Theme Papers

Health information systems: the foundations of public health

Carla AbouZahr[1] & Ties Boerma[1]

Abstract Public health decision-making is critically dependent on the timely availability of sound data. The role of health information systems is to generate, analyse and disseminate such data. In practice, health information systems rarely function systematically. The products of historical, social and economic forces, they are complex, fragmented and unresponsive to needs. International donors in health are largely responsible for the problem, having prioritized urgent needs for data over longer-term country capacity-building. The result is painfully apparent in the inability of most countries to generate the data needed to monitor progress towards the Millennium Development Goals. Solutions to the problem must be comprehensive; money alone is likely to be insufficient unless accompanied by sustained support to country systems development coupled with greater donor accountability and allocation of responsibilities. The Health Metrics Network, a global collaboration in the making, is intended to help bring such solutions to the countries most in need.

Keywords Information systems/organization and administration; Statistics/organization and administration; Delivery of health care; Health status indicators; Data collection; Decision making; Policy making (*source: MeSH, NLM*).

Mots clés Système information/organisation et administration; Statistique/organisation et administration; Délivrance soins; Indicateur état sanitaire; Collecte données; Prise décision; Choix d'une politique (*source: MeSH, INSERM*).

Palabras clave Sistemas de información/organización y administración; Estadística/organización y administración; Prestación de atención de salud; Indicadores de salud; Recolección de datos; Toma de decisiones; Formulación de políticas (*fuente: DeCS, BIREME*).

Arabic

Bulletin of the World Health Organization 2005;83:578-583.

Voir page 582 le résumé en français. En la página 582 figura un resumen en español.

Measurement and public health

"Nothing exists until it is measured" the physicist Niels Bohr famously stated in 1930. He was referring to quantum mechanics, but the idea is also relevant to public health. The definition of public health suggested by Beaglehole et al. as "collective action for sustained population-wide health improvement" presupposes the ability to measure and monitor the health of populations (*1*). Epidemiology, demography and biostatistics are the key disciplines of public health.

The work of John Snow during the cholera epidemics that ravaged London in the mid-1800s is an early example of an epidemiological investigation (*2*). Snow's groundbreaking work was made possible by the registers of births and deaths maintained by local authorities in every English parish from the early 1880s. Without information on numbers of deaths from cholera and the street address of each victim, Snow's mapping of mortality in relation to the siting of water pumps would not have been possible. Thanks to the almost universal coverage of death registration in countries such as Denmark, England and Wales, the Netherlands, Norway and Sweden, the study of changing patterns of death became possible during the nineteenth century and led directly to the development of the endeavour of public health (see, for example, *3–5*).

A century and a half on, comprehensive registration of deaths and their causes is undertaken in some 78 countries covering around one-third of the world's population. The countries with comprehensive registers are generally the more developed countries in the Americas, parts of Asia and Europe. In low-income countries, efforts to better document and count basic demographic events have increased over the past two decades. Some 150 countries now produce estimates of child mortality based on empirical data collected in the previous 5 years. Advances in information technology enable large volumes of data to be processed and analysed in ever shorter periods of time. Population-based surveys have become the predominant mode for collection of data on health and progress is being tracked using indicators measurable through surveys. Examples include the 30 cluster-sample surveys of the Expanded Programme on Immunization (EPI) and the family planning and maternal and child health-oriented Demographic and Health Surveys (DHS).

[1] Measurement and Health Information Systems, World Health Organization, Avenue Appia 20, CH-1211 Geneva 27, Switzerland. Correspondence should be sent to Dr AbouZahr at this address (email: abouzahrc@who.int).

Ref. No. **04-014951**

(*Submitted: 21 October 2004 – Final revised version received: 17 January 2005 – Accepted: 31 January 2005*)

Nonetheless, there is much that remains unknown, especially about adult mortality, causes of death and the burden of suffering associated with non-fatal health outcomes.

A basic prerequisite of any health programme is its ability to state unequivocally how many deaths it will avert and what proportion of the global burden of disease it will address. To fill missing data gaps, there has been a proliferation of model-based approaches to generating global, regional and national estimates of mortality, morbidity and burden of disease. Some of these models are relatively simple, others are complex and based on debatable assumptions and underlying philosophical frameworks. The complexity of such models often masks a stark reality: most people are born and die uncounted, the reasons for their deaths unknown.

