Predicting Post Traumatic Stress symptom prevalence and local distribution after an earthquake with scarce data.

Francisca Dussaillant, PhD, School of Government, Universidad del Desarrollo

Mauricio Apablaza, PhD, School of Government, Universidad del Desarrollo

Abstract

Background. After a major earthquake, the assignment of scarce mental health emergency personnel to different geographic areas is crucial to the effective management of the crisis. The scarce information that is available in the aftermath of a disaster, may be valuable in helping predict where are the populations that are in most need.

Objective. To derive algorithms to predict Post Traumatic Stress (PTS) symptom prevalence and local distribution after an earthquake and test whether there are algorithms that require few input data and are still reasonably predictive.

Methods: A rich database of PTS symptoms, informed after Chile's 2010 earthquake and tsunami was used. Several model specifications for the mean and centiles of the distribution of PTS symptoms, together with PTS disorder prevalence were estimated via linear and quantile regressions. The models varied in the set of covariates included.

Results: adjusted R^2 for the most liberal specifications (in terms of numbers of covariates included) ranged from 0.62 to 0.74, depending on the outcome. When only including peak ground acceleration (PGA), poverty rate, and household damage in linear and quadratic form, predictive capacity was still good (adjusted R^2 from 0.59 to 0.67 were obtained).

Conclusions: information about local poverty, household damage and peak ground acceleration can be used as an aid to predict Post Traumatic Stress (PTS) symptom prevalence and local distribution after an earthquake. This can be of help to improve the assignment of mental health personnel to the affected localities.

Keywords

Chile, natural disaster, earthquake, post-traumatic stress, disaster relief.

Abreviations

PTS: Post traumatic Stress

PTSD: Post Traumatic Stress Disorder

EPT: Post Earthquake Survey (Encuesta Post Terremoto)

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition

NOAA: National Oceanic and Atmospheric Administration

USGS: United States Geological Survey

MMS: Modified Mercalli Scale

PGA: Peak Ground Acceleration

1. Introduction

Managing the psychological impact of a disaster is a critical public health challenge. The informed incidence and prevalence of mental health disorders are essential to effective service planning in the aftermath of a disaster [1, 2]. But, as recent academic literature discusses, there is a shortfall of reliable systems for translating available data into public health tools [3]. This gap in knowledge must be filled, especially since such systems may be of significant use for emergency response planning. One of the biggest challenges faced by disaster researchers and disaster management and prevention practitioners is identifying the population at risk as precisely as possible [4].

In terms of mental health, post-traumatic stress disorder (PTSD) is probably one the most frequent and debilitating consequences of a disaster [5]. The disorder can become chronic and enduring, with lifelong effects that might even escalate in time [1,2]. But research has shown that early interventions have some effectiveness in the prevention of the disorder [6,7,8]. Therefore, PTSD should be one of the main targets of emergency mental health interventions in the aftermath of a disaster. However, these interventions are costly. To improve their cost-effectiveness, the right choice of intervention targets is central and achieving this is important to predict population risk [9]. The predictability of local aggregate measures of PTS (post traumatic stress) symptoms is valuable in that it points to localities where the problem may become most extended. Thus, it aids disaster management professionals in the challenge of assigning scarce mental health personnel to different geographic locations.

The enormous individual heterogeneity of response to the environmental shock, in terms of the emergence (or not) of PTS symptomatology, has not been duly understood even now. Yet, some risk factors have been identified through abundant research, including several reviews [1,5,10,11,12,13]. For example, the development of PTSD has been consistently found to correlate with disaster exposure, type and severity. Human loss (death of a relative) and physical injury have also been found to be associated with the symptoms. Material loss, especially (but not exclusively) housing damage, is related to the occurrence of post disaster PTSD. Previous history of mental problems has also been found to be closely associated with the development of PTSD in the aftermath of a disaster. Females appear to be more prone to acquire the disorder, whereas the elderly seem to be more resilient. Socioeconomic status and poverty have been found to be risk factors for post disaster PTSD.

