Using Internet reports for early estimates of the final death toll of earthquake-generated tsunami: the March 11, 2011, Tohoku, Japan, earthquake

Xiaolin Yang[1,2], Zhongliang Wu[1,*], Yingchun Li[3]

[1] Institute of Geophysics, China Earthquake Administration, Beijing, China
[2] Earthquake Administration of Shaanxi Province, Xi'an, China
[3] Jiangsu Province Earthquake Administration, Nanjing, China

ABSTRACT

A retrospective case study was conducted for the March 11, 2011, earthquake that occurred off the east Pacific coast of Tohoku, Japan, with reference to the Sumatra-Andaman earthquake and Indian Ocean tsunami on December 26, 2004. The regularities of the temporal variations of the number of deaths reported through the Internet is used to make an early stage quick-and-rough estimate to assist in the emergency decision making, and to continue the revision of the estimate as new data comes in. The quick-and-rough estimate of the final fatalities that was obtained on the second day after the earthquake is shown to agree with the final reported fatalities to within an order of magnitude. This also has an added advantage over other estimation models, in that the estimates can be updated with new data as they become available.

Introduction

Estimations of the final fatalities caused by earthquakes and/or earthquake-generated tsunamis have critical roles in the emergency decision-making, and they are a topic that has been under continuous study [Lomnitz 1970, Ohta et al. 1983, Samardjieva and Oike 1992, Tsai et al. 2001, Samardjieva and Badal 2002, Ashkenazi et al. 2005, Badal et al. 2005, Gao and Jia 2005, Gutiérrez et al. 2005, Liu and Wu 2005, Liu et al. 2005, Wyss 2005, Koshimura et al. 2006, Lamontagne 2008, Zhao et al. 2008, Wu et al. 2009, Li et al. 2011, Spence et al. 2011]. Previous studies on the regularities of death tolls have mainly focused on their relationships with the intensities or magnitudes of earthquakes [e.g., Samardjieva and Oike 1992, Samardjieva and Badal 2002, Wyss 2005] or with the destroyed buildings [Ohta et al. 1983], which can be used as a proxy for early estimates. For this reason, several public, governmental and commercial agencies, such as the National Earthquake Information Center (USA), the China Earthquake Administration

(China), and the World Agency of Planetary Monitoring and Earthquake Risk Reduction (Switzerland), provide such (open or internal) information services shortly after earthquakes. On the other hand, however, many factors contribute to the uncertainty of these 'first-glance, database-based estimates', such as site condition, heterogeneity of intensity distribution, status of infrastructure and economy, and emergency response capability of the local community. Therefore, a mechanism for on-going revisions of estimates using feedback from field reports is apparently needed.

Based on our previous data [Liu and Wu 2005, Liu et al. 2005, Zhao et al. 2008, Li et al. 2011], for the present study we conducted a retrospective case study of the March 11, 2011, great earthquake off the east Pacific coast of Tohoku, Japan, that was followed by the earthquake-generated tsunami (http://www.eqh.dpri.kyoto-u.ac.jp/src/etc/tohoku/; http://www.iris.edu/news/events/japan2011/). These data are considered with reference to the December 26, 2004, Sumatra-Andaman earthquake and the Indian Ocean tsunami that followed. We investigated the use of the temporal variations in the number of deaths reported by a single source on the Internet (http://www.sina.com.cn), to determine the early stage quick-and-rough estimates (QREs) to assist in emergency decision-making, and to continue to revise the estimates as new data come in.

Study methods

The temporal variations in fatalities reported after earthquakes and/or earthquake-generated tsunamis have been studied for quite some time. In 2001, Tsai et al. [2001] analyzed the 1999 Chi-Chi earthquake for the temporal cumulative numbers of killed and missing people found by the search and rescue teams. Our present approach

originated, to large extent, from that of Gao and Jia [2005], who modeled the temporal changes of deaths found by the earthquake rescue operations. They used a complicated polynomial function with five degrees of freedom. We modified the data-fitting function of Gao and Jia [2005], and based on the cases of the 2004 Indian Ocean tsunami and the 2005 Pakistan earthquake, we proposed a simple model with two degrees of freedom [Liu and Wu 2005, Liu et al. 2005]. As a further trial, Zhao et al. [2008] collected more cases, and showed that the rescue-efficiency parameter of the model of Liu and Wu [2005] is magnitude dependent. Li et al. [2011] further developed the model with more earthquake cases, and proposed the QRE approach. In this model, the time-dependent number of deaths reported at time t, $N(t)$, can be represented by [Liu and Wu 2005]:

$$N(t) = N_0\left[1 - \exp(-\alpha t)\right], \qquad (1)$$

in which N_0 and α are constants. Equation (1) was shown to be a good approximation of the real cases for 12 earthquakes and/or tsunamis [Zhao et al. 2008]. The meaning of Equation (1) can be shown by taking the time derivative of both sides:

$$\frac{dN}{dt} = \alpha(N_0 - N), \qquad (2)$$

in which the term dN/dt on the left side is the rate of discovery of dead bodies, the term $(N_0 - N)$ on the right side is the dead bodies that remain undiscovered, and the constant N_0 is the final death toll. The proportional relationship between these two quantities, dN/dt and $(N_0 - N)$, is a natural outcome of the rescue operation. Qualitatively, the more bodies of the dead that remain to be discovered, the higher the probability that the rescue team will discover them. The coefficient α that links these two terms reflects the efficiency of the rescue practice. Zhao et al. [2008] and Li et al. [2011] showed that the order of magnitude (i.e. the logarithm) of α decreases linearly with the order of magnitude of the radiated energy of the earthquake (as represented by its surface-wave magnitude M_S), albeit with significant uncertainties.

Whether these suggested parameters can be used for early estimates of final death tolls is a question with practical importance. In the deployment of the task forces for rescue and relief, estimates of the final death toll is one of the critical factors for decision making. A realistic estimate has to be available as early as possible, to take full advantage of the 'golden time' for rescue and relief. Meanwhile, for the follow-up relief and aid practice, this information that is up-dated with time is also of practical importance. To this end, a QRE was proposed by Li et al. [2011], as represented by:

$$N_0^{\mathrm{QRE}} = \frac{N(t)}{\left[1 - e^{-\alpha t}\right]}, \qquad (3)$$

in which α is taken from the empirical $\log\alpha$-M_S relation [Zhao et al. 2008, Li et al. 2011]. A case study by Li et al. [2011] using earthquakes and an earthquake-generated tsunami showed that for earthquakes, the QRE on the first day tends to be an underestimation, while that on the second day is much closer to the actual value. For tsunamis, it takes more time for the QRE to match the acual situation. Using Equation (3), the QREs cannot be improved by the accumulation of data; i.e. the QREs at different times are independent of each other. This is apparently a shortcoming, although considering that the early data are not sufficient for stable exponential-function-like curve fitting, this independence characteristic turns out to be an advantage when compared to other approaches, such as, for example, curve fitting.

The reason for the underestimation of the early estimate is also clear: Equation (1) only reflects the global property of the temporal variation over a long period. Examining the initiation of the temporal variation shortly after an earthquake [Li et al. 2011], it can be seen that a better description should be:

$$\frac{dN}{dt} = \alpha'N(N_0 - N), \qquad (4)$$

which defines a curve similar to logistic curves in ecology. Therefore, neither the curve-fitting of the initiation part of the time series using Equation (1) nor the QRE using Equation (3) are good estimators. For earthquakes it takes two days to overcome this underestimation problem, while for tsunamis it takes longer. Using a small number of data samples, Equation (4) is not easy to fit, due to the nonlinearity. To overcome this difficulty, a '2nd-day-QRE times 2 to 3 rule' was derived empirically from past earthquake-tsunami cases [Li et al. 2011]; although this is rough, it can be used as a simple but practical solution. Note that at the early stages of emergency responses, even if the rough estimate of the final death toll is only within an order of magnitude, this provides important information for decision making, and thus an assessment of this estimate is worthwhile.