Health information systems

Counting deaths is one component of the health information system, an "integrated effort to collect, process, report and use health information and knowledge to influence policy-making, programme action and research" (*6*). The health information system is heavily biased towards quantitative data — descriptions of health status and mortality of populations over time, analysis of causation of health problems, quantification of associations between health outcomes and risk or protective factors, and assessment of the effectiveness of public health interventions. Thus it is clearly distinguishable from health-care information for professionals or more general health-related knowledge (see, for example, *7*).

The word "system" implies a connected whole or organized process. In practice, most country health information systems lack such cohesion, having developed in a piecemeal way, fashioned by administrative, economic, legal or donor pressures and are invariably highly complex. One way of simplifying the issue is to describe the dimensions of demand: who needs data and for what? of supply: the tools and methods available to generate the needed information; and of level: the level of the system at which data are generated and used.

In terms of demand, the domains that the health information system should address include:

- health determinants (socioeconomic, environmental behavioural and genetic factors) and the contextual and legal environments within which the health system operates;
- inputs to the health system and related processes including policy and organization, health infrastructure, facilities and equipment, costs, human and financial resources and health information systems;
- the performance or outputs of the health system such as availability, quality and use of health information and services;
- health outcomes (mortality, morbidity, disability, well-being, disease outbreaks and health status); and
- health inequities in determinants, coverage and use of services, and outcomes, including key stratifiers such as sex, socioeconomic status, ethnic group and geographical location.

A range of data generation methods is available including health facility data, administrative returns, household surveys, censuses, vital registration, national health accounts and health research. Matching the data item or indicator with the most appropriate and cost-effective tool for generating it is an essential function of the health information system, but is not always straightforward. Discrepancies can arise when different data collection methods are used for the same indicator. For example, indicators such as "the proportion of children immunized against measles" or "the proportion of population with access to an improved water source" can be generated both using household surveys and administrative data. However, the two sources use different tools to measure the same thing and tend to contain systematic biases in one or the other direction. Where the distinction is not clearly understood, confusion and contradiction can arise (*8*).

There is no universal formula for a precise combination of data sources that will be optimal in every setting. Much depends on existing systems that are themselves the products of history and social development. However, during the technical discussions underpinning the development of the Health Metrics Network (see below), there was consensus that all countries should aspire to a mix of the data-generation approaches described in Box 1. In some settings, certain approaches will be absent or rudimentary; elsewhere they may exist, but require strengthening. The sequence of events in establishing or strengthening data-generation approaches will depend on existing capacity and resources and a long view is essential. Vital registration systems cannot be established overnight and the capacity to implement a household survey programme effectively is built up gradually. The availability of appropriately trained human resources with analytical, numerical and statistical skills is critical.

The health information system is part of both the health system and the wider statistical system. Accountability for health-related statistics is often distributed across different line ministries (e.g. ministries of employment, agriculture and education) or agencies. Countries vary in the extent to which there are good working links between ministries of health and national statistics offices, notwithstanding the central role of the ministry of health as a generator of data and, perhaps more importantly, as a primary user of data for public health action.

Both demand for and supply of data vary in complex ways at different levels of the health system. Although populists argue that there should be no data collection other than at the level at which the data will be used, in practice, things are not so simple. For example, the census is nationally planned, but data are generated from individuals and the data are analysed and used at both the national and subnational levels. National household surveys aggregate data collected from individuals at the national level, but these data cannot generally be disaggregated to the community or even the district level. Surveillance for human immunodeficiency virus (HIV) or other notifiable diseases generates information that is used at the national and subnational levels.

At the level of individuals and communities, information is needed for effective clinical management and for assessing the extent to which services are meeting the needs and demands of communities. At the level of the district, health information enables health planners and managers to take decisions regarding the effective functioning of health facilities and of the health system as a whole. At higher levels, health information is needed for strategic policy-making and resource allocation. Although the data requirements for patient care, system management and policy-making are somewhat different they are also linked along a continuum (Fig. 1).

