of diseases and their impact on WHO collaborative programmes.* Technical paper presented to the Forty-first Session of the Regional Committee for the Eastern Mediterranean, 1994 (EM/RC41/7).

- Determining the purpose of the forecast, how will it be used and who will use it. As already stated, the forecast should be looked at as a tool which will be useful if interpreted with a certain degree of caution. The user of the forecast should know its limitations and should have the expertise to understand the outlines of the process and how to use the outcomes.

- Determining the level of targeted forecasting accuracy. This depends on many factors including the characteristics of the data available, the forecast horizon, the purpose of forecast and the forecasting method.

2. *Identifying the variables of interest*

- Investigation of sources of available data and their time-span. It is important to know how far back in time historical data are available for and over what time interval they have been measured and whether any change in the system, such as reclassification, has occurred during the past. Various sources of data should be approached taking into account their completeness and means of collection.

- Determining factors that might have caused changes in the past or that might help to explain the historical pattern of the disease in question. In this regard all epidemiological factors, whether related to the host, the environment or the agent, that might explain the pattern should be thoroughly investigated. In addition, investigation will include all external factors or unusual circumstances that might have had a role in shaping the pattern.

- Identifying the relevant epidemiological factors, both intrinsic and external, that might effect changes in the pattern of the communicable disease under study in the future. They will be either intrinsic factors (host, environment or agent) or external factors, such as introduction of specific intervention techniques and their effect on herd immunity or modes of transmission.

3. *Checking reliability of data and preparing it for analysis*

- Checking reliability of sources of data. Primary sources, such as surveillance systems, should be investigated to determine how reliable their data are. Reliable sources should be sought also for supplementary data corresponding to the candidate influential factors, which will be used in the explanatory methods of forecasting.

- Checking data integrity. The quality of data is critical in producing reliable forecasts. Thus, it is important to examine the data carefully prior to the development of a forecasting model. Various statistical techniques, such as calculation of autocorrelations, summary measures and exploratory data analysis, can be used as tools for examination of data.

- Correcting any inconsistency in data, such as missing values, outliers (unusually large or small values) and time period data aggregation.

4. *Analysing the data*

- Determining the characteristics of the data, which means identifying the presence of a trend, whether it is increasing or decreasing, and whether it is linear or non-linear, and also identifying the presence of seasonal and cyclic changes.

- Determining the relationships between various variables affecting the communicable disease pattern. In this regard it is important to discover and quantify the relationship between variables of interest. This is usually done by using simple statistical techniques, such as measuring correlation.

5. *Selecting one or more appropriate forecasting method.* Forecasting methods are either qualitative or quantitative. Qualitative methods are used when there is little or no quantitative information; they rely on judgement, intuitive thinking and accumulated knowledge. One such technique uses a series of questionnaires to obtain information on a particular topic from a panel of experts.

   Quantitative methods are used when historical numerical data are available and when patterns found in the past are expected to continue in the future. Quantitative forecasting methods might be called quantitative projection techniques and they vary greatly in sophistication, from naive and intuitive methods based mainly on judgement and experience, to more formal methods developed on sound statistical principles. They can be grouped into two major approaches: the explanatory and the time pattern. Explanatory forecasting, which is our main interest, is based on the concept that changes in the variable or communicable disease under question are caused by or related to changes in certain influential factors. The time pattern approach focuses on the past patterns in the data and does not attempt to explain why the data changes over time. Of the quantitative explanatory projection techniques, it is preferable, where possible, to select more than one forecasting method for the particular problem at hand.

6. *Applying the selected forecasting method(s) to determine tentative forecasts.* This step implies the development of a forecasting model and its application to obtain forecasts of the communicable disease of interest.

7. *Comparing the results of the forecasting methods.* Results of the forecasting methods used might differ significantly, in which case the forecaster should benefit from the additional information provided by each method, or the forecasts might be very similar, in which case the methods tend to validate each other.

8. *Determining the final forecast.* In determining the final forecast, consideration is given to the reasonableness of results in relation to past changes and assumptions regarding the future. If more than one forecast with reasonable results is available a final forecast consisting of the combined forecasts may produce more accurate results than any one of the individual methods.

9. *Monitoring the forecast and updating it.* A tracking signal is usually used to investigate significant deviations of the actual observed values from the forecast values. The forecast should be modified and updated as appropriate.