Although these risk factors have been identified, even after accounting for the standard socio-demographic controls or the abovementioned risk factors, it is still very difficult to predict whether a specific person will suffer post-traumatic stress. Human beings are heterogeneous in multiple dimensions not accounted for in most studies, such as the biological predisposition to mental disorders, cognitive and emotional types, and personality. These factors influence mental health in ways that are only partially understood [2,9,13]. This unobserved individual heterogeneity makes it difficult to predict the emergence of PTSD at the individual level using standard covariates. But when predicting aggregate measures of PTSD (i.e. local mean scores, prevalence) the within group variation is removed and better predictive power may be obtained. This is what is intended to be checked through this study.

This analysis tries to advance in this line of research by deriving simple algorithms to predict the prevalence of PTSD and the distribution of symptoms in locations where an earthquake has struck. The starting point in the search for this "rules of thumb" is a very rich database with plenty of possibilities to model post-traumatic stress prevalence. There is also, therefore, the possibility to estimate more complex models to start with.

But the main interest of this specific study is to derive simple but predictive algorithms to be applied in an emergency context. This, and not a complex algorithm requiring difficult-to-find data, is what is needed in a real world setting to rapidly assign emergency mental health professional assistance to the different locations hit by a disaster. The objective is to derive a predictive algorithm under the assumption of data scarcity, and therefore efforts will be made to obtain good prediction with the least possible information requirements. Data averaged at the local level is more immediately available after the disaster strikes, compared to individual level information. This, since local poverty levels, local average educational levels, or local unemployment statistics are usually obtained through representative surveys where only a fraction of the population is interviewed. Therefore, individual level data will probably not be readily available after the disaster strikes, but aggregate level data may be easier to obtain. It is for that reason that the paper focuses on aggregate level predictors.

2. Methods

2.1 Data

The *Post-Earthquake Survey* (EPT, Spanish acronym) database contains longitudinal (two-panel) data about the same persons before and after a major disaster [15]. The database, that can be downloaded from the Ministerios de Desarrollo Social site (http://observatorio.ministeriodesarrollosocial.gob.cl/enc_post_basedatos.php), are innominated, as compromised in the informed consent signed by the responders.

It encompasses nationally representative data from a household survey gathered in November and December 2009, a few months before the 2010 earthquake and tsunami that hit Chile. The database was complemented by post-disaster follow-up information, since the Chilean government re-interviewed a representative subsample of 22,456 of the original 71,460 households between May and June 2009.

The follow-up asked about several disaster-related and socioeconomic issues, and respondents were requested to complete the Davidson trauma battery [16], a self-report instrument used to evaluate post-traumatic stress symptoms.

This trauma battery consists of seventeen items, each corresponding to a post-traumatic stress symptom, as described by the DSM-IV. The questions were worded so that PTSD symptoms were assessed specifically in relation to the earthquake/tsunami. Each item is rated twice on a five-point scale, once in terms of frequency (increasing from "not at all" to "every day") and once in terms of severity (increasing from "not at all distressing" to "extremely distressing"). The Appendix provides a list of the 17 items. Since the frequency scores for each item range from 0 to 4 and the severity scores for each item also range from 0 to 4, the total score per item ranges from 0 to 8 points. When adding up the scores for the 17 items a PTS total is obtained, which ranges from a minimum score of 0 to a maximum of 136. Only respondents who were present at the moment of the interview were asked to answer the battery, and this resulted in 23907 valid PTS score values for individuals aged 18 or older. At least one adult person from 21059 of the households included in the sample responded to the battery.

Several municipality level variables were generated using the EPT. Poverty, unemployment and rurality prior to the earthquake were constructed as the weighted average of individual indicator variables. EPT defines as rural the zones with less than 1000 inhabitants or the zoenes with between 1000 and 2000 inhabitants in which less than 50% of the active population work in the secondary or tertiary sectors [17]. Local inequality indexes (Gini Index and Theil Index) prior to the earthquake were estimated using a weighted measure of the total household income. The proportion of complete household destruction and the proportion of severe household damage variables were constructed using the responses to a survey item that asked each respondent whether their house had been completely destroyed by the earthquake or the tsunami, had been severely damaged, had undergone some minor damage or no damage at all.