Data used for analysis and results

We used the Internet report from a single media source (http://www.sina.com.cn) that was collected continuously after the March 11, 2011, Tohoku earthquake. The news related to the death toll was collected and organized, to examine how the number of reported deaths varied with time. As a convention, for the reports in the form 'the death toll is at least X', we simply took X as the value of the data at that time. At present, although the data from news media have (the most critical) timing characteristics, they have various shortcomings relating to their accuracy and reliability. For example, in the 2004 Sumatra-Andaman earthquake, the data published by different governmental
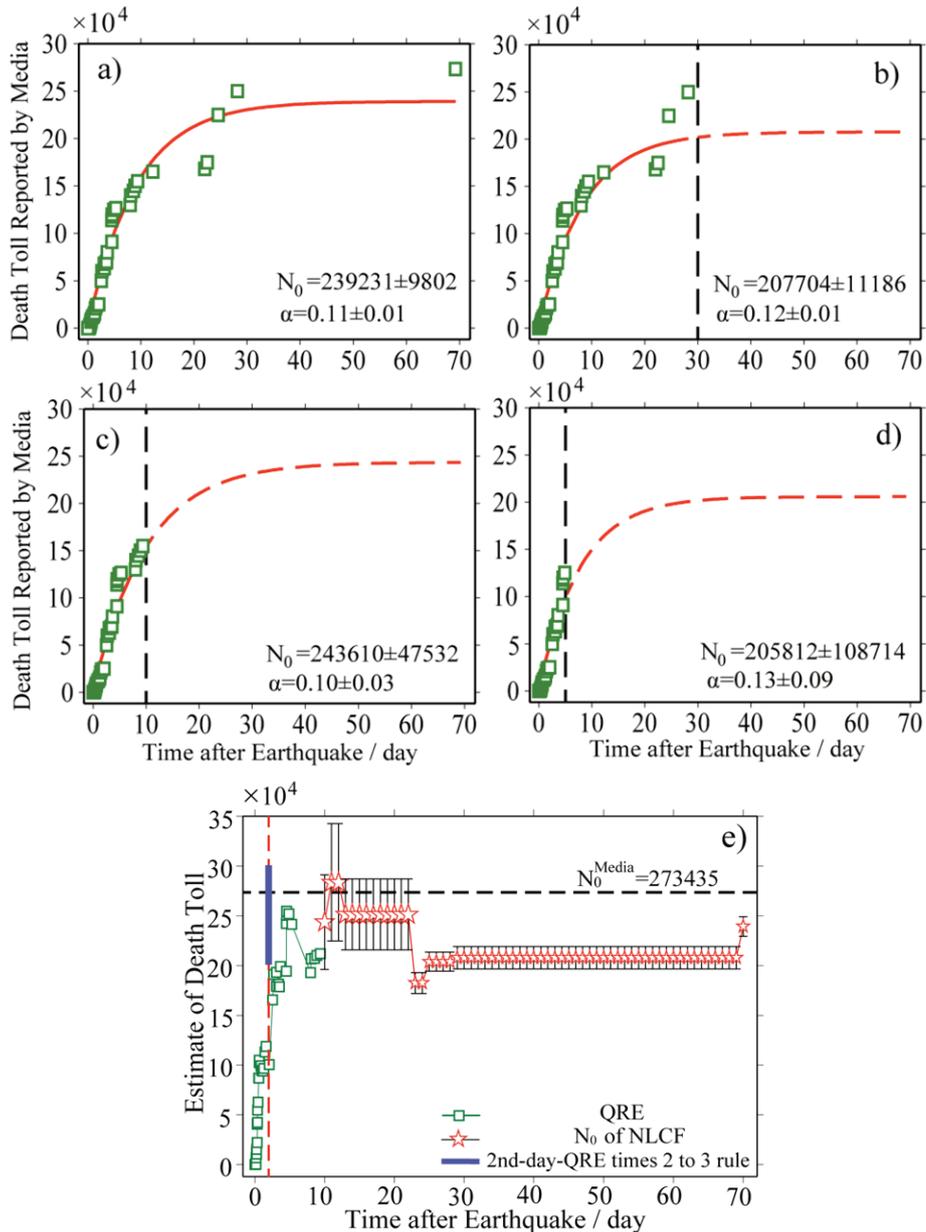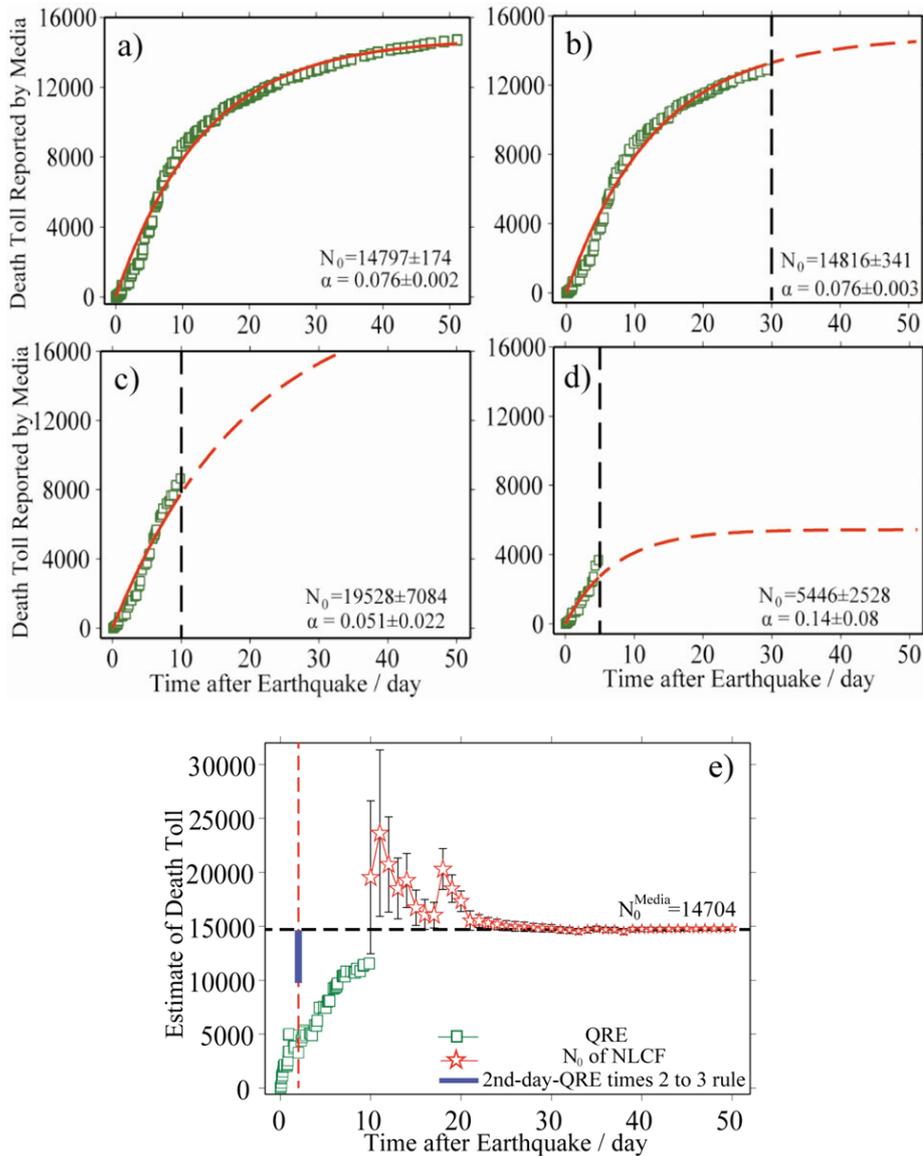
Figure 1. Variations in the reported death toll with time for the December 26, 2004, Sumatra-Andaman earthquake. (a) Temporal variations and fitting of the death toll reported on the Internet; (b-d) As for (a), but using only the early data points to estimate the final death toll (i.e., the data from before the time of the vertical dashed line). (e) On-going estimates of the final death toll by the QRE, according to Equation (3) (squares), and the non-linear 'complete' curve fitting, according to Equation (1) (stars). Empirically, α is taken as 0.14 for the QRE for the first 10 days. From the end of the 10th day, N_0 is estimated by nonlinear curve fitting for each 3 days, with error bars representing the uncertainty of the estimate. If there is no new data, then the estimate is taken as the same as that estimated 3 days before. For reference, the final death toll reported by the media is shown by the horizontal dashed line, indicated as N_0^{media}. The red vertical dashed line in (e) marks the second day QRE, with the blue bar showing the '2nd-day-QRE times 2 to 3 rule' estimate.