Box 1. **Essential sources of health-related information**

1. A decennial census
2. Continuous monitoring of births and deaths, with certification of cause of death. Where universal coverage and medical death certification is not feasible, consideration should be given to sample vital registration systems (i.e. registration of vital events in randomly selected samples of the population) coupled with verbal autopsy
3. A surveillance and response system focused on epidemic and vaccine-preventable diseases (e.g. cholera, human immunodeficiency virus (HIV) and polio) as well as on emerging diseases (e.g. severe acute respiratory syndrome (SARS))
4. A programme of household surveys designed to measure use of health care services and important household or individual behaviours, covering both demographic and health surveys and other surveys that can be used to generate health-related information, such as surveys of living standards
5. A system of service-generated data derived from facilities and patient–provider interactions covering aspects such as care offered, quality of care and treatments administered
6. Mapping of public health of facilities and services at national and district levels
7. Behavioural surveillance, focusing especially on risk factors such as smoking, unsafe sex and malnutrition
8. National health accounts
9. Financial and management information
10. Modelling, estimates and projections
11. Health research, including clinical, health systems and operations research

This continuum from patient care to strategic management implies that not everything needs to be known at every level of the system. The quantity and detail of data needed is generally greater at lower levels of the system, where decisions on the care of individuals are made, than at higher levels where broader policy-making takes place. Too often, lower level managers are required to report vast quantities of data to higher levels but rarely receive any feedback. At the same time, the information overload at higher levels is such that the data are in practice seldom used effectively.

We know what we don't know

How well are health information systems performing currently? The sobering response is not very well; far less sound information is available than is needed for effective decision-making. The failings of health information systems have been brought into sharp focus by the Millennium Development Goals (MDGs) which have a strong health component. Although the MDGs have been widely endorsed as a framework for measuring development progress, in practice, few countries have sufficiently developed health information systems to permit regular monitoring. For example, several of the MDG targets are worded in terms of reductions in cause-specific mortality, which few developing countries are able to report.

Where country data are available, they are often based on different definitions, sources and methodologies with a resulting loss of comparability both between countries and over time. The challenge of assessing trends is particularly acute. Statements as to whether countries are "on track" or "off track" are largely based on informed opinion, extrapolations and estimates. For example, only 23 developing countries have two or more data points over time for maternal mortality. Because of problems such as these, many indicators commonly cited are actually based on predictive models rather than on empirical data.

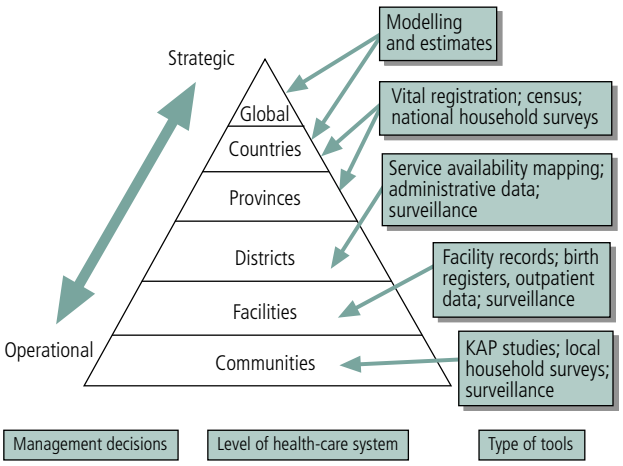
How did the public health community find itself in this sorry state? Is the problem lack of funds? Inadequate capacity? Abdication of responsibility? Failure of imagination? All of the above? Or something else entirely?

Money — necessary but not sufficient

Funding constraints offer a deceptively simple explanation for the problems experienced by the public health community. Certainly, establishing sustained and comprehensive systems of vital registration is an expensive, long-term proposition that carries little appeal for governments with short time horizons. Yet many European countries set up their registers of births and deaths well before they became economically powerful and a number of low- and middle-income developing countries have successful vital registration systems (Cuba, the Islamic Republic of Iran, Malaysia and Sri Lanka).

In any case, there is evidence that significant financial resources are being directed towards the generation of health information in developing countries. It is estimated that approximately US$ 1.25–2 billion are spent each year on different aspects of health information, of which around US$ 0.75–1 billion is spent by low-income and lower-middle-income countries and some US$ 0.5–1 billion by global organizations (McKinsey & Co., unpublished data, 2003). These are

Fig. 1. **Data needs and sources at different levels of the health-care system**



KAP = Knowledge, attitudes and practices.

WHO 05.43

considerable sums and yet data needs remain largely unmet. The problem is that the money is invested in a fragmented, duplicative and uncoordinated manner.