The EPT data was complemented with information on the strength of the earthquake and the tsunami, the history (and intensity) of aftershocks, and death rate at municipality level (203 municipalities). The intensity of the earthquake was quantified through peak ground acceleration (PGA), a measure that describes, in a broad sense, how hard the earth shakes in a given geographic area. Using the values provided by the United States Geological Survey (USGS), a research team leaded by José Zubizarrieta estimated the PGA in each of the municipalities where the EPS was collected[14]. They obtained one value for each municipality using the PGA grid provided by the USGS. Municipality values correspond to the inverse distance weighted average of the three closest grid estimates. Their interpolated data was used in the estimation.

To measure the intensity of the tsunami local geo-referenced data on height of the waves and horizontal inundation, from the Global Historical Tsunami Database at the National Geophysical Data Center (NOAA)[18], was used. The highest wave registered on the coast of each municipality and the longest inundation record were counted in when more than one observation was documented. In the case of locations with no information, the tsunami data was interpolated according to a north-south rank ordering of the coastal municipalities. In coastal locations where there was no tsunami information in any nearby municipality it was assumed that there was no alteration of the sea and assigned a value of 0 to the indicators. The same value of 0 was assigned to non-coastal municipalities.

Local data of intensity (measured in the Modified Mercalli Scale- MMS), date and location of the earthquake aftershocks occurring between February 27th and May 1st, 2010, when EPT fieldwork commenced were also available for this study. This data was obtained from the United States Geological Survey [19] that gathered the data using the DFYI method [20,21]. For municipalities where there was no measurement, the mean intensity for the province (the administrative division that follows in size, grouping several municipalities) was assigned. Where there was no measurement at the provincial level, null intensity (i.e. no aftershock) was assumed. Several variables that grouped aftershocks by intensity and counted them as they occurred between February 27th and May 1st, were constructed.

Finally, the number of deaths per municipality (due to the earthquake/tsunami) was obtained from the Statistics Unit at the Chilean Forensic Services Department [22]. These data was converted into death rates by using the municipality population figures obtained from the Chilean National Institute of Statistics [23].

Table 1 provides a list of variable names and explains their particular construction.

2.2 Statistical Methods

Several statistical analyses were performed using STATA 13.0. Different methods were used, depending on the variable to be predicted: PTS scores averaged at the municipality level, measures of the prevalence of PTS scores above thresholds, or centiles of the municipal PTS score distribution. To obtain these aggregate measures the weights provided in the survey were utilized. Since not every household integrant but only those present at the time of the interview responded to the PTS battery, some doubt may arise regarding the convenience of sample weight use to generate the aggregates. As a background check, parallel analyses with un-weighted aggregates (not reported) were performed, and returned similar results.

a. Predicting PTS average scores (within municipality): linear models were estimated with ordinary least squares with robust errors. Equations 1 to 4 depict the estimation process, with Y_{ij} representing individual PTS scores and X_j a set of covariates that will be described in the *Estimation* section of this paper. Coefficients estimated for the specification described in equation 1, $\hat{\alpha}$ and $\hat{\beta}$, were used to generate municipality level aggregate predictions (equation 2). These predictions were then regressed with the empirical municipality-level average values ($\bar{Y}$ as defined in equation 3) to check the similarity between empirical and estimated aggregates, as shown in equation 4. Coefficients of determination (R^2) and root mean square errors (RMSE) for this last step are the measures of goodness of fit that were chosen to report.

$$Y_{ij} = \alpha + \beta X_j + e \tag{1}$$

$$\hat{Y}_j = \hat{\alpha} + \hat{\beta} X_j \tag{2}$$

$$\bar{Y}_j = \frac{1}{N_j} \sum_{i=0}^{N_j} Y_{ij} \tag{3}$$

$$\bar{Y}_j = \delta \hat{Y}_j + u \tag{4}$$

b. Predicting PTS prevalence: prevalence was measured as the proportion of the sample that got a PTS score above a certain threshold. Thresholds were set at 20, 30, 40 points on the Davidson scale. The team that validated the Spanish version of the battery, proposes a cut-off score of 40 as the most efficient to determine clinical PTSD [24]. Nevertheless, several authors indicate that sub-syndromal PTSD does imply some form of disability (sometimes similar to that of the full-blown disorder) which deserves further study [1, 25, 26]. It was, therefore, decided to study lower thresholds too (cut-off score of 30 and of 20). Models for prevalence were estimated at municipality level in one stage using ordinary least square regressions with robust errors, as depicted in equation 5. Here, P_j represents prevalence level in any of its definitions. Adjusted R^2 and RMSE (root mean square error) for these estimations are the goodness of fit measures we chose to report.