agencies were different from each other. Despite all of these shortcomings, this set of 'crude' data still have a unique practical significance.

Figure 1a shows the temporal variation of the death toll reported on the Internet for the 2004 Sumatra-Andaman earthquake and the Indian Ocean tsunami (based on Liu and Wu [2005], Zhao et al. [2008], Li et al. [2011], revised with new data), together with curve fitting using Equation (1). The fitting uses the MatLab toolbox. Figure 1b-d shows the curve fitting using only the first part of the data, for different periods; i.e. to estimate the final value using only the early data. It can be seen that with the inclusion of more and more

data, the estimate has smaller and smaller uncertainties, and comes closer and closer to the final value. However, for the first several days, the fitting for the entire timescale does not work, and we have to consider a QRE.

According to the relation between coefficient α and the magnitude of the earthquake, in the case of this earthquake, an $\alpha = 0.14$ is chosen and used for the QRE calculation. Figure 1e shows the QRE (squares), together with the 'complete' fitting (stars). Taking the '2nd-day-QRE times 2 to 3 rule', a range for the early estimate can be given. Moreover, the result can be improved with time, as do the uncertainties of the estimate available.

Figure 2. Variations in the reported death toll with time for the March 11, 2011, Tohoku, Japan, earthquake. For details, see legends to Figure 1.

For the March 11, 2011, Tohoku, Japan, earthquake and tsunami, in the same way as for Figure 1, Figure 2 shows the fitting and the estimate. For this earthquake, the coefficient used for the QRE is taken as $\alpha = 0.14$. It can be seen from the evaluated time series that these two tsunami cases are quite similar to each other. In comparison with the cases of the earthquakes [Li et al. 2011], these two tsunamis both have much longer characteristic times, which reflects the differences between earthquake and tsunami rescue. Note that the empirically prescribed α value for the QRE is different from that obtained by the 'complete curve fitting', which reflects the empirical and approximate nature of the QRE. Taking the '*2nd-day-QRE times 2 to 3 rule*', an early estimate gives $0.97 \sim 1.46 \times 10^4$, which is consistent with the real situation to within one order of magnitude. In contrast, some of the early estimates based on 'static' models can be off by up to three orders of magnitude. As time lapses, the estimate is revised with the new incoming data. Figures 1e

and 2e indicate that a reliable estimate by curve fitting can be expected within 7 days of an earthquake.

Conclusions and discussion

An earthquake/tsunami emergency response program needs an early estimate of the life loss, shortly after the occurrence of the destructive/tsunami-generic earthquake, with on-going up-dating of this estimate based on field reports. In the present study, we analyzed the case of the March 11, 2011, Tohoku, Japan, earthquake, using the reported casualties from the Internet, with reference to the 2004 Indian Ocean tsunami. A QRE of the final death toll that fits the real situation within one order of magnitude can be obtained 2 days after the earthquake. In comparison with 'static' methods that use the relationships between magnitude/intensity/collapsed buildings and the final death toll to 'predict' the final death toll [e.g., Ohta et al. 1983, Samardjieva and Oike 1992, Samardjieva and Badal 2002,

Wyss 2005], our 'dynamic' approach keeps 'adjusting' the estimate according to new reports, and thus can be used to complement the 'static' approaches.

This early-estimate and up-dating approach has several empirical elements underlying it, based on previously studied earthquake/tsunami cases [Li et al. 2011]: 1) The number of deaths reported on the Internet at time t, $N(t)$, is represented by $N(t) = N_0 [1 - \exp(-\alpha t)]$, in which N_0 is the final number of deaths reported by the news media, and α is the coefficient reflecting the rescue process. 2) The logarithm of coefficient α is reversely proportional to the surface wave magnitude of the earthquake. 3) Reasonable QREs can be obtained 2 days after for the occurrence of earthquakes, and after longer times for the occurrence of tsunamis. For tsunamis, a '2nd-day-QRE times 2 to 3 rule' can be used for the early estimate. All of these three considerations are in need of further investigation for other earthquakes and/or earthquake-generated tsunamis.

In using this method, the nature and limits of the information from the Internet have to be taken into consideration, which implies that at least at the present time, the technical readiness level of this approach is still low, and thus further testing is needed. The exact number of casualties can take a long time to become known, and sometimes this can even remains unknown forever because of missing people. Many rumors occur after a disaster, and these can be repeated by many sources of information, thus modifying death estimates. Disasters that affect closed societies might not yield meaningful results, as the Internet is of little use as a source of information in such cases. Therefore, the approach proposed in the present study can be used to complement, rather than to substitute, other approaches.

On the other hand, being able to adjust estimates as new information becomes available is an advantage of this 'dynamic' approach. Moreover, because of the use of Internet-reported messages, this exercise also has the potential to be used for public awareness and social participation. This can help to mitigate earthquake/tsunami disasters, and it is an interactive part of 'citizen seismology'. Through using, confirming, improving, or falsifying the empirical relations presented in this study, and through using the Internet-reported death tolls, people can arrive at their own estimate of the final death toll of an earthquake, and take decisions related to their actions in terms of participating with and aiding the emergency, rescue and reconstruction services. Importantly, participation in interactive actions can enhance the involvement of people, which is the kernel spirit of 'citizen seismology'.

