



Emerging Infectious Disease Repository

A Repository and Web Application Dedicated to the
Emergence of Infectious Disease

Zachary S. Gold
Carla Tilchin

Andrew G. Huff
Amy M. Slagle

Nathaniel Breit
Russell S. Horton



EcoHealth Alliance

February 2015

Image: Bubonic plague, *Yersinia pestis* (yellow) in the digestive system of a rat flea (purple). Courtesy of B. Joseph Hinnebusch, Elizabeth Fischer and Austin Athman, National Institute of Allergy and Infectious Diseases, National Institutes of Health

EXECUTIVE SUMMARY

EIDR (Emerging Infectious Disease Repository) combines a curated, expansive and transparent repository of information on past emerging infectious disease (EID) events occurring between 1940 and 2013 with a user-friendly, interactive web application built to disseminate and discuss this information. Development of EIDR is ongoing. As of February 2015, the EIDR repository contains 369 EID events. For each EID event, data were collected on a set of variables chosen by EcoHealth Alliance experts and designed to capture critical spatial, temporal, clinical, epidemiologic, economic, pathogen, host, and zoonotic information. This project builds upon previous work at EcoHealth Alliance, like the 'hotspots' map published by Jones et al., in 2008 in the journal *Nature*.¹

Users can interact with EIDR in a variety of ways. The web application enables users to compare multiple EID events simultaneously, visualize historical disease emergence spatially, explore individual emergence events in depth, and comment on emergence events. Each EID event has a unique event page that contains a narrative abstract, tables of data pertaining to the event, a map of the event's location, and an interactive discussion board.

EIDR is powerful information sharing platform dedicated to EID. The combination of EIDR's EID event repository with a complex web application makes EIDR a unique and valuable tool for the study of EID.

BACKGROUND

In 2014, the World Health Organization estimated that in 2012 infectious and parasitic diseases accounted for 15.8% of all disability-adjusted life-years (DALYs).² Although, the 2012 estimate is a 4% decrease from an estimate of 19.5% of all DALYs in 2000, it is important to maintain a broad perspective on the current disease threat.³ The rate of pathogen emergence is increasing with the number of known human viruses alone predicted to increase by ten to forty by 2020.⁴ Globalization

has produced an unprecedented vehicle for the spread of these emerging disease threats. The global spread of human immunodeficiency virus (HIV) was shocking. A mere 20 years after the syndrome's discovery, HIV was the fourth highest cause of death worldwide.⁵ More recently, the acute vulnerability of global health has been visible in the ongoing Ebola Virus Disease epidemic in West Africa. The impact of the epidemic has been catastrophic, causing well over 8,000 deaths in

one year, and incurring an estimated total financial impact of between \$3.8 and \$32.6 billion United States (U.S.) dollars.⁶ To combat these dangerous disease threats we must fully understand the driving factors behind their emergence so that informed and effective prevention, preparation, and response strategies can be developed.

Studying the origins of EID is a complex task. The term itself is not easily defined. Despite this obstacle, valuable efforts have undertaken. The investigation of emergence completed by Jones et al., in 2008,⁷ although based on a fairly broad definition of EID event, resulted in many critical findings. Jones et al. defined EIDs as:

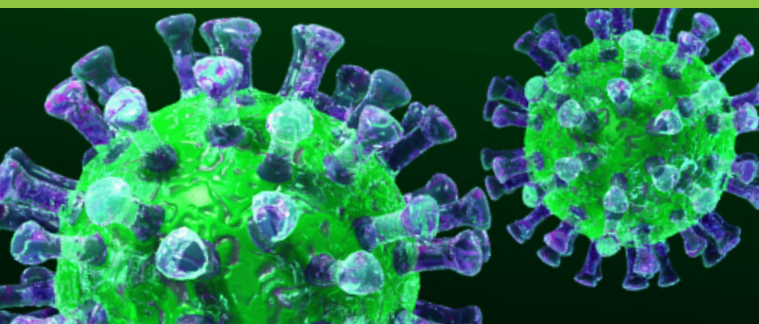
...the first temporal origination of an EID ... the original case or cluster of cases representing an infectious disease emerging in human

populations for the first time... caused by newly evolved strains of pathogens... pathogens that have recently entered human populations for the first time... and pathogens that have probably been present in humans historically, but which have recently increased in incidence.¹

Using that definition, Jones et al. identified 335 EID events from 1940 to 2004, finding a high prevalence of zoonotic emergence events (60.3%), a significant increase in the rate of emergence of novel drug resistance over time, an elevated risk of zoonotic, or vector-borne disease emergence in lower-latitudes, and several other important trends.⁸ Their work allowed for the creation of a map depicting 'hotspots' for disease emergence, a powerful tool for public health, and scientific research.

