BMC Public Health

RESEARCH

Open Access

What makes an epidemic a disaster: the future of epidemics within the EM-DAT International Disaster Database

Margo Tonnelier[1*], Damien Delforge[1], Regina Below[1], Juan Armando Torres Munguía[2], Claude Saegerman[3], Valentin Wathelet[1], Joris A. F. van Loenhout[1,4] and Niko Speybroeck[1]

Abstract

Background Reporting on and monitoring epidemics is a public health priority. Several initiatives and platforms provide epidemiological data, such as the EM-DAT International Disaster Database, which has 1525 epidemics and their impact reported since 1900, including 892 epidemics between 2000 and 2023. However, EM-DAT has inconsistent coverage and deficiencies regarding the systematic monitoring of epidemics data due to the lack of a standardized methodology to define what will be included under an epidemic disaster.

Methods We conducted a sequential study that included an online survey of experts in infectious diseases, public health emergencies, and related data, followed by committee discussions with disaster experts. This approach aimed to identify appropriate definitions and entry criteria for archiving disease outbreak events.

Results The survey had 21 respondents from universities and international organizations, with experts primarily specialized in infectious disease surveillance. Experts agreed that epidemics should be considered as disasters. Experts cited challenges in defining epidemic thresholds. However, they proposed pathogen-based criteria and agreed that disruption to society, especially to the healthcare system, serves as a determinant of epidemic disasters. The experts favored deaths and confirmed cases as key indicators, alongside suggestions on refining EM-DAT's entry criteria and improving epidemic impact assessment.

Discussion This article offers valuable insights into epidemic disasters, a topic previously underdefined in the literature, thereby enhancing understanding for policymakers and public health professionals.

Keywords Epidemic definition, Epidemiology of Disasters, Biological Hazards, Emergency Events Database (EM-DAT), Disaster Health Impact, Reporting Methodology, Epidemics data sources

*Correspondence:
Margo Tonnelier
margo.tonnelier@uclouvain.be
[1] Centre for Research On the Epidemiology of Disasters (CRED), Institute of Health and Society (IRSS), UCLouvain, 1200 Woluwe-Saint-Lambert, Belgium
[2] Faculty of Economic Sciences, Georg-August-Universität Göttingen, Göttingen, Germany
[3] Research Unit of Epidemiology and Risk Analysis Applied to Veterinary Science (UREAR-ULg), Fundamental and Applied Research for Animals & Health (FARAH) Center, Faculty of Veterinary Medicine, University of Liège, Liege 4000, Belgium
[4] Department of Epidemiology and Public Health, Sciensano, Brussels 1000, Belgium

Background

The United Nations (UN) defines a disaster as a serious disruption of the functioning of a community or a society [1] and the International Epidemiological Association as "an event that disrupts the normal conditions of existence" [2]. Accordingly, epidemics have such a potential and can lead to important human, material, economic, or environmental losses resulting from interacting conditions of exposure, vulnerability, and capacity [1]. The Dictionary of Epidemiology defines epidemics as "the occurrence of cases of an illness, specific health-related behavior, or other health-related events clearly above normal expectancy" [2] in a specific community and a specific period. Infectious diseases, owing to their modes of acquisition and the potential for rapid changes in transmission, contribute primarily to the occurrence of epidemics [3]. However, the term has also been extended to encompass non-communicable conditions, such as obesity [4] or opioid addiction [5].

Disease surveillance organizations use a variety of classifications for infectious diseases depending on their needs (e.g., detailed understanding or efficient response to epidemics) or specialties (e.g., epidemiologists, health geographers, or emergency coordinators). The most widely used classifications of infectious diseases are (i) anatomic, by bodily region or system affected [6]; (ii) by transmission mode [7]; (iii) aetiologic (causal agent) [6]; (iv) by geographic occurrence [7]; and (v) by severity (epidemiological rate) [7]. The Centers for Disease Control and Prevention (CDC) uses a hybrid approach [8], integrating the severity, geographical distribution, and characteristics of different diseases to define levels of response to epidemics. Similarly, the World Health Organization (WHO) classification [9] focuses on the level of response required for emergencies, taking into account severity and geographical spread. On the other hand, the International Federation of Red Cross and Red Crescent Societies (IFRC) adopts an approach primarily based on the type of transmission [10]. The United Nations Office for Disaster Risk Reduction (UNDRR) classifies epidemics as biological disasters in the Hazard Information Profiles (HIPs) [11]. Lastly, the INFORM (Index for Risk Management) initiative introduces a risk assessment tool focusing on infectious diseases grouped by mode of transmission [12].

Behind the reporting and monitoring of epidemics, the primary objective is to supply the necessary information for understanding the disease's spread and multiple impacts (e.g., health, economy, society). This concern gained prominence during the COVID-19 pandemic,[1] as the public, media, and politicians closely tracked epidemiological indicators daily. The pandemic underscored the absence of standardized data collection methods and disparities in data quality. Four pivotal categories of challenges can be identified in this regard [13]: (i) epidemiological factors, e.g., missing data, reporting delays [13]; (ii) data structure, e.g., date formats and place names lacking standardization [13]; (iii) data dissemination, e.g., divergent estimates and differences in variable data structures between reporting sources [14]; (iv) lack of availability and accessibility [15]. Beyond epidemics, disaster monitoring faces similar challenges of gathering accurate and timely data, constrained resources, and interoperability issues among diverse systems (e.g., official reports, real-time monitoring, hospital data). Integrating human, animal, and environmental data [16] necessitates harmonization, particularly in addressing public health challenges and anticipating secondary post-disaster health issues.

Numerous international data platforms play a crucial role in reporting and tracking epidemics [17]. A distinction can be made between early warning systems (potential epidemics), real-time surveillance (ongoing epidemics), and archiving services (historical epidemics). The WHO provides various tools, including the early warning systems Global Health Observatory (GHO) [18], offering health-related data, and the Disease Outbreak News (DON) [19], which delivers detailed reports on diseases and countries [20, 21]. Additionally, the WHO's Go.Data software [22] and EWARS-in-a-box [23] are applications for field data collection during epidemics, facilitating real-time entry and analysis, whose data are combined on the WHO Health Emergency Dashboard [24] and in the Assessment Capacities Project (ACAPS) [25], both real-time surveillance systems. The CDC contributes to surveillance through their early warning system Morbidity and Mortality Weekly Report (MMWR) [26], covering diverse health topics, including the occurrence of unusual clusters of illnesses, and the European Centre for Disease Prevention and Control (ECDC) managing the ECDC Atlas [27] with detailed indicators such as reported cases and mortality rates. ProMED-Mail [28], a reporting system, primarily provides early warnings and updates about outbreaks and emerging health threats worldwide [29–31]. Other initiatives include the Global Infectious Diseases and Epidemiology Network (GIDEON) [32], ReliefWeb [33], HealthMap [34],

[1] Pandemic: "An epidemic occurring over a very wide area, crossing international boundaries, and usually affecting a large number of people."(2).

BioCaster [35], the IFRC's Go platform [36], and Empres-i platform [37] (Additional file 1).