$$P_j = \delta X_j + u \qquad (5)$$

c. Predicting centiles of the local PTS score distribution: models for the 90th, 80th, 70th and 60th quantile of the distribution of PTS scores were estimated. Disaster exposure has a strong but heterogeneous effect on PTS symptoms. In representative samples, the distribution of PTS symptoms is highly skewed to the right, meaning that only a few individuals are high scorers. This is still the case even after a major disaster. The evidence indicates that PTS symptoms are dramatically but unevenly high among residents of strongly affected areas [14]. This is why an examination of the higher deciles of the score distribution might shed some light into understanding the phenomenon. To achieve this, the method for quantile regressions was used [27]. The process is similar to that described in subsection a. of this section: on a first step the model was estimated by quantile regression using individual level PTS scores as dependent variables, and municipal level covariates. A predicted value for the centile was obtained for each municipality. At the same time, the observed centile was obtained from the empirical distribution of each municipality. Finally, a regression of empirical centiles on predicted centiles is estimated. R^2 and RMSE for this last estimation are the goodness of fit measures chosen to be reported.

2.3 Estimation

Each of the aforementioned methods was applied to several sets of covariates, as shown in table 1 (the table also contains a brief description of each of the covariates used throughout the analysis, and how they are measured). The objective was to identify a parsimonious model with the predictors which, in the context of a disaster, are standard and the easiest to find. Since the number of variables at hand was manageable, variable inclusion and exclusion was performed manually, and the assessment of the model was guided mainly by human expertise. Forward stage-wise regression, lasso methods and least angle regression [28] were also used for the intermediate assessment of whether there was any important quality of the data that was being missed (results not reported). In these assessments *PGA* and some household damage variable (*Destruction, Damage or Destruction+Damage)* were always selected as the most informative, regardless of which of the methods was used. However, these statistical methods are not suitable for our final purposes, which include, in the choice of predictors, an assessment of the local availability of covariates.

Covariate sets described in table 1 include sometimes subsets of variables that are highly collinear. This was observed while conducting the analyses. Variance inflation factors (VIF) were estimated for each of the sets with mean results that are frequently above standard thresholds [29]. High collinearity of the regressors is a problem in this type of analyses. To investigate the magnitude of the problem, accuracy of the models was studied using 10-fold cross-validation [30,31] of the R^2 obtained in the first stage of each estimation. Results (not reported) indicate that the standard error of the R^2 estimate is at or below 0.015 in most models.

3. Results

Table 2 provides the descriptive statistics of the dependent and independent variables we used throughout the analyses. Although in some of the analyses used individual data as dependent variables, for the sake of space the table only describes the variables already aggregated to the municipality level. Also for the sake of space, descriptors for the quadratic forms of the variables or the addition of variables are not provided, although these transformations were included sometimes in the estimations.

The table is divided into two blocks of variables. The first describes the set of dependent variables we use throughout the analyses, and the second describes the set of covariates.

Table 3 shows results from the estimation of the different models. It shows the proportion of variance (adjusted R^2) that is explained by the different sets of covariates described in table 1. Adjusted R^2 was selected as a measure of fit since it is widely used and its scale is the same regardless of the scale of the independent variable. Table 4 reports the mean square error (RMSE), an alternative measure of fit. Since RMSE is scale-dependent, model comparisons based on this statistic are possible only across cells that belong to the same columns of table 4. To have an idea of magnitude of the statistic, it should be compared to mean, maximum and minimum empirical values as shown in table 2.

Finally, tables 5 and 6 show the estimated coefficients for two of the models estimated. These are the preferred specifications, given the constraints mentioned above. The choice of these preferred specifications is discussed in the next section of this document. Coefficients for the rest of the models are available from the authors upon request.

4. Discussion

The discussion section includes a discussion of main results and then proceeds to discuss the choice of model. It continues with the proposition of simple algorithms and finalizes by stating limitations of the research and providing directions for future research.