PROBLEM

Despite efforts to understand the origins of disease emergence, recent EID threats like Middle East Respiratory Syndrome (MERS); which emerged in Saudi Arabia in 2012, reveal the paucity of knowledge related to disease emergence. Available information surrounding historic emergence events is often complex, difficult to find, and even more difficult to validate. Efforts to improve and more effectively share knowledge on historical disease emergence are deeply needed.



SOLUTION

EIDR was created to serve as a centralized platform dedicated to the improvement, storage, display, dissemination, and discussion of current knowledge on historical disease emergence. The EIDR repository improves and expands knowledge of historical emergence events, and identifies unknowns and uncertainties in that knowledge. This information is displayed transparently in a centralized web application that allows for efficient exploration of disease emergence. Although the information in the EIDR repository is extensive, disease emergence is exceedingly complex. To address this complexity discussion boards have been built into EIDR to aggregate additional knowledge, and capture scientific debate.

Image: MERS (Middle East respiratory syndrome), Masoud Zamani

INTENDED AUDIENCE

EIDR is intended to serve public health officials, researchers, and the public. EIDR will provide substantial, transparent, and accessible data on EID events for public health officials and scientists. This data will improve disease surveillance efforts, and enhance the power of disease modeling. It will also aid in efforts to identify existing knowledge gaps in disease emergence, and allow for further investigation of trends in EIDs. Additionally, the discussion boards will allow for debate over the causes and epidemiologic characteristics of specific emergence events. Such debates may result in further research, subsequent improvements to our understanding of disease emergence, and investment in EID research.

The user-friendly web application built for EIDR provides an informative, accessible and engaging platform for exploring emergence. It is designed to appeal to a broad audience, including the general public. By educating the public on the threats associated with EID the web application will act as a public health tool.

COMPONENTS AND FEATURES

Repository of knowledge on historical EID events

The EIDR repository contains information on infectious disease emergence events occurring between 1940 and 2013. Events are considered emerging if they satisfy any of the following characteristics: expanded geographic region or host range, increased incidence or virulence, novel drug resistance, evolution of an existing organism, or reappearance after control or elimination. Potential emergence events that could not be verified, because of historical complexities, or a lack of reliable information, are also included. An adjoining request for more information is included with these events.

For each EID event data were collected on a set of variables determined by a team of EcoHealth Alliance experts. These variables are designed

to capture critical spatial, temporal, clinical, epidemiologic, economic, pathogen, host, and zoonotic information. Additionally, data were collected on potential drivers associated with a particular emergence event, like war and famine, antimicrobial agent use, or proximity to wildlife. If no information could be found on a particular variable this absence was captured. Short abstracts providing a summary of the emergence event are included for most events. The information collected on each event was obtained through intensive investigation of the literature. Taxonomic, economic, general pathologic, and general host information was acquired from varying sources, many of which are not directly related to the emergence events.