The extent of overlap is particularly evident in the area of household surveys and is not confined to the health sector. Over the past 20 years, national capacity to collect and process data has been strengthened, data dissemination has improved, and compliance with international standards has increased, largely thanks to international survey programmes. (These survey programmes include the Living Standards Measurement Study (LSMS) (the World Bank); Integrated Survey/Priority Survey (IS/PS) (the World Bank); Demographic and Health Survey (DHS) (United States Agency for International Development (USAID)/Macro-International Inc.); Multiple Indicators Cluster Survey (MICS) (United Nations Children's Fund (UNICEF)); Core Welfare Indicators Questionnaire (CWIQ) (the World Bank); Labor Force Survey/Child Labor Survey (LFS/CLS) (ILO) and the World Health Survey WHS (WHO)).

Because few developing countries have the financial and technical resources to implement large-scale survey programmes without external support, timing and content are to a large extent donor driven. The many examples of duplicated or conflicting data collection activities are not only wasteful, but also place a heavy burden on national statistics offices and line ministries. An effort to reduce duplication is currently under way through the International Household Survey Network (*9*).

Capacity matters

Statistical capacity-building has been identified as a core need by many countries and the efforts of the PARIS21 initiative (*10*) and The World Bank STATCAP programme (*11*) are intended to remedy the situation. Both the USAID-supported DHS and the UNICEF-supported MICS pay explicit attention to local capacity-building for data production and analysis.

Within the health sector itself, the need to build capacity for health information is often overlooked. The need for people with numeric and statistical skills to generate and analyse data is rarely mentioned in analyses of human resource requirements (*12*). The assumption seems to be that health-care workers can take on the duties of health information officers. Yet providers are understandably reluctant to divert their attention from patient care to data recording.

Recent changes in health policy and practice have added to the capacity shortfall. Health sector reform generally comprises decentralization of decision-making and resource allocation to the district level, yet neither the tools nor the capacities for information generation and analysis at this level have been sufficiently developed. Where capacity exists, it is largely concentrated centrally; the national level focus of the MDGs exacerbates this tendency.

Division of responsibility

Who is responsible for health data? The obvious response is the ministry of health. Yet ministries of health do not manage important components of the health information system. Data on determinants of health, on risk and protective factors, and on the consequences of ill-health are often to be found outside the health sector, for example, in agriculture, labour, education, water and sanitation. Although it is up to the health sector to identify the necessary information elements, much of the data collection is likely to be in other hands.

A further complication results from the fragmented nature of the health sector. Where the health care model is that of public provision and public financing, ensuring data flows is a relatively simple matter. But few countries have this "pure" state-controlled model. In most settings, public and private provision and financing coexist. Generating good data in such mixed systems can be a challenge. In theory, the regulatory authorities should oversee information reporting. In practice, regulatory power is a theoretical construct rather than a practical tool.

Responsibility is further diluted as an unintended consequence of donor actions. Driven by the demands of accountability and anxious to maximize comparability between countries, donors often support and implement their own data collection platforms. Anxiety about the availability and quality of data reported by individual countries can fuel the drive for the establishment of independent bodies primarily concerned with global monitoring (*13*). The result is separate and parallel mechanisms that respond to donor requirements rather than to the needs of country decision-makers. The problem is particularly acute when modelling is used to fill in missing data elements. Countries perceive this as an externally driven process, designed to meet donor needs and of little relevance to country action (*14*). Development partners have become more sensitive to this problem in recent years and have made greater efforts to involve country partners in modelling and projection efforts, for example in the production of national estimates of HIV prevalence.

A failure of imagination

Resources, capacity and responsibility address the data supply side of the health information equation. Equally important, but less scrutinized, is the demand side. Who needs the data and what do they need them for? An obvious answer to the question is that health policy-makers need data for decision-making. An unfortunate feature of health care systems in many parts of the world is that decisions are taken despite the absence of reliable information. In practice, decision-making in health is all too often based on political opportunism, expediency or donor demand. There is a growing awareness that this leads to inefficient and ineffective use of resources. Increasing emphasis on results-based management and performance-based funding is focusing minds on the need for sound data generated through reliable and transparent systems.

Data for decision-makers are clearly important, but the circle of stakeholders for health-related information is much larger. Access to health care and to the benefits of scientific knowledge are human rights which encompass also the right to sound health information. The public has the right to know the status of public health. Communities have the right to know why people die before their time, why they get ill, what care is available and how they can protect themselves. Health information is too important to be left to statisticians and politicians. Strengthening health information systems is also about imagining and creating a better world for all.

