The Development of a Database on Disasters

DEBARATI G. SAPIR and CLAUDINE MISSON

The need for systematic and organised data for disaster response and management has been an increasing concern of both international and national relief agencies. Until recently, each agency approached the problem in a relatively *ad hoc* manner collecting the information at the time of the emergency. As a result, data were incomplete, outdated or unuseable for a variety of reasons, even if better quality information existed. The time pressure to respond quickly for fund raising or relief planning was usually paramount and the quality and availability of information suffered. This reflected a lack of preparedness by a community of professionals who have, in recent years, actively promoted disaster preparedness.

Why has such a database not been constructed or compiled earlier, especially since many of the data were available? The principal answer lies in the reactive nature of disaster management. Action was always taken post-facto and pre-planning was rarely considered a funding or policy priority. However, with the increase in the number of disasters requiring external assistance and the recurrent nature of some disasters (e.g., floods in Bangladesh), donors (governments, bilateral and multilateral aid agencies) are increasingly aware of the need for a rational approach to disaster relief. The need to improve the efficacy of emergency missions, including needs assessments, has been highlighted with particular reference to the requirement for systematically organised information.

The Emergency Preparedness and Response Office of the World Health Organisation noted the importance of a planned response and the rationalisation of relief, as well as the need for speedily provided country briefs for emergency missions. The Office undertook an examination of the feasibility and the conceptual design of systems to answer these needs.

DISASTER EVENTS DATABASE (EM-DAT)

The Centre for Research on the Epidemiology of Disasters (CRED) has explored possible designs for, and the feasibility of, a system of databases for the management of and response to disasters globally. An Emergency Management Information System (EMIS) was designed to provide rapid and accurate information for eventual use by the World Health Organisation and other agencies involved in disaster preparedness and response. The original design of the system is presented in Figure 1.

In the second phase, CRED developed Subsystem B in greater detail with three databases that were of greatest general need in preparedness, rapid response planning and emergency briefing (Figure 2). These were: (a) disaster events and their essential characteristics; (b) key information related to disasters or relevant to planning response by country; and (c) disaster-related institu-