EM-DAT is an archiving service and a global database that collects data on disaster occurrence and human and economic impact from 1900 to the present date [38] (Additional file 1). EM-DAT follows a systematic methodology in reporting disasters – not only epidemics – that meet at least one eligibility criteria: 1) 10 or more people killed; 2) 100 or more people affected; 3) a declaration of a state of emergency or a call for international assistance [38]. While data collection has improved for many disasters related to natural hazards since EM-DAT's inception in 1988, epidemics have remained sub-optimally categorized as disasters. To improve the classification of epidemics and the entry of epidemics data in EM-DAT, we undertook an expert consultation study. This study aimed to generate a list of criteria for collecting data on and reporting epidemic disasters in an operable but systematic way. This information can be used to improve the quality and scope of the EM-DAT database.

The study focused on the following six operational objectives: (i) Develop a standardized definition for epidemics, (ii) Establish inclusion criteria for reporting epidemics, (iii) Identify essential human impact metrics for reporting epidemics, (iv) Develop a standardized classification for epidemics, (v) Specify the key informational elements required to comprehensively describe epidemic, and (vi) Create a prioritized list of reliable data sources for epidemics.

Methods

Data sources

EM-DAT is compiled from various sources of information, including UN agencies, non-governmental organizations, insurance companies, research institutes, and press agencies [38].

Pandemics, which are comprehensively covered elsewhere, are excluded from EM-DAT due to the system's limited capacity for continuous updates. Between 1900 and 2023, EM-DAT recorded 1,525 epidemic disaster events, with 892 between 2000 and 2023. Of the epidemic events recorded, 26 (2%) have no information on the total number of deaths or people affected, and 142 (9%) outbreak events are not classified into disease subtypes. Also, 290 (19%) events have no information on symptoms or disease names. While the data structure of EM-DAT requires the country of the event to enter the impact figure, only 20% ($n = 304$) of the events have more detailed information on the sub-national location of the outbreak. The current classification system in EM-DAT categorizes epidemics aetiologically, i.e., based on the biological nature of the causal agent, while allowing for a general "Infectious disease" category when not explicitly presented in the sources. The non-classified events (i.e., the causal agent remains unidentified) account for about 2.7 million deaths (the cumulative death toll from all epidemics stands at 9.6 million individuals) [39]. Of the 1,525 events recorded, a state of emergency was declared for 52 (3%) and an appeal for international assistance for 7 (0.5%) [39].

Study design

To ensure the usefulness and relevance of disaster impact data for infectious disease experts utilizing platforms like EM-DAT, we used a sequential approach to establish standardized definitions for collecting epidemic disaster data.

First, we conducted an expert consultation through an online survey. Second, findings from the online survey were presented in an online extended EM-DAT scientific committee specializing in disaster and epidemic data. These discussions enabled experts to review the preliminary results, provide additional input, and refine key areas needing alignment.

Identification and selection of experts

Experts for the online survey were selected as follows: background investigation of known users and partners of EM-DAT, including insurance companies, non-governmental agencies, and the United States Agency for International Development; a literature review of scientific articles published in the last 15 years and reviewing the corresponding author list; additional publications by selected authors and their co-authors; identification of international groups dedicated to the surveillance and control of infectious diseases; investigation of websites of universities with a research department on infectious diseases. Inclusion criteria specified that experts must have at least two related publications in the last 15 years. Experts with strict technical or operational backgrounds were included based on recommendations obtained through initial participants via the *snowball* method [40], with a maximum of three experts per department to maintain a diverse panel.

The final list consisted of 78 experts, whom we pre-classified according to the following areas of expertise: epidemiological surveillance, control and prevention of infectious diseases, emerging infections, integrated epidemic analysis management, medical specialists in zoonoses, and infectious and tropical diseases.

The survey results were presented to an extended EM-DAT Scientific Committee. The Scientific Committee's primary objective is to uphold rigorous data quality standards by providing strategic oversight and refining protocols for data handling, validation, and analysis, including evaluating the integration of international standards, terminology, and key indicators. The committee comprises experts in disaster databases and data management, and three infectious disease specialists from the selected pool of 78 experts were also invited to contribute their insights.

Questionnaire development and pre-testing

To ensure clarity and relevance, two sociologists specializing in health research reviewed the survey, refining it for an international and multidisciplinary audience. This pre-test ensured alignment with the study's six objectives. The final questionnaire consisted of 23 questions, including 15 open-ended questions designed to gather in-depth qualitative insights (Additional file 2).

Data collection

We conducted the survey using the Qualtrics© software (version March 2023). We generated unique participation links, facilitating the sending of personalized reminders. These participation links were sent to each expert by email in April 2023 with May as a deadline to respond.[2] We sent two reminders via email to participants in May.

Data analysis

We processed qualitative data in two stages: a first phase of repeated and cross-referenced readings to identify significant recurrences of terms and define common categories, and a second phase of reading to classify responses. The sequential analysis enables categories of responses to emerge progressively [41]. Following this analysis, we transformed qualitative data into categorical outcomes to identify trends.

We performed descriptive analyses of the quantitative results in R (version 4.3.1). Owing to our small sample size and to gain more robustness, we bootstrapped the data to obtain credibility intervals (CIs) using 2500 random samples of size 21 [42].

Data presentation

On June 25, 2024, the online survey results were presented at the EM-DAT Scientific Committee meeting. The 15-min oral presentation was followed by a 25-min discussion, introduced to participants as a phase to consolidate and refine responses. Recommendations were integrated and further explained.

[2] Three participants explicitly requested a time extension, so the survey remained open for them until June 6, 2023.

Table 1 Demographic characteristics of survey respondents

Gender		
(n = 21)	**n (%)**	**95% credibility interval**
Female	11 (53)	[29; 76]
Male	10 (47)	[24; 71]
Age group		
(n = 21)	**n (%)**	**95% credibility interval**
18–34	1 (5)	[0; 14]
35–64	18 (86)	[71; 100]
>65	2 (10)	[0; 24]

Results

Stage 1: online survey

In this section, the quantitative results are presented and supplemented by respondents' comments (qualitative results). A selection of verbatim responses is provided in Additional file 3.

Demographic profile of survey respondents

The number of survey respondents was 21. The sample indicates a balanced gender representation (Table 1). The majority (86%) belong to the 35–64 age group.

Surveillance of infectious disease epidemics emerges as the predominant area of specialization, constituting 63% of the sample (Fig. 1A). Universities and international organizations were well represented in the sample, with 33% and 38%, respectively (Fig. 1B). The overall breakdown by world region reveals a strong presence of experts from Europe (48%) (Fig. 1C).

Perspectives on epidemic reporting in EM-DAT

A consensus (95%) emerged among respondents that epidemics should be categorized as disasters (Table 2). The need for a common approach to public health emergencies and disasters, considering environmental change and climate-related events, is emphasized in the qualitative information. Respondents who argue for the inclusion of all epidemics vs. only those of a certain size (Table 2) cite their potential magnitude and stress the importance of understanding patterns beyond numerical thresholds. Possible political considerations over the terminologies 'epidemics' and 'outbreaks' are also put forward by respondents as an argument to include all infectious events. Depending on the country and the political stakes, an epidemic might not have the same threshold.