Health Metrics Network

Today we find ourselves in the paradoxical situation in which a large unmet need for reliable information coexists with considerable investment of effort and resources in data collection.

The United Nations Statistics Division supports countries in implementing the census and vital registration according to internationally agreed standards (*15*). WHO provides technical guidance on the use of mortality and morbidity classifications (*16*) and on the establishment of disease surveillance and response programmes (*17*). Attention has been paid in recent years to ways of making better use of routine service statistics (*18*). UNICEF has established a multi-country survey programme to monitor child health. USAID and other bilateral donors are investing substantial and growing amounts in Demographic and Health Surveys, and disease surveillance and response. New institutions and partnerships such as the Global Fund to Fight AIDS, Tuberculosis and Malaria, and the Global Alliance for Vaccination and Immunization (GAVI) offer significant new resources for health information. International agencies, bilateral donors and countries are collaborating to develop core sets of indicators for disease-specific fields and invest jointly in monitoring the leading indicators. What has been missing in these efforts has been an overall vision of a comprehensive health information system and the interlocking of its various component parts.

In 2003, a group of stakeholders came together to devise innovative solutions to the health information conundrum. Discussions resulted in a plan for the way forward, to be implemented through a new global collaboration called the Health Metrics Network, with a significant injection of venture capital from the Bill and Melinda Gates Foundation. This is the first global health initiative to have focused not on a specific disease but on a core component of the health system on which progress in disease-focused efforts is dependent. Partners in the Network have agreed to bring together their energies and expertise to provide coherent and coordinated support to the reform and strengthening of country health information systems.

The papers in this special issue of the *Bulletin* consider work undertaken during the development phase of the Health Metrics Network. They cover a wide range of health information challenges at the global and country levels. We learn about the costs of different ways of generating key indicators in the United Republic of Tanzania. We learn how China is working to improve cause of death attribution. The paper from South Africa describes the challenges of introducing data production and use into health facilities in settings with little tradition of using data for decision-making. Other papers address specific technical challenges such as how to extend coverage of vital registration systems in resource-constrained settings where deaths generally occur at home and how to deal with the striking lack of data on health inequalities.

Conclusion

It is not because countries are poor that they cannot afford good health information; it is because they are poor that they cannot afford to be without it. Good examples exist of the use of data for evidence-based decision-making leading to better health (*19*). Such examples need to be extended and scaled up. The time has come to put serious effort and resources into building health information systems that can effectively support public health. The papers in this issue describe the beginnings of change. ■

Competing interests: none declared.

Résumé

Les systèmes d'information sanitaire piliers de la santé publique

La prise de décisions de santé publique repose de manière déterminante sur la disponibilité en temps voulu de données solides. Le rôle des systèmes d'information sanitaire est de produire, d'analyser et de diffuser de telles données. Dans la pratique, il est rare que ces systèmes fonctionnent méthodiquement. Résultant de l'action de diverses forces historiques, sociales et économiques, ils sont complexes et fragmentaires et ne répondent pas convenablement aux besoins. Les donateurs internationaux dans le domaine de la santé ont une large responsabilité dans cette situation, car ils ont accordé la priorité aux besoins urgents en matière de données plutôt qu'à la mise en place de moyens nationaux d'action à long terme. Cette option se traduit malheureusement par l'incapacité de la plupart des pays à produire les données nécessaires pour suivre les progrès dans la réalisation des objectifs de développement pour le millénaire. Ce problème exige des solutions globales ; il ne suffira probablement pas d'allouer des fonds si cette action ne s'accompagne pas d'une assistance durable au développement des systèmes nationaux et, s'agissant des donateurs, de l'instauration d'obligations de rendre des comptes plus contraignantes et d'une plus grande responsabilisation. Le Health Metrics Network, un effort collaboratif mondial en cours, a pour objectif de contribuer à la mise en place de telles solutions dans les pays qui en ont le plus besoin.