6. EXAMPLES OF APPLICATIONS

There are many examples of the use of forecasting in communicable diseases to predict unusual occurrence (epidemics), to project incidence and mortality of specific diseases, to select the most cost-effective intervention strategy and to design control programmes. The following are a few examples of the applications in the various areas.

a) *Predictions of epidemics of certain communicable diseases*

- Forecasting has been used to predict epidemics of such communicable diseases as cholera, malaria, meningitis and Rift Valley fever. The list of diseases includes more than those mentioned, but discussion will be limited to these four since the Regional Office had had some experience in forecasting epidemics of these and has provided support to prevent or mitigate their impact.

- Cholera outbreaks in the Eastern Mediterranean Region countries in the Horn of Africa were successfully predicted in recent years by intuitive methods of quantitative projection based mainly on the occurrence of the disease in neighbouring countries, on prevailing environmental sanitation, on past experience and on good scientific judgement. Although it was difficult to prevent such epidemics, it was possible to mitigate their impact by being well prepared. The value of such anticipation and preparedness was reflected in the case fatality rates which were maintained at their lowest possible despite the difficult situations.

- Malaria outbreaks were predicted in Sudan and the Republic of Yemen using very similar methodology and measures were taken to limit their toll. The timely application of these measures was successful in limiting the magnitude of the outbreak. More success could have been achieved had the forecasts been developed and taken more seriously.

- A number of efforts have been made to develop systems/indicators to anticipate epidemics of meningococcal meningitis. Outbreaks during the past three years in the meningitis belt countries resulted in unprecedented morbidity and many thousands of deaths.

  The outbreak of meningitis in Sudan in the early months of 1999 was rightly predicted by the Regional Office and the national authorities who were well prepared for it. Meningococcal meningitis in Sudan has not reached epidemic proportions since 1989. The increase in the number of cases, which started at the beginning of the season (December 1998), and the available information on cessation of preventive vaccination of high-risk groups during the past few years were indicators of the potential for a serious epidemic. WHO collaborated with national authorities in preparing an emergency plan and in immediate implementation of that plan, and it is believed that this had substantial impact on the epidemic, despite the many constraints. Analysing the epidemic, and particularly its geographic spread in Sudan, it appears that the epidemic started spreading late in some provinces and therefore there is a possibility of its continuation in the next season, particularly in areas marginally hit by the 1999 epidemic. It is hoped that these data will be used to avert this potential in a timely manner.

- Prediction of epidemics in Africa is now the subject of several studies that depend on developing a climate-driven model, using data from meteorological satellites and weather prediction models. Climate-driven systems are also being used to predict outbreaks of Rift Valley fever in the endemic areas of sub-Saharan Africa.

- *Projecting measles outbreaks.* The introduction of measles vaccination and achievement of very high coverage may result in stopping the transmission of natural infection with measles virus and accumulation of susceptibles over the years. When the virus is introduced into such communities the disease will spread rapidly among the accumulated, non-immune population causing an outbreak. Here, anticipation and forecasting and timely intervention are essential tools in the prevention of such an outbreak.

  In countries where routine immunization of infants against measles reaches 95% and given that measles vaccination during infancy will result in approximately 90% protection, by simple calculation, this means that $95 \times 90 = 85.5\%$ of newborn children will be immune and 14.5% will remain susceptible. If these are not vaccinated again or exposed to natural infection, they will accumulate year after year and in about 7 years they will be of the size of a birth cohort which can support the occurrence of an epidemic once infection is introduced into the community. These facts have been used to forecast potential epidemics of measles and in the timely introduction of preventive measures; namely, giving an additional dose of measles vaccine to those age groups among whom susceptibles have accumulated (those born since the last epidemic or last mass vaccination campaign). Vaccination of schoolchildren was carried out in Great Britain in 1994 when it was forecast that an epidemic might occur as a result of accumulation of a sufficient number of susceptibles and the epidemic was averted. These principles are now the basis for the global strategy for elimination and eventual eradication of measles which is being implemented in many countries of the Region.

b) *Projecting incidence and mortality*

- *Projecting HIV incidence and mortality.* The WHO Global Programme on AIDS prepared projections using EPIMODEL (developed by WHO) which fits a gamma distribution to reported AIDS cases and generates predictions of HIV incidence and mortality. As a result of these projections it was suggested that in the year 2020 there would be essentially little or no HIV incidence anywhere. Later that was found to be too optimistic a scenario, and it was modified by the Global Burden of Disease exercise so as to be more realistic. As a result the number of incident cases per year is assumed to stabilize once incidence has fallen to one-half of peak incidence (Figure 1).