While 57% preferred including all epidemics irrespective of scale, 70% noted that the current EM-DAT entry criteria do not fully reflect the actual severity of an epidemic and may be influenced by political considerations,

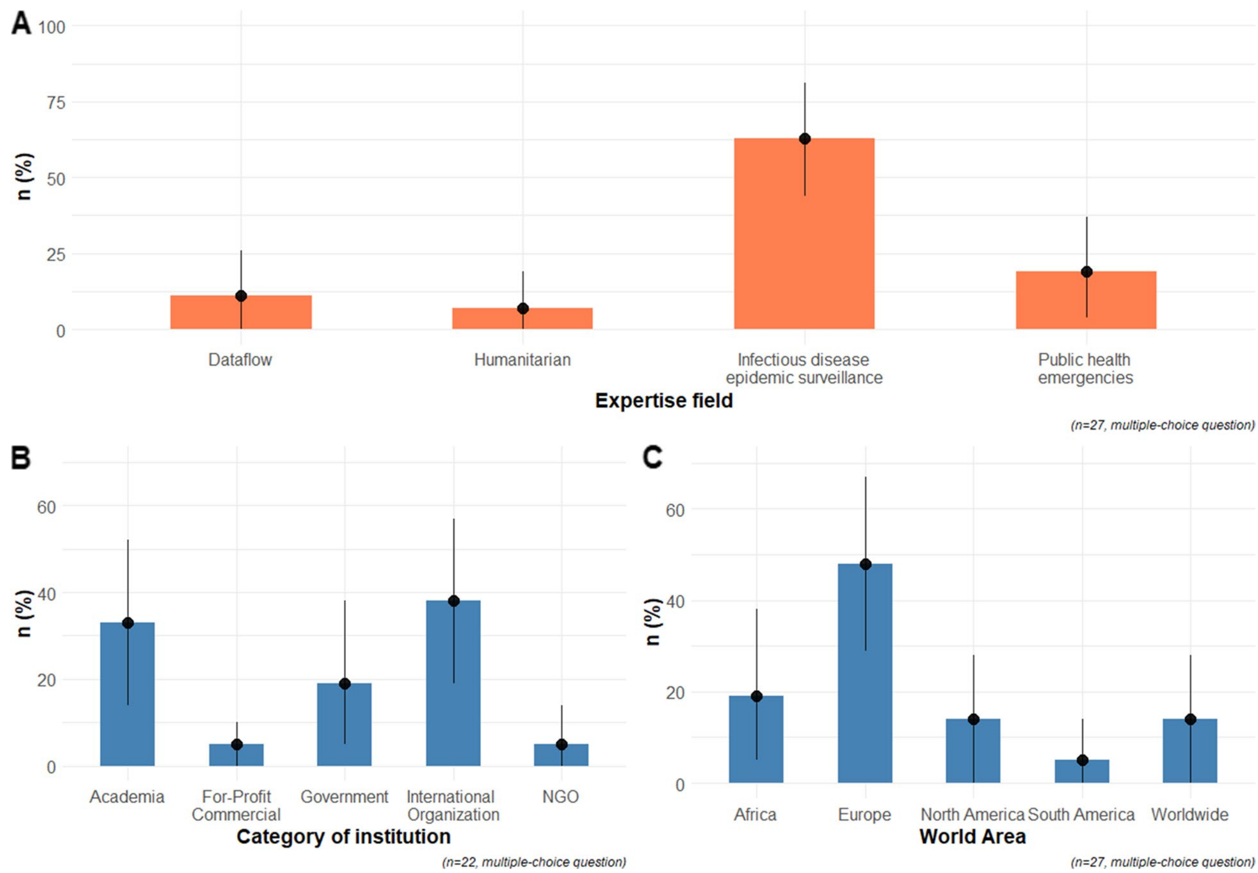

Fig. 1 Professional information on survey respondents. **A** Fields of expertise of survey respondents. **B** Categories of institutions where the survey respondents are employed. **C** World areas where the survey respondents are involved. *For each result, the percentages along with their corresponding 95% credibility intervals are reported*

Table 2 Epidemic reporting in EM-DAT

Is an epidemic a disaster?		
(n = 21)	**n (%)**	**95% credibility interval**
Yes	20 (95)	[86; 100]
No	1 (5)	[0; 14]
Should EM-DAT include epidemics of a certain size and not include outbreaks and pandemics?		
(n = 21)	**n (%)**	**95% credibility interval**
Yes	9 (43)	[24; 67]
No	12 (57)	[33; 76]
Can we use the current EM-DAT entry criteria for epidemics?		
(n = 20)	**n (%)**	**95% credibility interval**
Yes	6 (30)	[10; 50]
No	14 (70)	[50; 90]

particularly when declaring a state of emergency or calling for international aid. Most respondents (58%) favored pathogen-dependent criteria (Fig. 2). In their comments, the experts added that using epidemic thresholds established for specific diseases may enable a more tailored, evidence-based approach and that the lethality

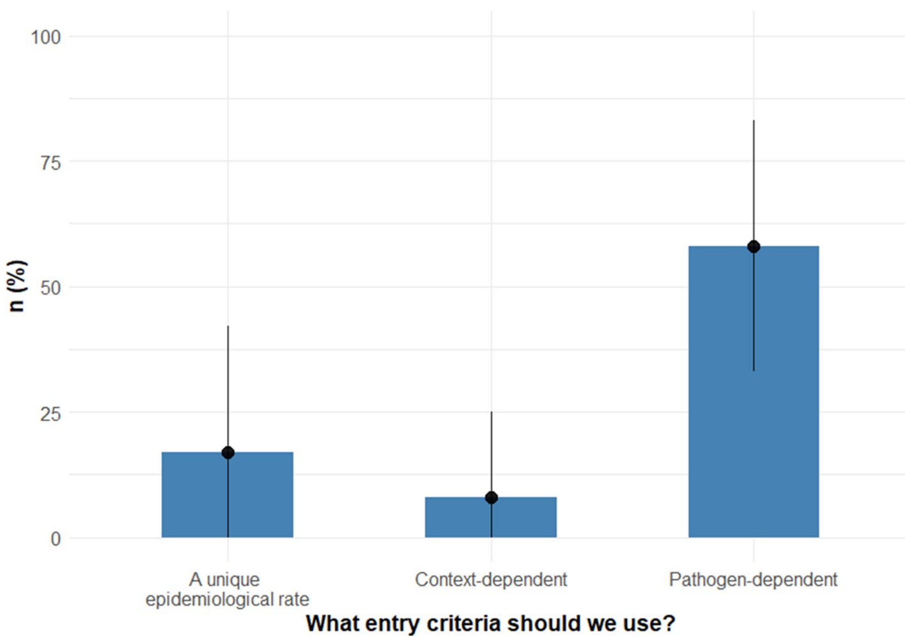

Fig. 2 Entry criteria of an epidemic. *For each result, the percentages along with their corresponding 95% credibility intervals are reported*

and transmission characteristics of the pathogen are important.