Resumen

Los sistemas de información sanitaria, pilares de la salud pública

La adopción de decisiones de salud pública depende de forma decisiva de la posibilidad de disponer oportunamente de datos sólidos. Los sistemas de información sanitaria, cuya finalidad es generar, analizar y difundir tales datos, en la práctica rara vez funcionan sistemáticamente. Producto de fuerzas históricas, sociales y económicas, son sistemas complejos, fragmentados e insensibles a las necesidades. Los donantes internacionales que actúan en el campo de la salud tienen bastante responsabilidad en ese problema, pues han priorizado las necesidades urgentes de datos antes que la creación de capacidad en los países a más largo plazo. Una consecuencia tristemente patente de ello es la incapacidad de la mayoría de los países para generar los datos requeridos para vigilar los progresos hacia los Objetivos de Desarrollo del Milenio. Las soluciones al problema deben ser integrales; la mera aportación de dinero será probablemente insuficiente a menos que se acompañe de un apoyo sostenido al desarrollo de los sistemas de los países, unido a una mayor asignación de responsabilidades y rendición de cuentas por los donantes. La Red de Sanimetría, una iniciativa de colaboración de reciente andadura, tiene por objeto ayudar a encontrar soluciones de ese tipo para los países más necesitados.

Arabic

References

1. Beaglehole R, Bonita R, Horton R, Adams O, McKee M. Public health in the new era: improving health through collective action. *Lancet* 2004;363:2084-6.
2. Snow J. *Snow on cholera*, second edition. London: Oxford University Press; 1936.
3. Marfarlane A, Mugford M. *Birth counts: statistics of pregnancy and childbirth.* London: Her Majesty's Stationery Office; 1984.
4. De Brouwere V, Tonglet R, Van Lerberghe Wl. La "Maternité sans Risque" dans les pays en développement: les leçons de l'histoire. [Safe motherhood in developing countries: lessons from history.] Antwerp: ITG Press; 1997. Studies in Health Services Organization and Policy, 6. In French.
5. Evans T, Stansfield S. Health information in the new millennium: a gathering storm? *Bulletin of the World Health Organization* 2003;81:856.
6. Lippeveld T. Routine health information systems: the glue of a unified health system. *Keynote address at the Workshop on Issues and Innovation in Routine Health Information in Developing Countries, Potomac, March 14-16.* Washington, DC: JSI; 2001.
7. Godlee F, Pakenham-Walsh N, Mcayiyana D, Cohen B, Packet A. Can we achieve health information for all by 2015? *Lancet* 2004;364:295-300.
8. Carraro L, Khan S, Hunt S, Rawle G, Robinson M, Antoninis M, et al. Monitoring the millennium development goals: current weaknesses and possible improvements. Glasgow: Oxford Policy Management, Department for International Development; 2003.
9. International Household Survey Network. *A global partnership to make better use of survey data for development results.* Washington DC: IHSN. Available from: www.surveynetwork.org
10. Development Co-operation Directorate. *Partnerships for Statistics in Development in the 21st Century PARIS21.* Paris: Organisation for Economic Co-operation and Development (OECD). Available from: http://www.paris21.org/
11. World Bank. *Statistical Capacity Building Program. A new lending program to support more efficient and effective statistical systems in developing countries.* Washington, DC: World Bank. Available from: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/SCBEXTERNAL/0,,contentMDK:20100851~menuPK:244214~pagePK:229544~piPK:229605~theSitePK:239427,00.html
12. World Health Organization. *High-level forum on the Health Millennium Development Goals, Geneva, January 2004.* Geneva: WHO. Available from: http://www.who.int/dg/lee/speeches/2004/mdgforum_geneva/en/
13. Murray CJL, Lopez AD, Wibulpolprasert S. Monitoring global health: time for new solutions. *BMJ* 2004;329:1096-100.
14. World Health Organization. *The world health report 2000 – Health systems: improving performance.* Geneva: WHO; 2000.
15. United Nations Statistics Division, Department of Economic and Social Affairs. *UNSTATS.* New York: United Nations Statistics Division. Available from: http://unstats.un.org/unsd/demographic/standmeth/principles/default.htm
16. World Health Organization. *The WHO Family of International Classifications.* Geneva: WHO. Available from: http://www.who.int/health_topics/classification/en/
17. World Health Organization. *The Global Outbreak Alert and Response Network.* Geneva: WHO. Available from: http://www.who.int/csr/outbreaknetwork/en/ and http://www.who.int/health_topics/epidemiologic_surveillance/en/
18. Lippeveld T, Sauerborn R, Bodart C. *Design and implementation of health information systems.* Geneva: World Health Organization; 2000.
19. Tanzanian Ministry of Health. *The Tanzania Essential Health Interventions Project (TEHIP). An overview.* Tanzania: Ministry of Health; 2002.