

Valparaiso's 2014 Fire: Evaluation of Environmental and Epidemiological Risk Factors During the Emergency Through a Crowdsourcing Tool

Sebastián Eduardo Espinoza Espinoza, MSc; Anibal Enrique Vivaceta De la Fuente, MPH; Constanza Andrea Machuca Contreras, DDS

ABSTRACT

Objective: To describe and relate the main environmental risk factors in the emergency process after a large urban fire in Valparaiso, Chile, in April 2014.

Methods: An observational, cross-sectional descriptive study was performed. All 243 reports from an ad hoc web/mobile website created on the Ushahidi/Crowdmap platform were reviewed. Reports were recorded in a new database with dichotomist variables based on either the presence or absence of the relevant category in each report.

Results: Almost one-third of the reports presented data about garbage (30%) and chemical toilets (29%). Reports related to water, infrastructural damage, and garbage had significant associations with 4 categories by chi-square test. In the logistic regression model for chemical toilets, only the variable of water was significant (P value = 0.00; model P value: 0.00; R^2 : 11.7%). The “garbage” category confirmed infrastructural damage (P value: 0.00), water (P value: 0.028), and vectors (P value: 0.00) as predictors (model P value: 0.00; R^2 : 23.09%).

Conclusions: Statistically significant evidence was found for the statistical dependence of 7 out of 10 studied variables. The most frequent environmental risk factors in the reports were garbage, chemical toilets, and donation centers. The highest correlation found was for damaged infrastructure, vectors, and garbage. (*Disaster Med Public Health Preparedness*. 2016;page 1 of 5)

Key Words: urban fire, environmental risk factors, epidemiology, emergencies and disasters, crowdsourcing

On April 12 and 13, 2014, Valparaiso suffered the worst urban fire in its history. On this occasion, a usual situation in Valparaiso, the spread of a wildfire to poor peripheral urban areas, went further than usual. More than 3000 houses were burned in 2 days. The fire also destroyed the family health care facility that served a high proportion of the affected population. Two of the authors were part of a team that created a specific website using the Ushahidi/Crowdmap (Nairobi, Kenya) platform to collect data in real time. This allowed data capture from the field from April 14 through June 26. Environmental epidemiological risk factors, described previously by other authors,¹ were explored and displayed.

It is well known that disasters are related to a broad spectrum of health problems, such as diarrhea, fever, and respiratory diseases in children.² Environmental risk factors like access to potable water, wastewater disposal, solid waste management, food security, vector exposure, and weather conditions play a key role in the risk of developing those diseases in the affected population.³

Health conditions during and after a disaster are a key part of the disaster's impact assessment and must be studied carefully because their effects are not immediately evident in the early stages of a disaster.⁴ Data on the health of a population and the interpretation of these data are fundamental for decision-making about the whole response to a disaster, including preventive actions for medium- and long-term effects. Epidemiological data and information management, bias control, and sanitary intelligence are basic for public health decisions, their implementation, and communication.⁵

Since the health care sector is responsible for just a small share of all the factors influencing the health of some populations, data must be useful as feedback for decision-making at all territorial scales and for different disciplines and management areas. Also, the data and information generated for health monitoring support further knowledge production through more detailed analysis and interpretation, which is of vital importance.

Disaster learning requires that researchers utilize these scarce, ominous real-life situations to test new

conceptual, technological, and operational developments. Too often, the valuable data set collected is not used to generate learning at a later stage. This is an insurmountable waste in such scarce opportunities, because the previous experience is one of the most important sources of preventive information, allowing modeling of future scenarios for better preparedness.⁶

Some similar experiences of data collection during uncontrolled wildfires affecting houses do exist, but the processing of these data was usually qualitative.⁷ These examples show the need for the mapping process to be collective instead of being a centralized effort restricted to first responders or official managers. Experiences describing the use of information technologies in collective mapping of disasters clearly show that Internet users are faster than seismological agencies in providing updated data.⁵ Vector-transmitted diseases can account for an important part of the consequences of a disastrous event.⁸ Furthermore, although most authors agree that a sizable portion of the effects of a disaster emerge in the medium and long term, most epidemiological studies only consider short-term effects.

The system disruption inherent to disasters demands tools able to deal with high levels of uncertainty for both their management and their description. It is worthless to expect a detailed description of highly controlled variables. Uncertainty will be a part of data management, from capture to interpretation. This problem surely exists and complicates the performance of managers, because the subject itself puts uncertainty in the center. Any model intended to describe such chaotic situations must bring along uncertainty.