When defining an epidemic as a disaster, respondents emphasized that the infectious event must disrupt the functioning of society. Disruption may refer to the healthcare system, limited capacity of public health infrastructure, widespread impact across multiple countries or regions, and anticipated longevity resulting in lasting economic. In contrast, adaptability and returning to normal or stabilized conditions signal the end. Most participants stressed the need to use local data and, if possible, hospital data. The experts suggested that specific criteria, such as the absence of cases for several consecutive days, can be used to declare the end of the epidemic for diseases such as Cholera or Ebola.[3]

Current EM-DAT human indicators and classification

Regarding the adaptation of the EM-DAT term 'Affected' to epidemics, a majority (53%) of the respondents proposed referencing confirmed cases, underlining the importance of quantifying people diagnosed (Fig. 3A). About adapting the term 'Dead' to epidemics, Fig. 3B highlights the respondents' suggestion to consider the number of deaths as the main criterion (40%) rather than case fatality ratios (20%). The 'Others' category includes all the isolated, non-representative responses.

[3] Due to their high virulence.

Most survey respondents (71%) felt that the classification currently used in EM-DAT could be improved (Table 3). Regarding alternative classifications, participants suggested adopting a classification based on the type of disease transmission. Some experts expressed concerns about classifying epidemics solely by the type of pathogen, suggesting a need for a purely biological approach that includes environmental and chemical hazards.

Other valuable data for describing an epidemic

Based on the responses, several key indicators should be considered to quantify the impact of an epidemic (Fig. 4). According to the survey participants, infectious data should include the basic reproduction number R_0, the attack rate, the case-fatality rate, and the median time from symptom onset to recovery.

Some experts suggested recording psychological sequelae to give an idea of the duration of the epidemic's impact. Tracking disability-adjusted life years (DALYs) and quality-adjusted life years (QALYs) related to mental health was reported to give insight into the broader impact on health, including mental well-being. A particular focus on child and maternal health data was also mentioned.

Data on the availability of and access to essential medicines and vaccination coverage would help to understand the epidemic's management capacity. According to participants, assessing the capacity of healthcare facilities and surveillance systems measures the ability to

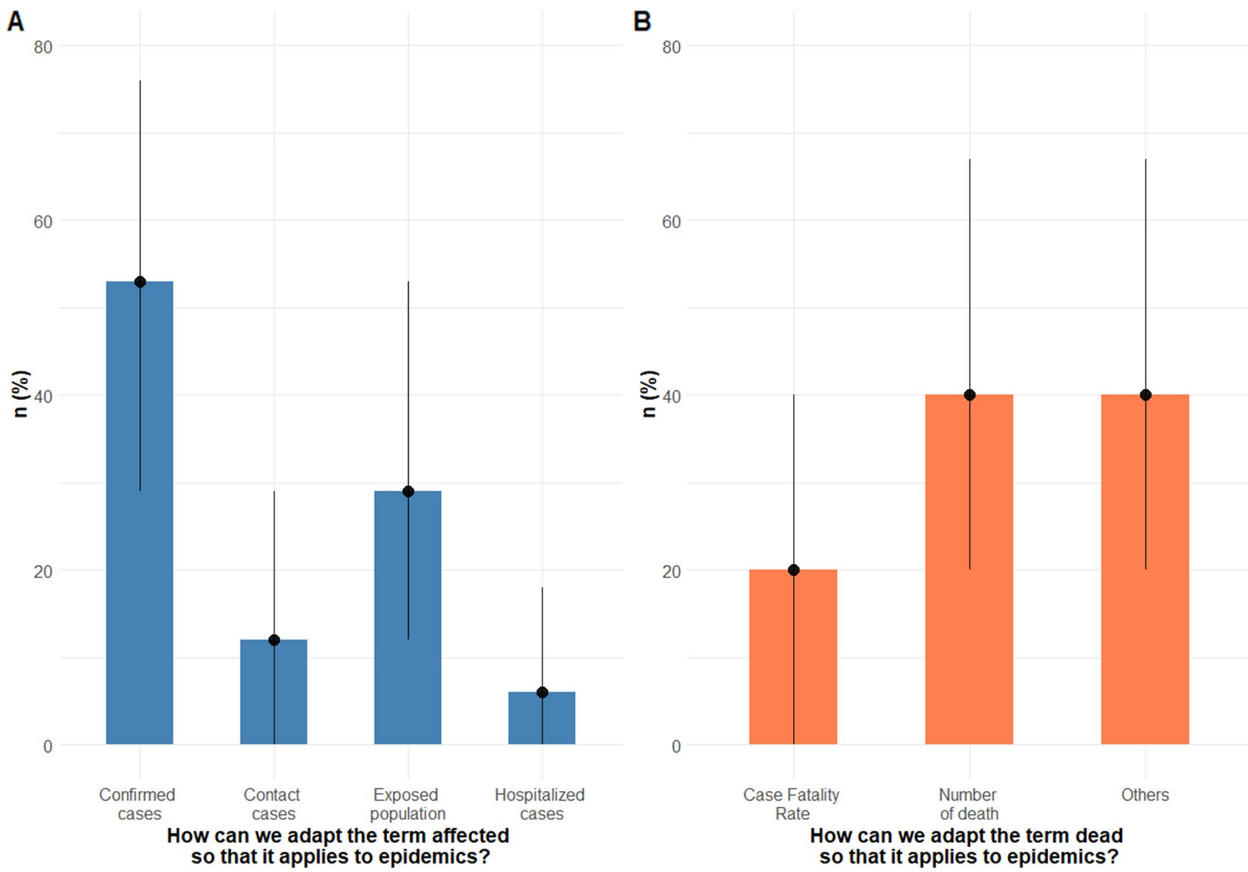

Fig. 3 Human indicators for epidemics in EM-DAT. **A** Expert s' proposal to adapt the term 'affected' as defined in EM-DAT to epidemics. **B** Experts' proposal to adapt the term 'dead' as defined in EM-DAT to epidemics. *For each result, the percentages along with their corresponding 95% credibility intervals are reported*

Table 3 Epidemic classification in EM-DAT

Can we improve the current classification of epidemics in EM-DAT?		
(n = 14)	**n (%)**	**95% credibility interval**
Yes	10 (71)	[43; 93]
No	4 (29)	[7; 57]

cope with increased demand, detect the epidemic, and respond effectively. This assessment can cover both personnel and the capacity of health services to detect new cases, identify transmission chains, manage severe cases, and provide timely treatment and preventive measures.

Data sources for epidemics

The experts identified WHO Disease Outbreak News, ACAPS, WHO Health Emergency Dashboard, and ProMED as valuable sources for outbreak reports and detailed information on emerging health threats. EWARS-in-a-box and similar tailored tools designed for managing outbreaks in humanitarian contexts were considered reliable sources. Hospital electronic medical records (EMRs) and integrated outbreak analytics systems were further mentioned as data collection and analysis avenues.

Stage 2: scientific committee

Following the survey, discussions with the EM-DAT Scientific Committee were held to evaluate and refine the findings, leading to specific focus areas.

Inclusion of epidemics and pandemics

The committee endorsed the classification of epidemics as disasters. However, the inclusion of pandemics was debated; the committee recommended recording pandemics retrospectively to improve data completeness, with potential oversight from WHO and ReliefWeb.