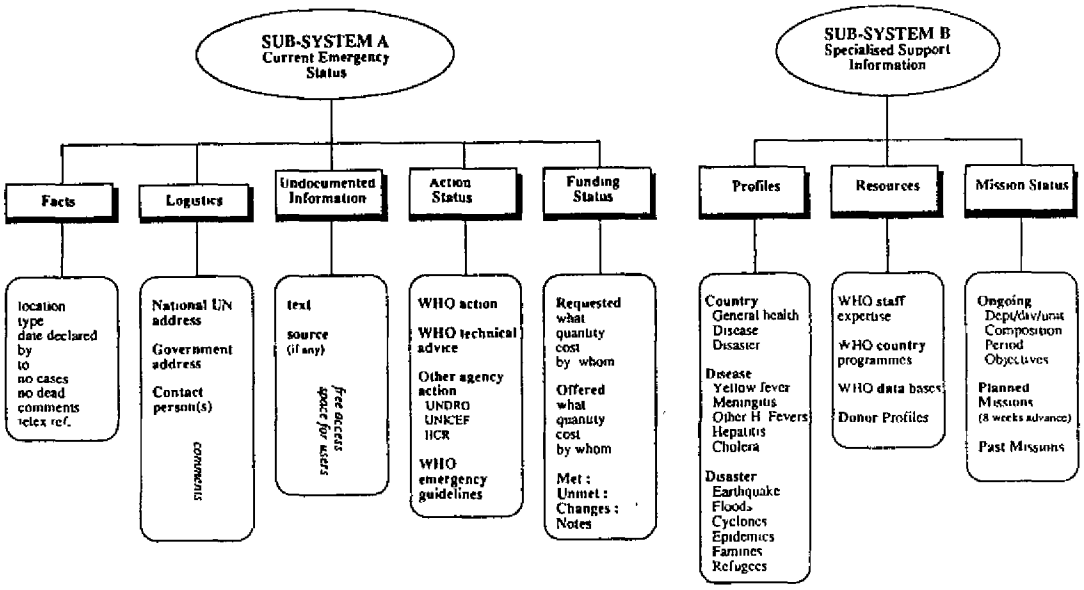

FIGURE 1 Emergency Management Information System: outline and contents

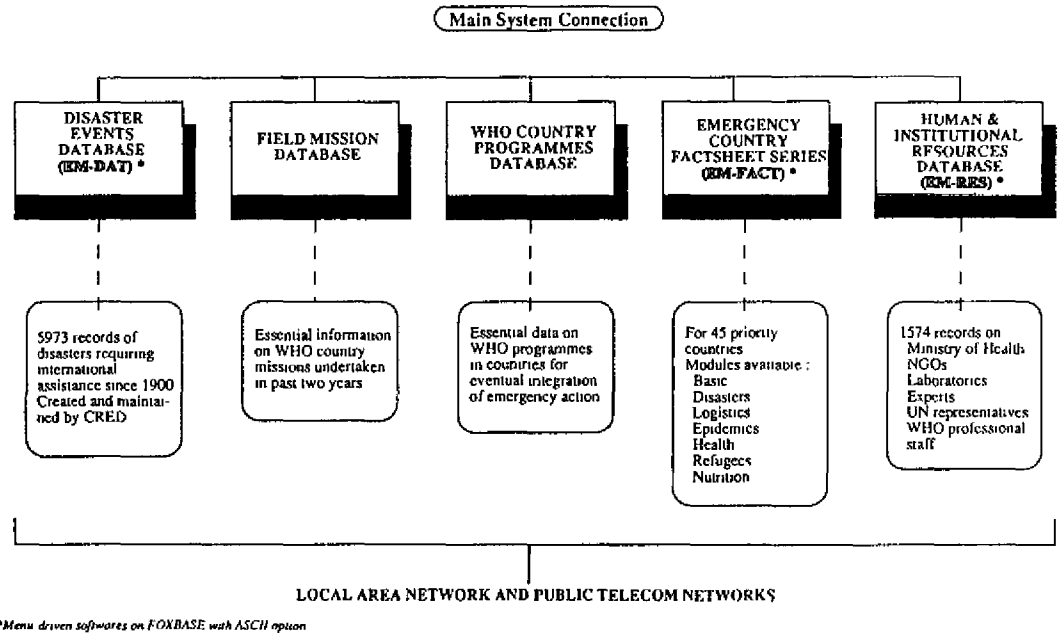

FIGURE 2 *Emergency Management Information System: layout for databases on sub-system B*

tional and human resources by country. This article presents the technical aspects of the first of these databases — Disaster Events Database or EM-DAT.

DESCRIPTION AND STRUCTURE

The database was developed using resources provided by the University of

Louvain and CRED. The existing computer support and documentation collection at CRED provided the basis for the start-up. Staff supported by the University provided daily management and data entry services for the system. Additional staff support was provided by CRED. The project currently involves four staff members (part-time) with two external assistants for short-term programming and research.

EM-DAT is currently operational with its own menu for updates, modifications and retrievals. It has 5870 records and is being checked case by case for redundancies, inconsistencies and completion of missing data.

METHODOLOGICAL ISSUES IN EM-DAT OPERATION

The definition of a disaster is much debated and various versions are used by different agencies, largely according to their needs. According to the World Health Organisation,

> A disaster is any occurrence that causes damage, ecological disruption, loss of human life, deterioration of health and health services on a scale sufficient to warrant an extraordinary response from outside the affected community or area. (WHO Manual, Section XV.4 article 20, 1/3/89.)