4.1. Discussion of main results

A quick inspection of table 3 shows that the different specifications are almost always capable of explaining more than 60% of the variance of the variables to be predicted.

- Local average PTS score: The highest R^2 (0.737) of all models is obtained when predicting the local average PTS score with the less restrictive of all the covariate sets of table 1 (covariate set 0). But many other simpler specifications of the model for local average PTS score can explain more than two thirds of its variance ($R^2 > 0.667$).

- Percentiles of PTS symptom distribution: fit improves for the higher percentiles (80th and 90th). This is good news since our main interest lies in predicting the right tail of the distribution. As discussed previously, a large majority of people will not present PTS symptoms even after a major natural disaster. The distribution of PTS symptoms is, therefore, highly skewed to the left and is not too informative about the real degree of the mental health problems that may have arisen with the disaster [14].

- Prevalence: Remember that a cut-off score of 40 is said to be the most efficient in the determination of clinical PTSD [24]. Nonetheless, two alternative measures that include some subclinical scores were defined; Three alternative prevalence values were constructed, one for the cut-score of 40 (*prevalence40*), another for the cut-score of 30 (*prevalence30*) and the last one for the cut-score of 20 (*prevalence20*). Table 3 shows that models for *prevalence30* achieve better fit than its alternatives and, in many covariate specifications, get R^2 values at or above 0.66. Models for *prevalence40* achieve lower R^2 but still, in most covariate specifications, the statistic is at or above 0.6.

4.2 Choosing a model

The task of choosing a model includes the need to take into account the potential availability or ubiquity of covariates. In the process of choosing a model it must be taken into consideration that some of the covariates are more difficult to find or require more detailed data than others. For example, *Horizontal Inundation* is easier to measure in steps of 200 meters than down to the nearest meter (and therefore the variable *Horiz200* is preferred to *Horiz* in tables 1 and 2 as a covariate). Also, length of inundation, *Horiz200* or *Horiz* in tables 1 and 2, is preferred to height of the highest wave (*Water Height*), since the latter will not be observable in the aftermath of the disaster. Inequality indexes (*Gini*, *Theil*) and *Unemployment* are more difficult to estimate than *Poverty*, since the latter is constructed by the aggregation of simple indicators of whether an individual is or is not poor. Regarding the aftershock covariates, the sum of aftershocks that are clearly perceived by the population (MMS 4 and over, see tables 1 and 2) are preferred to a disaggregated group of variables indicating each the frequency of aftershocks of a certain intensity (1 to 2 MMS, 2 to 3MMS, etc.). Overall, variables related to the aftershocks are not the most preferred because their construction requires that some time should elapse in the aftermath of the disaster. As already mentioned, the expectation is to make predictions as soon as possible after the disaster strikes. In addition, peak ground acceleration (*PGA*) is available worldwide from USGS at short notice after the earthquake has struck, and household destruction is evidenced immediately (although quantifying it in detail is more difficult and therefore the variable that combines complete destruction and severe damage is preferred, instead of considering them separately; see table 1). Household damage variables are preferred to *Death Rate*, since accurate information about the latter will be available within a few days or weeks. Nevertheless, a rough estimation of *Death Rate* can be obtained with some speed.

With this in mind, inspection of tables 3 and 4 can be performed, in search of a set of covariates ensuring a reasonable fit while at the same time comprising information relatively easy to get in the aftermath of a disaster. In this analysis covariate set 0 (see table 1) should be used as the reference, since it is the less restrictive in terms of covariate choice. Therefore, covariate set 0 yields the best predictions, independently of the variable that is being explained. However, availability of these covariates in the aftermath of a disaster is unlikely. Something similar happens with the simpler specification that uses covariate set I. Some predictive power should be sacrificed in order to make the predictions more attainable.