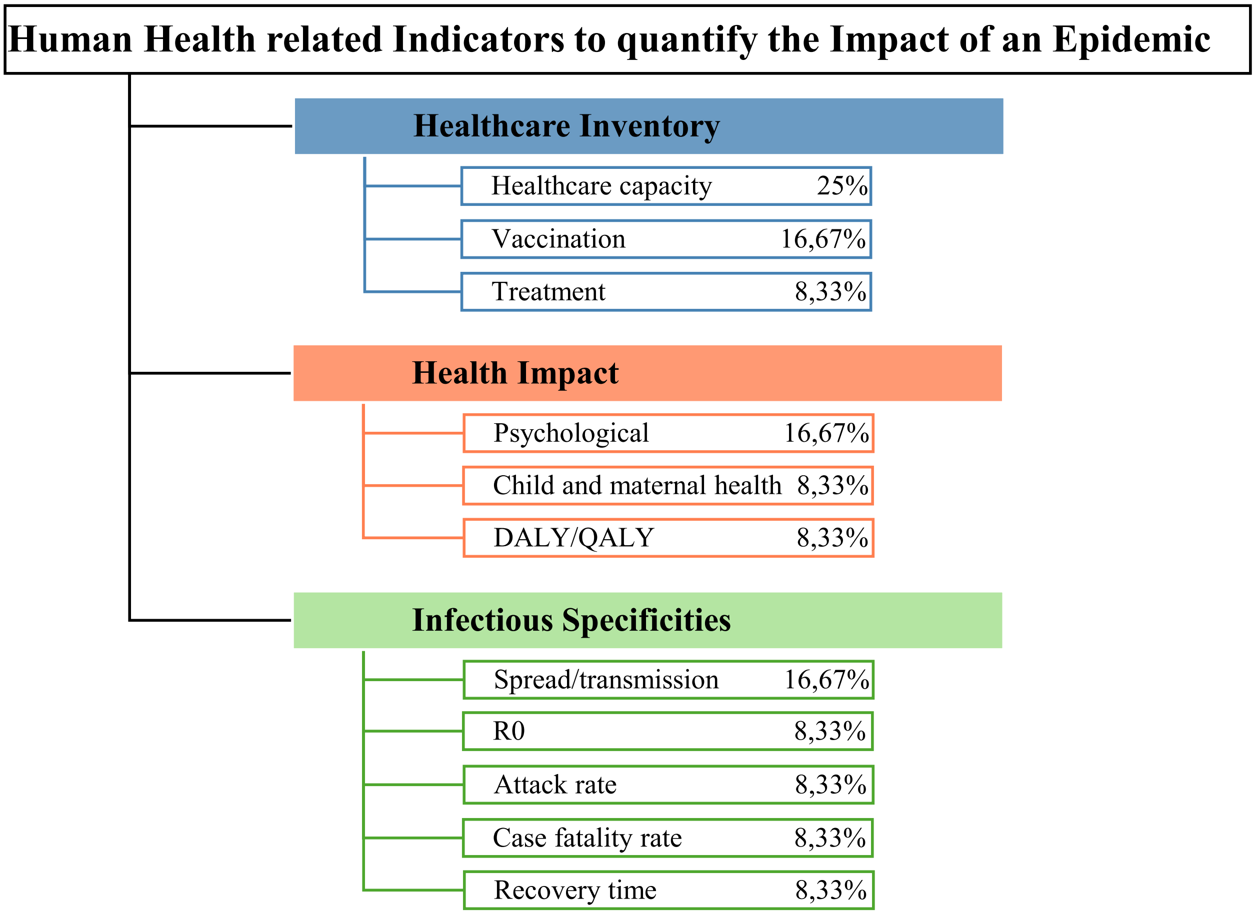

Fig. 4 Human health indicators describing epidemics impact

Inclusion criteria and flexibility

The committee upheld the current inclusion criteria but suggested providing a detailed explanation of any limitations, especially for diseases with varying thresholds. For epidemics with extended durations, they proposed aggregating cases to better align with WHO and ECDC methodologies, and it emphasized including subnational location data and accurate start and end dates for each epidemic.

Defining key indicators

Confirmed cases and deaths were reaffirmed as core indicators. The committee emphasized the importance of tracking the proportion of tested individuals to capture testing capacity and potential underreporting. Mortality measurement remained challenging, prompting a call for consistency between reported deaths and statistical mortality analyses.

Classification adjustments

The committee suggested omitting detailed subcategories within EM-DAT, while allowing users to subclassify epidemics themselves for their purposes based on the pathogen type. They recommended referencing classifications from WHO and ECDC to improve standardization across reporting systems.

Discussion

Define a disaster epidemic

The results from the survey indicate that an epidemic, consolidated by the opinion of a Scientific Committee, as defined by the occurrence of a health-related infectious spread beyond normal expectations, meets the disaster definition of a "serious disruption of the functioning of a society" [1]. While there is a consensus in the literature [1] and among experts that an epidemic is a disaster, there is more uncertainty regarding epidemiological thresholds. Having a clear operational definition of an epidemic disaster is instrumental in epidemic monitoring, as the lack of harmonization in data collection processes has been cited as an issue in epidemic reporting [13]. General definitions of an epidemic [3, 43] provide no precise statistical threshold above which an infectious disease is an epidemic. According to the survey respondents, the epidemiological thresholds should depend on the pathogen and its factors, including virulence and specific disease contexts. Prolonged healthcare system strain, workforce depletion, and supply chain disruptions are highlighted in recent studies as key indicators that an outbreak has reached a disaster level, warranting classification within EM-DAT and similar databases [44].

Furthermore, the acceptable risk level is sensitive to social contexts, with possibly different norms. This is linked to the significant variations in cultural contexts, justice, and symbols between different societal groups [45]. The accessibility of health infrastructures, particularly in developing countries, is also a factor to consider when establishing epidemiological thresholds, as these infrastructures play a key role in preventing and containing epidemics [46].

EM-DAT's scope for epidemic monitoring

The importance of standardized terminology and criteria is highlighted across multiple studies [47, 48], which show how differing definitions for 'epidemics' and 'outbreaks[4]' can lead to inconsistencies in global reporting and data collection efforts. For instance, this research [47, 48] stresses the necessity for context-sensitive parameters in epidemic thresholds, suggesting that disease-specific factors like transmission dynamics, host immunity, and healthcare infrastructure greatly influence an epidemic's spread and impact and should be integrated into database entry criteria.

The question of whether EM-DAT should only monitor epidemics rather than outbreaks and pandemics divided the survey panel. A slight majority rejected an EM-DAT monitoring exclusively limited to epidemics, wishing to include outbreaks and pandemics, which was strongly supported by the Scientific Committee. Instead, survey respondents advocated for expanding EM-DAT's use to include early warning in addition to its archiving functionality [38]. This perspective is not applicable to EM-DAT, whose scope remains the archiving of disaster events for analysis purposes and whose human and financial resources do not allow real-time encoding.