MATERIALS AND METHODS

An observational, descriptive, cross-sectional study was performed. Because of the diversity of the data generators and the potential users, the monitoring system was based on crowdsourcing/crowdfunding. All 243 reports were collected in a census-sampling mode.¹

During the emergency, different stakeholders, such as medical doctors, nurses, physical therapists, and students of the same discipline, collected data. The data were uploaded in real time through a website/map generated in Crowdmap (<http://incendiovalpo.crowdmap.com>; Ushahidi, Nairobi, Kenya). This free map-based tool allows users to create a thematic mapping project, with uploading of data in a distributed way, and is managed on a cloud-based server. Administrators can define categories, organize categories, and manage users' privileges and displays. Users can create their own reports as geographic objects (point, line, polygon) and associate attributes and media.

The authors began creating the website on April 14, 2 days after the start of the emergency. The website was started

simultaneously with intentional data collection in the field. Ushahidi's plasticity allowed us to enter the initial data "on the run," while the website was still in a basic configuration. We initially used categories defined by the Pan American Health Organization and Chile's Ministry of Health.⁶ The categories were later reorganized according to feedback in the field and the characteristics of the collected data.

Ten categories were finally defined, in order to better collect the different types of reports. These were as follows: (1) chemical toilets, (2) garbage, (3) feces, (4) damaged infrastructure, (5) health care facilities, (6) water, (7) landslides, (8) donation center, (9) vectors, and (10) others. Customization properties of Ushahidi were used to set 2 different levels of access privileges. Those reporting collective situations and areas were left open access. Personal disease surveillance data are covered by Personal Data Protection Law;⁹ thus, these data were set for restricted access, just by the Sanitary Authority and system administrators. Surveillance data were also transmitted directly to the Sanitary Authority by using a protocol agreed upon on the third day of operation, safeguarding all the ethical research criteria.

For the present research, a new database was created to ease descriptive analysis. Categories in each report and the content described by the user were recorded in 10 dichotomous categories based on the presence or absence of the relevant category in the report. Analysis did not consider the actual meaning of the variable in terms of its positive or negative impact on the population but only whether it was included in each report. Because reports could have data from more than one category, all of them were included. Data were processed by using STATA 13.0 (StataCorp, College Station, TX), characterizing the proportions of the different factors.

RESULTS

A total of 243 reports were analyzed, with an average presence of 1.72 (± 1.44) categories. Almost one-third of the reports presented data about garbage (30%) and chemical toilets (29%). One-fourth reported situations linked to donation centers (25%), water (24%), and infrastructural damage (24%). Noticeably, 12% of the reports included data about vectors, the same percentage as those labeled "others." Landslides and feces were included in few reports (7% each). Just 2% of the reports included health facilities (Table 1). If only absolute frequencies of positive observations, in each category, were selected ($n = 491$), three-fourths (76%) of the observations corresponded to 5 categories: chemical toilets (17%), garbage (17%), donation centers (14%), water (14%), and infrastructural damage (14%) (Figure 1).

One to one, contrasts were also performed, checking for dependence of study variables. Reports related to water, infrastructural damage, and garbage had significant associations with 4 categories. These were followed by reports about

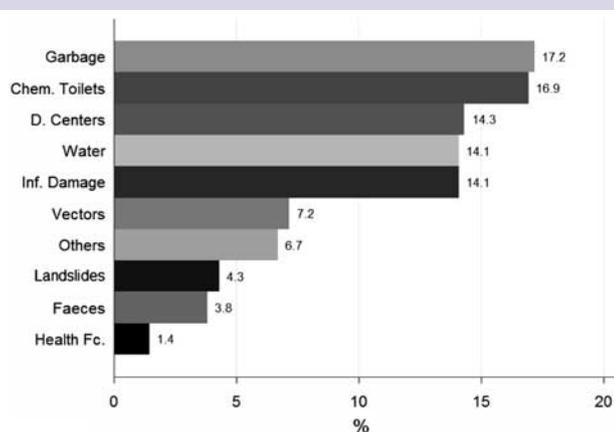
TABLE 1

Presence of Categories in the Reports ^a					
	Fr	%	SE	95% CI	
Chemical toilets					
No	172	71%	0.03	65%	77%
Yes	71	29%	0.03	23%	35%
Garbage					
No	171	70%	0.03	65%	76%
Yes	72	30%	0.03	24%	35%
Feces					
No	227	93%	0.02	90%	97%
Yes	16	7%	0.02	3%	10%
Infrastructural damage					
No	184	76%	0.03	70%	81%
Yes	59	24%	0.03	19%	30%
Health facilities					
No	237	98%	0.01	96%	99%
Yes	6	2%	0.01	1%	4%
Water					
No	184	76%	0.03	70%	81%
Yes	59	24%	0.03	19%	30%
Landslides					
No	225	93%	0.02	89%	96%
Yes	18	7%	0.02	4%	11%
Donation centers					
No	183	75%	0.03	70%	81%
Yes	60	25%	0.03	19%	30%
Vectors					
No	213	88%	0.02	83%	92%
Yes	30	12%	0.02	8%	17%
Other					
No	215	88%	0.02	84%	93%
Yes	28	12%	0.02	7%	16%

^aAbbreviations: CI, confidence interval; Fr, frequency; SE, standard error.