For the purposes of the database, a working definition had to be formulated to help the data manager and technical staff to decide what would constitute a valid case for entry. The following definition, worked out in collaboration with UNDRO, is currently in use:

> An interruption in time and space of normal processes beyond the coping capacity of the community, causing death, injury or homelessness, direct material losses and/or negative economic impact. The interruption can be either sudden or gradual onset.

This definition encompasses the following categories.

(1) Sudden natural disasters.
   - — disasters of geological origin, e.g. earthquakes, landslides, avalanche, volcanic activity, tsunami;
   - — disasters of hydro-meteorological origin, e.g. floods, high wind (tropical cyclones, hurricanes and typhoons), local storms, extremes of weather (heat or cold); and
   - — insect infestation.

(2) Natural disasters of gradual onset:
   - — e.g. drought, famine (food shortage plus malnutrition and their consequences), epidemic disease

(3) Man-made disasters:
   - — violent mass conflicts, e.g civil war, major riots, uprisings;
   - — political origins, not necessarily violent, e.g. displaced persons (mass evacuation, refugees, expellees);
   - — accidents, e.g. transportation, food or insecticide poisoning, explosions; and
   - — chemical accidents, e.g. factory explosion, nuclear accident but not including slow pollution.

The categorisation of disasters significantly affects the validity and accuracy of the information gathered. The typology used for the EM-DAT is shown in Table 1. It has many weaknesses and will be reviewed and refined. Many questions can be raised about the typology but it is easy to lose sight of what is practical and feasible in maintaining the database at the optimal level of detail without entailing a major investment of staff and materials.

The general approach is to have as many clearly defined categories as possible. Although this detailed categorisation means that the information is more useful for specialists, it poses logistical problems. It increases the number of categories so that the retrieval of data usually requires the selection of a number of categories, thus complicating procedures for retrieval and use. Hurricanes, typhoons and cyclones, for

TABLE 1
List of Disaster Types recorded in EM-DAT

Accident	Earthquake	Insect Infestation
Avalanche	Epidemic	Landslide
Chemical accident	Famine	Storm
Civil Strife	Fire	Tsunami
Cyclone	Flood	Typhoon
Displaced persons	Heat/cold wave	Volcano
Drought	Hurricane	

example, are disasters created by high winds and are largely differentiated by their place of origin. In simple terms, all of these phenomena are tropical cyclones. Typhoons and hurricanes are local terms used for those originating in the Pacific (China Sea region) and the Caribbean seas respectively. Thus, classifying them separately unnecessarily sub-divides a group which would normally be considered as a whole. Any retrieval, in this form, would require selecting all of the different sub-types that are relevant and adding them to obtain the total. In this case, it may make it simpler for most users if all these events were entered under one category, namely high winds. Obviously, in the comments section, the exact nature, specific location and path of the event would be provided for the interested user.

The problem is less obvious for other disasters. 'Displaced persons', for example, is listed as a category separate from famine, civil unrest and drought. But displaced persons are a consequence of the disaster, albeit often the principal one, not the disaster itself. From the user's point of view, the event should be recorded in the form most appropriate for the database's objectives — i.e., policy formulation, prioritisation and relief efficiency. It was decided, therefore, that 'displaced persons' should be entered as the event itself and that its cause should be entered in the comments section.

There are other ambiguities within this category. Currently, it includes only expellees (those persons expelled from their country of residence, such as Indians from Uganda); returnees (those persons obliged to return to their home countries, such as Mozambicans from South Africa) and refugees (those persons obliged to leave their home country because of religious, racial or political persecution). There are other groups, however, which also belong, potentially, to this category, such as persons forced to move due to the scarcity of food and water, or those who have moved due to civil war. Sometimes these groups move within their own country and do not cross international borders and are therefore not refugees or expellees. The question remains whether all types of displaced persons should be recorded within the same category or whether there should be separate categories for those outside and inside their national borders.

The problem of typology is currently addressed by classifying a disaster according to the list shown in Table 1. All other relevant information is then entered in the comments section. This allows for the deletion and creation of categories as necessary.