Based on inputs from most survey respondents, an epidemic should be considered a disaster when there is a disruption of the functioning of society, in particular, healthcare structures, starting at the local level. In addition, experts think that adapting the health system's capacity to treat new cases signals the end of an epidemic disaster. In that sense, a hospital's White Plan declaration [49] could serve as a pivotal marker in identifying the onset of an epidemic disaster. A White Plan is a system that defines a healthcare establishment's crisis management procedures. When the White Plan is lifted, it implies that the healthcare system can manage the number of new cases without disrupting day-to-day care. This may ask for the study of the applicability of a hospital White Plan declaration or equivalent (declaration on a local scale as recommended in the survey results) as criteria for EM-DAT entry.

Secondly, regarding epidemic classification, the transmission type is the most reported factor to be used as a basis in the elaboration of a classification for epidemics (Additional file 1). Data from survey responses and committee discussions highlight opportunities to refine EM-DAT's epidemic classification. A majority of survey participants suggested basing classifications on disease transmission type rather than focusing solely on pathogens. The Scientific Committee recommended simplifying EM-DAT's categories, allowing optional subclassification by pathogen and aligning with WHO [19, 22] or CDC [26] standards.

Quality considerations about human indicators

Accurate mortality estimations are an ongoing challenge. It could be improved by routinely tracking and reconciling excess mortality rates alongside reported deaths, thereby addressing discrepancies often caused by underreporting or variable testing capacities [47]. Regarding the choice of human impact statistics, the survey respondents propose considering case-fatality rates and the count of deaths for mortality data. In addition, the number of confirmed cases and the exposed population could replace the term 'Affected', which was considered less suitable for epidemics. The Scientific Committee reaffirmed confirmed cases and deaths as essential, emphasizing the inclusion of testing capacity metrics to better capture the risk of underreporting.

Yasobant et al. [48] highlight how mental health sequelae, alongside physical health impacts, provide a more holistic representation of an epidemic's toll. According to the experts surveyed, other indicators may further help to quantify the impact of an epidemic, such as the duration of the impact, the scale of transmission (attack rate, R_0), access to essential medicines, the capacity of health facilities and surveillance systems, and the exposure and vulnerability of the population. These data are less readily available in existing databases (Additional file 1) and would make the recording difficult. Vaccination is already included in EM-DAT data [50], but including other information requires further investigation and evaluation. The location of clusters (as an indicator of geographical dissemination) seems feasible (Additional file 1), as does

[4] Outbreak: "an epidemic limited to localized increase in the incidence of a disease, e.g., in a village, town, or closed institution"(2).

the recording of hospitalization rates (as an indicator of severity). R_0 rates, important estimates for epidemiologists, seem less often available. Although R_0 is a valuable tool for understanding disease transmission dynamics, it is rarely included in existing databases. Accessing and utilizing R_0 often requires additional efforts in data collection and analysis. However, incorporating R_0 could enhance the ability to compare and classify epidemics. It is becoming increasingly necessary to explore the feasibility of creating a unified database that compiles R_0 rates by disease [51].

Advancing epidemic data collection

The diversity of expertise related to infectious diseases within the panel allows for a well-rounded perspective in defining standardized epidemic disaster criteria. The sequential approach—incorporating an online survey and committee discussions—enabled ongoing feedback and refinement at each stage. However, specific weaknesses should be acknowledged. Despite using this sequential approach, a clear consensus was not achieved, highlighting that a Delphi study might have been more effective for consensus-building on such complex topics. The involvement of experts on a voluntary basis limited the number of participants, hence the generalizability of the findings. We recommend extending this survey to obtain more respondents in broadened contexts (e.g., stakeholders). Furthermore, the study did not explicitly address data quality and consistency challenges, which are essential for ensuring the reliability and usefulness of the EM-DAT database.

The perspective of using existing databases and the automation of data collection processes are important for platforms like EM-DAT to capture and archive epidemic information efficiently in real time. Additional literature underscores the potential for advanced data analytics and machine learning to enhance real-time epidemic reporting. A 2024 study [52] advocates for the use of these technologies to incorporate data from multiple sources, such as social media, electronic health records, and mobile health platforms. Furthermore, Garriga et al. [53] stress that collaboration with international health agencies, particularly the WHO and the ECDC, could enhance consistency in terminology and data reporting protocols. Such partnerships could facilitate the sharing of supplemental datasets, allowing disaster databases to reflect epidemiological trends and responses with greater precision.

Harmonization may be a challenge, which can potentially be countered by automatic data exchange protocols between databases. This asks to rely, among other factors, on agreed standard concepts, units, definitions, formats, and classifications to effectively translate information from other databases into the terms and definitions of EM-DAT. While this may be time-consuming, costly, and challenging, these costs can be significantly minimized by employing a systematic automated approach for data extraction, collection, and collation. For instance, leveraging artificial intelligence tools, such as Natural Language Processing (NLP), can establish a common language [54] to parse and reconcile how definitions are applied and operationalized across different classifications of infectious diseases (e.g., anatomical, transmission mode, aetiological[6]). This approach allows for a deeper integration and interpretation of classification systems, enabling the production of more comprehensive and harmonized information.

To enhance EM-DAT's capabilities in identifying and recording epidemic events, a systematic approach involving the definition of a data structure, identification of information sources, development of an automated machine learning system [54], implementation of an NLP algorithm for detection and recording, data enrichment from statistical sources, and ensuring flexibility and reusability are essential. The NLP algorithm could adapt to predefined and standardized data structures, creating a structured dataset from unstructured information [55]. By harnessing automation to extract relevant and complementary information from other external sources, EM-DAT may also enhance its internal readability by facilitating the establishment of logical linkages between different types of disasters that coincide in space and time [21]. For example, it can correlate epidemics with specific natural hazards, thereby providing a more comprehensive understanding of the interconnectedness of multi-hazard disasters.

Conclusions

This study supports characterizing epidemic disasters as significant health-related infectious events that spread beyond typical expectations, leading to substantial disruptions in societal function, especially by overwhelming local healthcare capacities. However, reaching a consensus on specific epidemiological thresholds to define an epidemic disaster remains challenging, highlighting the ongoing difficulty of achieving uniform definitions. The importance of clear operational definitions cannot be overstated, particularly given the lack of harmonization in data collection processes, which complicates accurate epidemic reporting. Findings from the survey and committee review indicated that inclusion criteria should be tailored to pathogen-specific factors, including virulence, transmission, dynamics, healthcare infrastructure resilience, and the socio-economic context. While challenges persist in identifying precise entry criteria for epidemic reporting platforms like EM-DAT, there is a consensus on the need for comprehensive indicators to quantify the epidemic's impact. Key indicators, ranging from confirmed cases to transmission dynamics and healthcare capacity, offer a holistic perspective on epidemic severity and societal resilience. As efforts to refine epidemic classification and reporting mechanisms continue, collaboration among stakeholders and ongoing research will be essential for improving the understanding and response to epidemic disasters.