FIGURE 1

Frequency Plot of the Impact of Each Category on the Total Reports.



Abbreviations: Chem. Toilets, chemical toilets; D. Centers, donation centers; Health Fc., health facilities; Inf. Damage, infrastructural damage.

TABLE 2

Significant Contrasts Performed on the Reports Showing Dependency of Study Variables			P Value
Chemical Toilets	Garbage		0.002 ^a
	Infrastructural damage		0.026 ^a
	Water		0.000 ^a
Garbage	Infrastructural damage		0.000 ^a
	Water		0.000 ^a
	Vectors		0.000 ^a
Feces	Inf. damage		0.000 ^a
	Water		0.007 ^a
Infrastructural damage	Donation centers		0.000 ^a
Water	Donation centers		0.005 ^b
Landslides	Others		0.002 ^a
Vectors	Others		0.002 ^a

^aP value by chi-square test.

^bP value by Fisher's exact test.

chemical toilets, which were associated with 3 categories. Donation centers and vectors were associated with 2 categories. Health facilities showed no associations (Table 2).

The study of these associations was expanded by using a logistic regression model. This showed that, for instance, for the dependent variable "chemical toilets," just the variable "water" showed as explicative in the model (P value = 0.00) (model P value: 0.00; R^2 : 0.117). The dependent variable "garbage" confirmed predictor variables "infrastructural damage" (P value: 0.00), "water" (P value: 0.028), and "vectors" (0.00) (model P value: 0.00; R^2 : 0.2309). The other variables were not explicative.

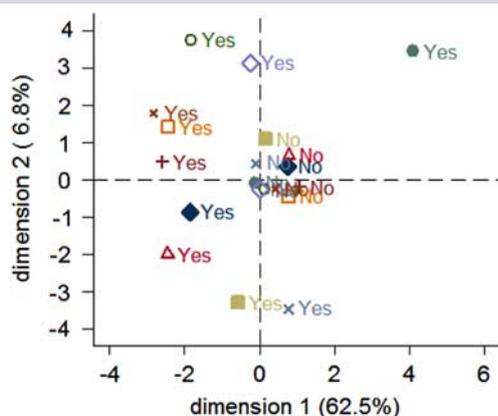
Finally, a multiple correspondence analysis showed that the variables "feces" and "landslides" could be more closely related to each other. Something similar happened with "vectors," "infrastructural damage," and "garbage." There was also an association in the groups "chemical toilets" with "water," and "donation centers" with "others." The health facilities variable did not seem to be related with any other category (Figure 2).

DISCUSSION

Implementing a sampling plan in an emergency or disaster situation is a complex task. On the one hand, resources are distributed to the event itself, and on the other, the unique conditions of each disaster often make it impossible to have a predefined research protocol before the event and its consequences. Because of the practical impossibility of having complete sampling frames, or clusters after the situation, it is not possible to infer through probabilistic sampling. Other authors have described this constraint.^{1,5,6} Collaborative sampling brings promising solutions but challenges remain regarding bias control. There are important restrictions to standardization when collectors with very different

FIGURE 2

Multiple Correspondence Analysis Coordinate Plot of Category Reports.



Abbreviations: D. Centers, donation centers; Health Fc., health facilities; Inf. Damage, infrastructural damage.

backgrounds are going to the field for other purposes and secondarily collect data. Written detailed instructions can help to reduce bias but have a restricted effect.

Census-based sampling, like that explored in this research, could solve such difficulties. Nevertheless, such sampling can entail problems with capturing an emergency's variability or its evolution. Also, volunteers may feel uncomfortable and even frustrated when they carry out data collection instead of physical, immediate-impact work.¹

The literature shows an imperative need for tools for rapid data collection in emergency and disaster situations.² Recent experiences show the potential utility of distributed knowledge facing that need. This research confirms previous findings from other sources in the literature.^{4,7}

Concomitancy of variables could be used as a predictor orienting future preparation. Predefined and preset ways of assessing this possible innovation should be designed in advance as part of a preparedness process. Probabilistic models could be implemented for specific scenarios. These research designs must take account of modeling uncertainty and the diversity of such scenarios.