Web application

The information stored in the EIDR repository

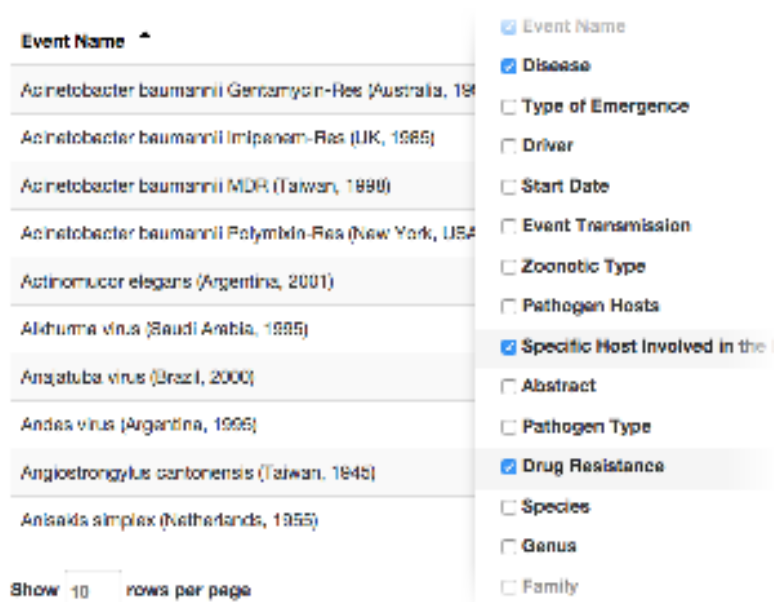


Figure 1. The customizable “Emergence Events” view allows users to sort, explore, and compare events by selecting EIDR variables for the table to display. In this particular “Emergence Event” view, data collected on disease, driver, type of emergence (hidden by dropdown menu), and city are displayed in the customizable table. Clicking any of these events brings the user to the page for that event.

is displayed in a user-friendly, interactive web-application. Emergence events can be sorted, compared, and investigated in a variety of ways. The “Emergence Events” view contains an interactive table of EID events. This table is customizable, allowing the user to choose which EIDR variables they would like to view (Fig. 1). Users can perform specific searches within the “Emergence Events” view using a filter feature. This feature allows users to search for events with a common variable, like a specific host, or pathogen. A broad spatial depiction of all EID events within the EIDR repository is offered through the “Event Map” view (Fig. 2).

Users can explore individual EID events in greater detail through individual event pages. Event pages can be accessed from the “Emergence Events” list by clicking on an event

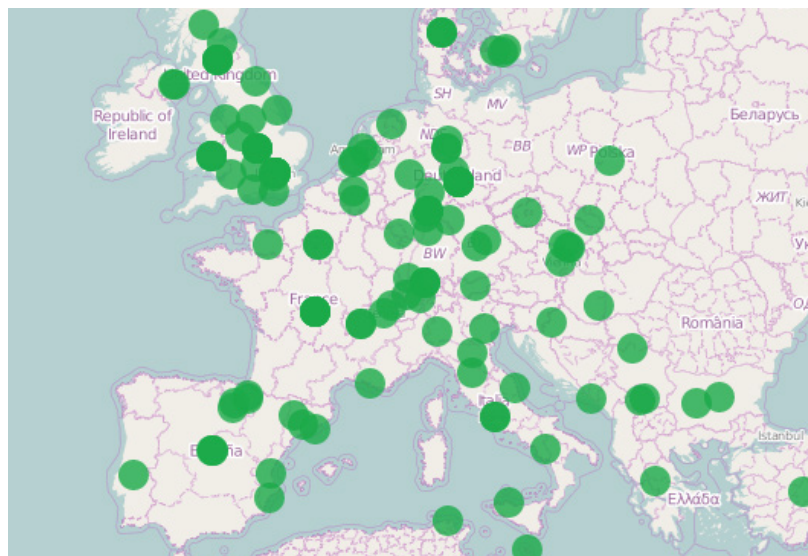


Figure 2. “Event Map” view depicting the location of all EID events that occurred in a given region. Individual events can be accessed by clicking on their corresponding pins.

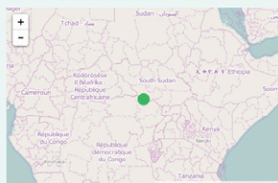
in the table. Each EID event page contains a detailed report on the event, including a map showing the location of the event, a narrative abstract, tables of additional data, and a discussion board that allows users to comment on the event (Fig. 3). In some cases data is displayed with supporting textual evidence. References used for each event are available in the event pages.