Abbreviations

ACAPS	Assessment Capacities Project
CDC	Centers for Disease Control and Prevention
CIs	Credibility Intervals
DALYs	Disability-Adjusted Life Years
DON	Disease Outbreak News
ECDC	European Centre for Disease Prevention and Control
EM-DAT	Emergency Events Database
EMRs	Electronic Medical Records
GHO	Global Health Observatory
HIPs	Hazard Information Profiles
IFRC	International Federation of Red Cross and Red Crescent Societies
INFORM	Index for Risk Management
IRSS	Institute of Health and Society
MMWR	Morbidity and Mortality Weekly Report
NLP	Natural Language Processing
QALYs	Quality-Adjusted Life Years
R_0	Basic reproduction number
UN	United Nations
UNDRR	United Nations Office for Disaster Risk Reduction
WHO	World Health Organization

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-024-21026-2.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements

The authors would like to acknowledge the contribution of the survey participants: Diana Manilla Arroyo, Marie-Amélie Degail, Carlos Machado de Freitas, Cristina Estrada, Flavio Finger, Mary Jackman, Mory Keita, Justin Lessler, Anne Perrocheau, Peter Piot, Jonathan Polonsky, Ruwan Ratnayake, Wbeimar Alejandro Sanchez, Maya Schaerer, Egbert Sondorp, Thérèse Umuhoza, Sophie Vanwambeke, Paula Vasconcelos, Richard B Yapi, and the other participating experts.

The authors would like to acknowledge the contribution of the Scientific Committee participants: Aglaé Jezequel, Albert Kettner, Ilan Noy, Hamish Patten, Jonathan Polonsky, Sylvain Ponserre, Juan-Armando Torres-Munguia, Iria Touzon Calle, and Shiomi Yumi.

The authors thank Messaline Fomo Fodom from the University of Antwerp, Gilles Louis and Bernard Voz from the University of Liège for reviewing the survey form.

The authors would like to thank Lieven Desmet of UCLouvain for his support in the boostrapping stages.

Authors' contributions

All authors have made substantial contributions either to the writing of this paper or to prior work associated with epidemic data collection in the EM-DAT database. M.T. drafted the first versions of the survey, selected the experts consulted, managed the survey, processed the data, and drafted the manuscript. D.D. supervised all stages of the paper, from survey to editing. C.S. supported methodology in developing the stages of the online survey. R.B. has been managing the EM-DAT data content since 1992. R.B. and V.W. revised the survey form and tested its interface. J.A.T.M. contributed to the discussion on automating the inclusion of outbreak information in EM-DAT. J.A.F.L. and N.S. revised the paper, and N.S. supervised the survey process. All authors reviewed the manuscript.

Funding

The EM-DAT project, within the Université catholique de Louvain (UCLouvain), receives funding from USAID/BHA [ref no.39772OFDA20CA00072].

Data availability

EM-DAT epidemic data can be downloaded for non-commercial use from https://public.emdat.be/. The questionnaire used for this research is attached as Additional file 2. The anonymised answers to the questionnaire used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The processing of survey participants' data complies in all respects with the obligations of the General Data Protection Regulation (GDPR) in the current version of the OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. the official text of the Regulation (EU) 2016/679 is available here https://gdpr-info.eu/. All study participants completed the survey online and provided informed consent to participate in the published study. Two participants asked not to be named.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 20 July 2024 Accepted: 9 December 2024
Published online: 03 January 2025

References

1. UN. Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction. United Nations; 2016 Dec [cited 2022 Aug 25] p. 41. Report No.: A/71/644. Available from: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N16/410/23/PDF/N1641023.pdf?OpenElement.
2. Porta MS, Greenland S, Hernán M, Silva I dos S, Last JM. A Dictionary of Epidemiology. New York: Oxford University Press; 2014. p. 377.
3. CDC NERD Academy. What is a pandemic - Lesson Plan. [cited 2024 Jan 2]. Available from: https://www.cdc.gov/scienceambassador/documents/module-1-defining-a-pandemic-lesson-plan.pdf.
4. Temple NJ. The Origins of the Obesity Epidemic in the USA–Lessons for Today. Nutrients. 2022;14(20):4253.
5. Wilkerson RG, Kim HK, Windsor TA, Mareiniss DP. The Opioid Epidemic in the United States. Emerg Med Clin. 2016;34(2):e1-23.
6. Zumla A, Ustianowski A. Tropical Diseases: Definition, Geographic Distribution, Transmission, and Classification. Infect Dis Clin. 2012;26(2):195–205.
7. Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, Bhatt S, et al. Global mapping of infectious disease. Philosophical Transactions of the Royal Society B: Biological Sciences. 2013;368(1614):20120250.
8. WHO. ICD-11 for Mortality and Morbidity Statistics. 2023 [cited 2024 Jan 9]. ICD-11 for Mortality and Morbidity Statistics. Available from: https://icd.who.int/browse11/l-m/en#/http%3a%2f%2fid.who.int%2ficd%2fentity%2f1435254666.
9. WHO. World Health Organization. 2024 [cited 2024 Jan 9]. WHO grading of public health events and emergencies. Available from: https://www.who.int/emergencies/grading.
10. IFRC. The International Federation of Red Cross and Red Crescent. 2024 [cited 2024 Jan 9]. Diseases. Available from: https://epidemics.ifrc.org/volunteer/disease-per-type.