There are several other important issues related to the typology of disasters that require conceptual clarification. The main consideration in deciding on a typology for a database is that there is sufficient detail for it to be meaningful without being so detailed as to make it over complicated for the wide user community.

CRITERIA FOR ENTRY

The following criteria are used for entering an event in the database.

> A disaster has to have killed 10 or more persons or affected 100 or more persons. An international appeal for assistance, however, takes first precedence for entry, even if the first two criteria are not fulfilled.

Where there is conflicting information, priority is given to data from the government of the affected country, followed by UNDRO and OFDA, in that order. If, however, two of the above sources provide the same estimate, this figure would then have priority over the third, regardless of its source. All other estimates are registered in the comments section and any research or analytical work should be undertaken using all the information available on that record. It should be emphasised that this does not reflect the value placed on the quality of data. It is well known that, especially in disasters, most reporting sources have a vested interest in the numbers they report. Figures are inflated or deflated according to socio-political considerations.

Certain events require different criteria for entry. For displaced persons, any displacement of 2000 or more persons (but not necessarily any killed) would be eligible for entry into the database. Similarly, for drought and famine, the threshold for entry is 2000 or more affected persons. Fires present another special case. There are industrial or urban fires and bush or forest fires. Some of the latter can go on for months or even years. Fires and accidents must be sudden to be eligible and the criteria for entry are those for natural disasters (i.e., 10 dead or 100 affected). As for chemical accidents, all reported events are entered, even if there are no dead or affected.

Finally, dates have been a source of ambiguity. The declared date for an event such as a famine is on the one hand necessary but on the other almost meaningless. Neither famines nor population displacements occur on one particular date. Civil strife occasionally erupts on a precise date but it also builds up over time. Similarly, there is rarely an exact date for the outbreak of an epidemic. For these events, the declaration of an emergency by an appropriate body has to be used as the date for entry and, as always, further information is entered in the comments section.

One might have thought that, for natural sudden-onset disasters (cyclones, earthquakes, flash floods), the date of the occurrence and date of declaration would be identical. In practice, however, there are often discrepancies between these two dates. The declaration of an emergency varies widely according to who declares it and the government concerned. Sometimes external agencies declare an emergency before national authorities and sometimes it is the opposite. The declaration often takes place one or two days after the actual occurrence. For natural disasters, the solution is relatively easy. The physical occurrence of earthquakes, cyclones and other sudden natural events are generally well monitored and scientifically accurate information is available. The problem is much more difficult for the slow onset and man-made disaster, where declaration is as much a political decision as an issue of fact.

STATISTICAL INFORMATION

There are five variables for statistical information: number of persons killed, number of persons injured, number of persons homeless, number of persons affected and the monetary value of losses sustained. Each of these has definitional ambiguities that require clarification.

The number of persons killed as an immediate consequence of the disaster should include all confirmed dead, all missing and all presumed dead. If the event is entered immediately after its actual

occurrence, the number of 'missing' is often not included in the estimated dead, particularly if the scource provides preliminary figures. Estimates of dead currently include all confirmed dead in the immediate post-impact period. Later estimates of dead will include those missing for a lengthy period. With no international standards, this definition varies from source to source. Currently, CRED checks each entry for clarification on this point.

'Persons injured' is meant to include all those with physical injuries requiring medical attention or hospitalisation. It remains to be decided whether ambulatory care, provided at the site of the disaster, should be included, as well as what constitutes medical care It should be remembered that care provided by para-medics and other auxiliary medical personnel is often the main source and form of care in the immediate aftermath. Whether all persons receiving these services would qualify as injured is also an issue. A simpler alternative could be only to include physical trauma, such as fractures and crushes This has the advantage of clearly defining the scope of the variable, but it excludes important health effects that are not traumatic. In addition, the classification of severe psychological stress and its effects remains unresolved. Injuries are currently registered only for those disasters that may entail acute physical injury, such as earthquakes, cyclones and avalanches. Famines and epidemics do not have injuries recorded, only numbers dead or affected.