Value

EIDR is the combination of an expansive, reliable and transparent repository of data on disease emergence with a user-friendly, engaging and interactive web application. The data collected in EIDR could improve our understanding of disease emergence, by allowing for the creation of more accurate disease emergence models, and more insightful analyses of the trends in EIDs. An initial application of EIDR could be the creation of an improved ‘hotspots’ map of disease emergence. Additionally, EIDR data could be used to study trends in disease

Ebola virusDisease: Ebola Hemorrhagic Fever
1976

Between June and November 1976 a deadly outbreak of Ebola hemorrhagic fever infected 284 people living in southern Sudan. A total of 151 people died. Patients presented with fever, headache, joint and muscle pain, diarrhea, vomiting, chest-pain, pain and dryness of the throat, rash and hemorrhagic manifestations. The index case was a storekeeper in a cotton factory located in the small town of Nzara. The patient became ill on June 27, 1976, and died on July 6. The virus was transmitted within the factory and to family members through close contact with acutely ill patients. The outbreak spread to Maridi after a patient from Nzara was admitted to a hospital there. A large outbreak subsequently occurred in the hospital due to nosocomial transmission. Fruit bats are the natural hosts of the virus, although this was unknown at the time of the event. Proximity to wildlife likely drove the emergence of this virus into the human population.



Location: Nzara
*This represents the most specific location information found for the event.

Emergence Type	Driver	Infected	Deaths
expanding host range, evolution of an existing organism	ecosystem change, proximity to wildlife	284	151
Event Transmission	Zoonotic Type	Host	
	zoonotic from wildlife	fruit bats, primates, other mammals	

Stats		
Type of Emergence	expanding host range, evolution of an existing organism	
Driver	ecosystem change, proximity to wildlife	
Start Date	1976-06-27	
Start Date Description	first report	
End Date	1976-11-25	
End Date Description	last report	
Duration of Event	5 months	
Number Infected	284	
Number of Deaths	151	
Event Transmission	close, nosocomial	
General Transmission	close, nosocomial	
Reported Symptoms	Fever (100%), Headache (100%), joint and muscle pain (100%), Chest pain (83%), Diarrhea (81%), Vomiting (89%), Dry painful throat (83%), Rash or desquamation (52%), Cough (48%), hemorrhagic manifestations (71%)	
Sample Type	serology	

Pathogen		
Pathogen Type	virus	
Drug Resistance	0	
Initial Reported Name	Ebola virus, subtype E. Sudan	
Species	Sudan ebolavirus	
Genus	Ebolavirus	
Family	Filoviridae	
Order	Mononegavirales	
Economics		
Occupation	cotton factory workers, healthcare workers	
Per Capita National GDP in Year of Event	360 (current US\$)	
Life Expectancy in Country During Event Year	53.8 years	

Figure 3. The event page includes a map with the location of the EID event (left and right), an abstract describing the event (left), and tables displaying different sets of data for a particular event (right). By hovering over variables with their mouse, users can obtain a short definition of each variable (right).

emergence that were hitherto prohibitively difficult due to a lack of quality data. Such studies could better analyze categorical trends in disease emergence based on host taxonomy, or the impact (quantified by deaths) of initial emergence events over time.

The web application associated with the repository will help bring the data to the attention of interested parties and ensure that this information is not abandoned. Too often useful datasets are developed for a single task and promptly forgotten. Debates sparked by the information displayed on the web application will provoke interaction with the data, hopefully leading to its improvement and use in future studies and models. The EIDR initiative is the marriage of a robust repository of scientific data, with an effective, and powerful web-based application.

¹Jones KE, Patel NG, Levy MA, Storygard A, Balk D, Gittleman JL, Daszak P (2008) Global Trends in Emerging Infectious Diseases. *Nature*. Vol. 451:990-993.

²World Health Organization (2014) Global Health Estimates 2014 Summary Tables: DALY by cause, age and sex, by WHO Region, 2000-2012. Geneva: WHO. http://www.who.int/healthinfo/global_burden_disease/en/

³Ibid.

⁴Woolhouse EJ, Howey R, Gaunt R, Gaunt E, Reilley L, Chase-Topping M, Savill N (2008) Temporal Trends in the Discovery of Human Viruses. *Proc. R. Soc. B*. Vol. 275:2111-2115.

⁵Smolinski MS, Hamburg MA, Lederberg J (2003) Microbial Threats to Health: Emergence, Detection, and Response. Institute of Medicine (US) Committee on Emerging Microbial Threats to Health in the 21st Century. Washington (DC): National Academies Press (US).

⁶World Bank Group (2014) Update on the Economic Impact of the 2014 Ebola Epidemic on Liberia, Sierra Leone, and Guinea. Washington, DC: World Bank.

⁷Jone et al., (2008) Global Trends in Emerging

⁸Ibid.



EcoHealth Alliance



View the application at eidr.ecohealth.io