11. UNDRR. PreventionWeb. 2023 [cited 2024 Jan 9]. The Disaster Risk Reduction (DRR) Glossary. Available from: https://www.preventionweb.net/drr-glossary/hips.
12. The European Commission Disaster Risk Management Knowledge Centre. European Comission. 2022 [cited 2023 Feb 5]. INFORM Risk Methodology. Available from: https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Risk/Methodology.
13. Badker R, Miller K, Pardee C, Oppenheim B, Stephenson N, Ash B, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health. 2021;6(5):e005542.
14. Majumder MS, Cusick M, Rose S. Measuring concordance of data sources used for infectious disease research in the USA: a retrospective data analysis. BMJ Open. 2023;13(2):e065751.
15. Humblet MF, Vandeputte S, Mignot C, Bellet C, De Koeijer A, Swanenburg M, et al. How to Assess Data Availability, Accessibility and Format for Risk Analysis? Transbound Emerg Dis. 2016;63(6):e173–86.
16. Panel (OHHLEP) OHHLE, Adisasmito WB, Almuhairi S, Behravesh CB, Bilivogui P, Bukachi SA, et al. One Health: A new definition for a sustainable and healthy future. PLOS Pathogens. 2022;18(6):e1010537.
17. Warsame A, Murray J, Gimma A, Checchi F. The practice of evaluating epidemic response in humanitarian and low-income settings: a systematic review. BMC Med. 2020;18(1):315.
18. WHO. World Health Organization. 2024 [cited 2024 Jan 9]. Global Health Observatory. Available from: https://www.who.int/data/gho.
19. WHO. World Health Organization. 2024 [cited 2024 Jan 9]. Disease Outbreak News (DONs). Available from: https://www.who.int/emergencies/disease-outbreak-news.
20. Carlson CJ, Boyce MR, Dunne M, Graeden E, Lin J, Abdellatif YO, et al. The World Health Organization's Disease Outbreak News: A retrospective database. PLOS Global Public Health. 2023;3(1):e0001083.
21. Torres Munguía JA, Badarau FC, Díaz Pavez LR, Martínez-Zarzoso I, Wacker KM. A global dataset of pandemic- and epidemic-prone disease outbreaks. Sci Data. 2022;9(1):683.
22. Hollis S, Stolow J, Rosenthal M, Morreale SE, Moses L. Go.Data as a digital tool for case investigation and contact tracing in the context of COVID-19: a mixed-methods study. BMC Public Health. 2023;23(1).
23. PAHO. Pan American Health Organization. 2024 [cited 2024 Jan 9]. Early Warning, Alert and Response System. Available from: https://www.paho.org/en/health-emergencies/health-emergency-information-and-risk-assessment/early-warning-alert-and.
24. WHO. WHO Health Emergency Dashboard. 2024 [cited 2024 Feb 29]. WHO Health Emergency Dashboard. Available from: https://extranet.who.int/publicemergency.
25. ACAPS. ACAPS. 2024 [cited 2024 Jan 9]. ACAPS Data. Available from: https://www.acaps.org/en/data.
26. CDC. Centers for Disease Control and Prevention. 2024 [cited 2024 Jan 9]. Morbidity and Mortality Weekly Report (MMWR). Available from: https://www.cdc.gov/mmwr/index.html.
27. ECDC. European Centre for Disease Prevention and Control. 2023 [cited 2024 Jan 9]. Surveillance Atlas of Infectious Diseases. Available from: https://www.ecdc.europa.eu/en/surveillance-atlas-infectious-diseases.
28. International Society for Infectious Diseases. ProMED-mail. 2024 [cited 2024 Jan 9]. Home - ProMED. Available from: https://promedmail.org/.
29. Desai AN, Anyoha A, Madoff LC, Lassmann B. Changing epidemiology of Listeria monocytogenes outbreaks, sporadic cases, and recalls globally: A review of ProMED reports from 1996 to 2018. Int J Infect Dis. 2019;1(84):48–53.
30. Tarnas MC, Desai AN, Lassmann B, Abbara A. Increase in vector-borne disease reporting affecting humans and animals in Syria and neighboring countries after the onset of conflict: A ProMED analysis 2003–2018. Int J Infect Dis. 2021;102:103–9.
31. Bonilla-Aldana DK, Holguin-Rivera Y, Cortes-Bonilla I, Cardona-Trujillo MC, García-Barco A, Bedoya-Arias HA, et al. Coronavirus infections reported by ProMED, February 2000–January 2020. Travel Med Infect Dis. 2020;1(35):101575.
32. GIDEON. GIDEON. 2024 [cited 2024 Jan 9]. Global Infectious Diseases and Epidemiology Network. Available from: https://www.gideononline.com/.
33. OCHA. ReliefWeb. 2024 [cited 2024 Jan 9]. ReliefWeb - Informing humanitarians worldwide. Available from: https://reliefweb.int/.
34. HealthMap. HealthMap. 2024 [cited 2024 Jan 9]. HealthMap. Available from: http://healthmap.org.
35. Meng Z, Okhmatovskaia A, Polleri M, Shen Y, Powell G, Fu Z, et al. BioCaster in 2021: automatic disease outbreaks detection from global news media. Bioinformatics. 2022;38(18):4446–8.
36. IFRC. IFRC Go. 2024 [cited 2024 Jan 9]. IFRC Disaster Response and Preparedness. Available from: https://go.ifrc.org/.
37. FAO. Empres-i Global Animal Disease Information System. 2024 [cited 2024 Jun 26]. Empres-i. Available from: https://empres-i.apps.fao.org/.
38. Delforge D, Wathelet V, Below R, Lanfredi Sofia C, Tonnelier M, van Loenhout J, et al. EM-DAT: the Emergency Events Database. Research Square. 2023 Dec 27 [cited 2024 Jan 27]. Available from: https://doi.org/10.21203/rs.3.rs-3807553/v1.
39. CRED. Public EM-DAT. 2024 [cited 2024 Jan 10]. Public EM-DAT platform. Available from: https://public.emdat.be/.
40. Lupo C, Wilmart O, Van Huffel X, Dal Pozzo F, Saegerman C. Stakeholders' perceptions, attitudes and practices towards risk prevention in the food chain. Food Control. 2016;1(66):158–65.
41. Paillé P, Mucchielli A. L'analyse qualitative en sciences humaines et sociales. 4e éd. Malakoff: Armand Colin; 2016. (Collection U).
42. Yung YF, Chan W. Statistical analyses using bootstrapping: Concepts and implementation. Stat Strateg Small Sample Res. 1999;1:81–105.
43. Porta M. Epidemic. In: A Dictionary of Epidemiology [Internet]. 6th ed. Oxford University Press; 2016 [cited 2024 Jan 2]. Available from: https://www.oxfordreference.com/display/https://doi.org/10.1093/acref/9780199976720.001.0001/acref-9780199976720-e-637.
44. Demongeot J, Magal P, Oshinubi K. Forecasting the changes between endemic and epidemic phases of a contagious disease, with the example of COVID-19. Mathematical Medicine and Biology: A Journal of the IMA. 2024:dqae012. https://academic.oup.com/imammb/advance-article-abstract/doi/10.1093/imammb/dqae012/7737412?redirectedFrom=fulltext.
45. Anna K, McLeod LT. Resilience: Lessons to be learned from safety and acceptable risk. J Saf Sci Resilience. 2021;2(4):253–7.
46. Peiter PC, Pereira R dos S, Moreira MCN, Nascimento M, Tavares M de FL, Franco V da C, et al. Zika epidemic and microcephaly in Brazil: Challenges for access to health care and promotion in three epidemic areas. PLOS ONE. 2020;15(7):e0235010.
47. Pardo-Araujo M, García-García D, Alonso D, Bartumeus F. Epidemic thresholds and human mobility. Sci Rep. 2023;13(1):11409.
48. Yasobant S, Bhavsar P, Lekha KS, Patil S, Falkenberg T, Bruchhausen W, et al. One Health Risk and Disease (OHRAD): a tool to prioritise the risks for epidemic-prone diseases from One Health perspective. Global Health Research and Policy. 2024;9(1):20.
49. Or Z, Gandré C, Durand Zaleski I, Steffen M. France's response to the Covid-19 pandemic: between a rock and a hard place. Health Econ Policy Law. 2022;17(1):14–26.
50. CRED. EM-DAT - The international disaster database. [cited 2024 Jan 10]. EM-DAT - The international disaster database. Available from: https://www.emdat.be/.
51. Guerra FM, Bolotin S, Lim G, Heffernan J, Deeks SL, Li Y, et al. The basic reproduction number (R0) of measles: a systematic review. Lancet Infect Dis. 2017;17(12):e420–8.
52. Kakulu RK, Kimaro EG, Mpolya EA. Effectiveness of Point of Entry Health Screening Measures among Travelers in the Detection and Containment of the International Spread of COVID-19: A Review of the Evidence. Int J Environ Res Public Health. 2024;21(4):410.
53. Garriga C, Valero-Gaspar T, Rodriguez-Blazquez C, Diaz A, Bezzegh P, Daňková Š, et al. Identification of methodological issues regarding direct impact indicators of COVID-19: a rapid scoping review on morbidity, severity and mortality. Eur J Public Health. 2024;34(Supplement_1):i3–10.
54. Spasic I, Nenadic G. Clinical Text Data in Machine Learning: Systematic Review. JMIR Med Inform. 2020;8(3):e17984.
55. Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. J Biomed Inform. 2017;1(73):14–29.

Publisher's Note