- *Projecting the impact of HIV on tuberculosis.* There is a strong interaction between HIV infection and tuberculosis. In areas where HIV incidence is high, more than one-third of HIV positive individuals are expected to die from tuberculosis. The increased incidence of tuberculosis in HIV positives will certainly influence the annual risk of infection of tuberculosis for HIV-negatives in the community. This will lead in due time to higher incidence and mortality from tuberculosis in HIV-negatives. Mathematical models of tuberculosis were developed to project the impact of HIV on tuberculosis. Results from these models show that the expected decline in mortality from tuberculosis as a result of introduction of tuberculosis control measures will not begin to appear in high HIV-prevalence areas until the year 2010.

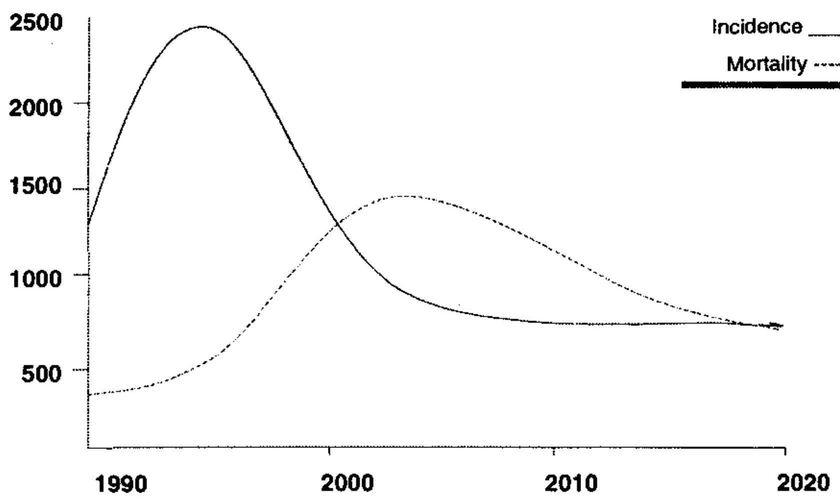

Figure 1. Projected HIV incidence and mortality

c) *Determining the cost–effectiveness of control strategies*

Forecasting change in the epidemiology of diseases as a result of the introduction of control measures is useful also in determining cost–effectiveness of control strategies

A model has been developed to evaluate the outcome of immunization strategies to prevent hepatitis B virus transmission. The model was used to determine the incremental effects (the reduction in infections and costs of HBV-related liver disease) of several immunization strategies (prevention of prenatal HBV infection, routine infant vaccination and routine adolescent vaccination) in a birth cohort. The results showed that routine vaccination of infants in successive birth cohorts to prevent HBV transmission is the most cost-effective strategy over a wide range of assumptions.

Similarly a mathematical model was used to determine the most effective option for the control of *Ascaris* infection by chemotherapy. It was found that it is more cost–effective to treat as many people as possible in each round of treatment rather than increase the frequency of treatment. This is largely because of the effect of treatment on those untreated.

d) *Designing control programmes*

Forecasting has also been used extensively to help design many control programmes in which vaccination was the main intervention strategy. In this regard mathematical models have been used to determine the minimum coverage rate required in continuous mass vaccination to eliminate infection and to determine the optimum age for vaccination, the choice of vaccine and the need for booster doses.

7. CONCLUSIONS AND RECOMMENDATIONS

In order to develop a more comprehensive strategic plan to address the problem of communicable diseases information may be needed about the future outlook of priority communicable diseases so that decisions and choices can be taken to cope with them. Forecasting is the process by which such information can be obtained, providing the pre-requisites and the steps for effective forecasting are soundly met.

Forecasting in communicable diseases necessitates good epidemiological knowledge and practice on the one hand and a functional surveillance system on the other. Statistical and mathematical expertise is also needed to develop the models. However, without sound epidemiological thinking, the process can be misleading.

In this regard, it is recommended that Member States, with WHO support:

1. Continue to improve their disease surveillance systems by implementing the regional strategy for development and strengthening of epidemiological surveillance.
2. Introduce the concept of forecasting for priority communicable diseases and strengthen their capabilities in this regard.