The necessary flexibility and improvisation of the whole process of data management—from capture to decision-making and learning—requires, reciprocally, a careful and

detailed preparation in this specific area and cannot be improvised. A well-trained team should be able to start data collection in a predefined, flexible way.¹⁰

Evidence of the linkage of some categories, especially chemical toilets, water, damaged infrastructure, and vectors, makes necessary further exploration of incidences and prevalence of diseases with a geostatistical perspective. However, most of the patient's personal data files were destroyed together with the Primary Health Centre. Current data do not provide a reliable baseline.

The use of internationally accepted categories eases the reproducibility of studies. However, it is not certain that during the development of the emergency such categories will be *stabilized*.

Health facilities were not related with the presence of any variable. This could be explained by the fact that sanitary facilities are covered early in an emergency, reducing the probability of finding sanitary risk conditions.

Emergency managers have to monitor the whole landscape of the emergency, and thus need data feeding in from the field. Official field staff usually cannot cover the whole territory, and the bureaucratic pattern of data management reduces management capabilities during emergencies. Distributed-knowledge-based technologies can be effective tools for increasing updated data streaming. Categorization and filtering, like those shown in this article, can be automatized for better results.

Limitations

The main limitations of this study were due to variable recategorization and not distinguishing the negative from the positive connotations of each report. This could bias the presence of the variables considering that in emergencies, attention is focused on deficits much more than on well-functioning aspects.

CONCLUSIONS

During the uncontrolled wildfire of April 2014 in Valparaiso, the most frequent environmental risk factors in the reports uploaded to a Crowdfunder-based platform were garbage, chemical toilets, and donation centers. Multiple correspondence analysis enlightens the emergence of environmental risk factors after an uncontrolled wildfire. The highest correlation found was for damaged infrastructure, vectors, and garbage. Maps are a fundamental tool for disaster management, allowing a deeper and broader understanding of a disaster's development and territorial impact. A pyramidal style of data flux and data management paradigm must be replaced by one based in the acknowledgment of the distributed condition of knowledge/data. This is the greatest challenge for disaster and emergency information systems.

About the Authors

Schools of Dentistry (Drs Espinoza and Machuca) and Medicine (Dr Vivaceta), Public Health Department, Universidad de Valparaíso, Valparaíso, Chile.

Correspondence and reprint requests to Sebastián Espinoza E, #211, Subida Carvallo, Valparaíso, Chile (e-mail: sebastian.espinoza@uv.cl).

REFERENCES

1. Astorga M, Eguren D, Espinoza S, et al. An experience in combining disaster data with tools of distributed knowledge. *Revista Chilena de Salud Pública*. 2014;18(2):199-202. <http://dx.doi.org/10.5354/0719-5281.2014.31980>.
2. Datar A, Liu J, Linnemayr S, et al. The impact of natural disasters on child health and investments in rural India. *Soc Sci Med*. 2013; 76(1):83-91. doi: 10.1016/j.socscimed.2012.10.008.
3. Pan American Health Association, Ministry of Health Office. *Guía de Vigilancia Epidemiológica en Emergencias y Desastres*. Santiago de Chile: Elaborado por el Ministerio de Salud; 2010; ISBN: 978-956-8823-52-8.
4. Noji, Eric K. *Impacto de los Desastres en la Salud Pública*. Bogotá, Colombia: Pan American Health Association/WHO; 2000.
5. Zook M, Graham M, Shelton T, et al. Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. *World Med Health Policy*. 2010;2(2):7-33. <http://dx.doi.org/10.2202/1948-4682.1069>.
6. Van den Berg B, Grievink L, Gutschmidt K, et al. The public health dimension of disasters - health outcome assessment of disasters. *Prehosp Disaster Med*. 2008;23(suppl):55S-59S.
7. Petersen K. Producing space, tracing authority: mapping the 2007 San Diego wildfires. *Sociol Rev*. 2014;62:91-113. <http://dx.doi.org/10.1111/1467-954X.12125>.
8. Korteweg HA, van Bokhoven I, Yzermans CJ, et al. Rapid health and needs assessments after disasters: a systematic review. *BMC Public Health*. 2010;10(1):295. <http://dx.doi.org/10.1186/1471-2458-10-295>.
9. Ministry of Health Office, Government of Chile. Public law number 20.584.
10. Dominici F, Levy JI, Louis TA. Methodological challenges and contributions in disaster epidemiology. *Epidemiol Rev*. 2005;27(1):9-12. <http://dx.doi.org/10.1093/epirev/mxi009>.