Last but not the least, we include the caution that this study is only retrospective, and thus it has limited significance in the real 'prediction' perspective. Verification or falsification of the effectiveness of this method needs further tests and 'forward prediction' experiments.

References

Ashkenazi, I., B. Isakovich, Y. Kluger, R. Alfici, B. Kessel and O.S. Better (2005). Prehospital management of earthquake casualties buried under rubble, Prehospital and Disaster Medicine, 20, 122-133.

Badal, J., M. Vázquez-Prada and Á. González (2005). Preliminary quantitative assessment of earthquake casualties and damages, Nat. Hazards, 34, 353-374.

Gao, J.G. and Y. Jia (2005). A study on the time of promulgating earthquake disaster – An index of earthquake rescue ability, Journal of Catastrophology, 20, 31-35 (Chinese with English abstract).

Gutiérrez, E., F. Taucer, T.D. Groeve, D.H.A. Al-Khudhairy and J.M. Zaldivar (2005). Analysis of worldwide earthquake mortality using multivariate demographic and seismic data, Am. J. Epidemiol., 161, 1151-1158.

Koshimura, S., T. Katada, H.O. Mofjeld and Y. Kawata (2006). A method for estimating casualties due to the tsunami inundation flow, Nat. Hazards, 39, 265-274.

Lamontagne, M. (2008). Casualties directly caused by an earthquake in Canada: First contemporaneous written accounts from the M6.5 Charlevoix, Quebec, earthquake of 20 October, 1870, B. Seismol. Soc. Am., 98, 1602-1606.

Li, Y.C., Z.L. Wu and Y.Z. Zhao (2011). Estimating the number of casualties in earthquakes from early field reports and improving the estimate with time, Nat. Hazards, 56, 699-708; doi: 10.1007/s11069-010-9583-9.

Liu, Z. and Z.L. Wu (2005). A simple model of reported casualties during earthquakes and earthquake-generated tsunamis, Earthquake Research in China, 21, 526-529 (Chinese with English abstract).

Liu, Z., Z.L. Wu, M.J. Cai and C.S. Jiang (2005). Discussion on casualties during 2004 earthquake-generated tsunami in Indian Ocean, In: S.C. Li, Y.J. Wang and P. Huang (eds.), Progress in Safety Science and Technology, Vol. 5, Beijing, Science Press, 2075-2077.

Lomnitz, C. (1970). Casualties and behavior of populations during earthquakes, B. Seismol. Soc. Am., 60, 1309-1313.

Ohta, Y., N. Goto and H. Ohashi (1983). An empirical construction of equations for estimating number of victims by earthquake, Zisin II, 36, 463-466 (in Japanese).

Samardjieva, E. and K. Oike (1992). Modelling the number of casualties from earthquakes, Journal of Natural Disaster Science, 14, 17-28.

Samardjieva, E. and J. Badal (2002). Estimation of the expected number of casualties caused by strong earthquakes, B. Seismol. Soc. Am., 92, 2310-2322.

Spence, R., E. So and C. Scawthorn, eds. (2011). Human Casualties in Earthquakes: Progress in Modelling and Mitigation, Amsterdam, Springer, 1-322.

Tsai, Y.B., T.M. Yu, H.L. Chao and C.P. Lee (2001). Spatial distribution and age dependence of human-fatality rates from the Chi-Chi, Taiwan, earthquake of 21 September, 1999, B. Seismol. Soc. Am., 91, 1298-1309.

Wu, X.Y., J.H. Gu and H.Y. Wu (2009). A modified exponential model for reported casualties during earthquakes, Acta Seismologica Sinica, 31, 457-463 (Chinese with English abstract).

Wyss, M. (2005). Human losses expected in Himalayan earthquakes. Nat. Hazards, 34, 305-314.

Zhao, Y.Z., Z.L. Wu and Y.T. Li (2008). Casualty in earthquake and tsunami disasters: Internet-based monitoring and early estimation of the final death toll, In: Proceedings of the 14th World Conference on Earthquake Engineering (WCEE; Beijing, October 12-17, 2008), CD, Paper ID 09-01-0044.

★Corresponding author: Zhongliang Wu,
Institute of Geophysics, China Earthquake Administration,
Beijing, China; email: wuzl@cea-igp.ac.cn.