Homelessness is defined as loss of personal residence. Another discrepancy in reporting units arises, in the case of the affected and homeless, where the units may be individuals or families. The current procedure is to convert all figures into individuals by multiplying the families by the average family size of the region (5.6 for developing countries and 3.5 for developed countries). For a less crude estimate one could use the specific average family size for the country in question but, given the quality of the base figure, such refinements may be unnecessary.

Defining 'persons affected' is extremely difficult because there is a wide variety of parameters. In famines and civil strife, for example, major structural breakdowns in the social and economic framework are the essential elements that are responsible for the human impact. An estimate of the population affected by such a breakdown must always be relatively broad.

Disparities in units of reporting have also posed a problem. The monetary value of damages may, for example, be provided in U.S. dollars or the local currency. On the one hand, it is simpler to leave the currency unit in the form in which the source has reported it and convert it only when the event is of interest. On the other hand, this precludes comparisons or computations that are often required by users, such as the total annual financial damage caused by a type of disaster, or the total losses sustained by a country. The uniform conversion of all reported losses into one currency unit (US dollars or Swiss Franc) may be easily programmed in but it should be standardised to a base index value (consumer price index) for it to be truly comparable.

Insurance and re-insurance companies would make a good source for damage estimates but there can be serious biases in their data. They would, for example, only report insured damages, which do not reflect the real damage caused. In addition, in developing countries, personal property is only insured by the wealthier members of the community and therefore the use of insurance data can be seriously misleading for relief and rehabilitation purposes.

CONCLUSIONS AND FUTURE DEVELOPMENTS

The creation of a valid and useable database for disasters is fundamentally dependent on the clarity of the concepts and definitions

underlying it. In a field that has only recently engaged a professional and scientific interest, this conceptual development is in its infancy. Individual disciplines that are traditionally involved in disaster research, such as geology, engineering, metereology and other earth and atmospheric sciences, have their own definitions and information systems. The issue becomes complex when a multi-disciplinary database has to be compiled that brings together, in an optimal mix, data from different disciplines with different conceptual bases.

Experience so far in putting together the EM-DAT database has given every indication that such an enterprise is feasible, but with one very important qualification: the collaboration of different disciplines and the establishment of a working network is essential for its success. An informal association with some of the major figures in the field of disaster data collection has already been established by CRED as the project has developed.

Two main directions are envisaged for the development of the project through the extension of its scope to include the following information groups.

(a) *Donor response:* This concerns amounts and types of relief. The exact layout and detail of data to be entered remains to be defined. The projected use of the data will determine the nature of its contents.

(b) *Physical characteristics of the event:* These are its physical indices (e.g. Richter, Mercalli and other readings, wind speeds and paths), location and precise time of occurrence. Ambiguities in the declared date of the event will be avoided by using only such scientifically based data.

It is clear that substantial research is needed into the design of methodologies for estimating the numbers of dead or in need of food or life-line services in disasters. This is important, not only in order to gain more reliable information, but also to improve disaster response management. The adoption of standard methods and agreed definitions by major international agencies would improve the quality of estimates quickly, with minimum investment.

Note

We should like to acknowledge the collaboration of the World Health Organisation, UNDRO, Office of Foreign Disaster Assistance (USAID), Munchener Ruckversicherungs-Gesellschaft (Munich), and Compagnie Suisse de Reassurances — SIGMA. J. Duppenthaler (WHO), J.B Miller (World Meteorological Organisation), A. Velasquez (Department of Geography, University of Yale, Columbia) and Ms. R. Maier (UNDRO) merit special thanks for their detailed comments and interest in the development of this database.

Debarati G. Sapir
Claudine Misson
Centre for Research on Disaster Epidemiology (CRED)
University of Louvain School of Public Health
B-1200 Brussels
Belgium