The close inspection of tables 3 and 4 indicates that covariate set III seems to work better than the less parsimonious set II, independently (with few exceptions) of the variable to be explained. However, covariate set III is still too liberal for our purposes. When comparing the results from covariate sets IV to XII a similar model fit is found, once again, independently of the variable being predicted. Covariate set IX is especially interesting since it seems to dominate the rest regardless of the variable being explained (except when the dependent variable is the 90^{th} percentile of the PTS score local distribution). Covariate set IX includes PGA, percentage of households destroyed or severely damaged, death ratio and local poverty levels. Of these, death ratio is maybe the most difficult to find in at the immediate aftermath of a disaster. If excluded, models can be estimated using covariate set

XII without sacrificing too much predictive power. Covariate sets XIII and XIV are much worse at predicting, meaning that it would not be advisable to use PGA and household destruction alone to guess PTS prevalence or score distribution.

Covariate sets XV to XVII were included to check whether entering the main covariates in a linear (and not polynomial) fashion would suffice. A linear specification is preferred to the quadratic form since it would give a very straightforward rule of thumb for calculation. But tables 3 and 4 show clearly that all these linear specifications are outstripped by the very simple specification XIII, which only includes PGA and household destruction data, both entering in a linear plus quadratic form.

These considerations lead us to conclude that our preferred models are those comprising covariate sets IX or XII. Linear specifications are not advisable since they are overcome by covariate set XIII, giving rise to a very simple two-covariate (*PGA* and *Destruction+Damage*) model where each covariate enters as a second-degree polynomial. But still, specification XIII is easily improved when *Poverty* is added as a predictor (and further improved with the *Death Ratio*).

The three covariate sets at the end of the list (XVIII, XIX and XX) were devised to answer several questions that arise after having chosen sets IX and XII as the preferred covariate sets. First of all, covariate set XVIII permits to check how model fit improves specification IX when the *death ratio* enters the equations as a second degree polynomial instead of only linearly. Results indicate that although fit is slightly improved with the new specification, improvement is low in absolute terms. Since *Death Ratio* data will probably be a very rough estimate (if it is available at all) in the immediate aftermath of a disaster, it may be preferable to insert it in the model only in a linear form.

Covariate sets XIX and XX are alternative specifications devised to check how model fit is improved (compared to our preferred specification XII) when the *Destruction+Damage* covariate is separated into its components. In covariate set XIX, both *Destruction* and *Damage* enter linearly, and in covariate set XX, both enter as polynomials. Specification XIX is worse than specification XII across models. It is preferable, then, to use one rough measure of *Destruction+Damage* alone as long as it enters the model as a quadratic polynomial. When each component of *Destruction+Damage* enters separately and as polynomials we are in specification XX. The latter is superior to specification XII only in predicting the average PTS score. The main interest of this research, though, lies in predicting the right tail of the distribution. With this in mind, preferred specifications are still those that use covariate sets IX and XII.

4.3 Simple algorithms

Tables 5 and 6 provide the coefficients that arise from the preferred specifications that have been just been decided upon. As discussed above, only PGA, Poverty, *Destruction+Damage* and *Death Rate* are the covariates included in the final choices. The importance of these specific covariates in the assessment of the risk of PTSD has been documented in previous literature. The exposure level has been documented as a fundamental determinant of mental health disorders [32, 33]. Specifically, the relevance of earthquake intensity as an important predictor of PTSD and other mental disorders has been documented for disasters similar to Chile's 2010 in other latitudes (for example, local effects were found in the analysis of the Christchurch 2010/11 earthquake [34]). Also, severe, lasting and pervasive mental health effects have been found to be associated with the degree of damage to property, loss of lives and socioeconomic status [12, 33].

Estimated coefficients on tables 5 and 6 can be used as simple algorithms to predict PTS symptom prevalence and distribution. Estimation is straightforward when the data is available: the practitioner must multiply each coefficient with the corresponding variable and add the results. For example, according to table 5, to predict average PTS score in a location where PGA is 22 (g/100), rate of poverty is 0.5 and 20% of the households were destroyed or severely damaged, with no information about the death rate, the calculation would be $0.947+0.072x25+0.006x25^2+69.297x0.2-65.658x0.2^2+48.070x0.25-87.580x0.25^2$, and the prediction would render a local average of 39.4 PTS score points. Since we are predicting Average PTS score using covariate set IX, a root mean square error of 5.823 PTs score points can be associated to this estimation from table 4. But, more interesting than the point estimates are the comparisons among locations that can be made using this tool. In other words, locations can be identified in terms of the *relative* importance of the mental health problem to be tackled, and such comparisons may be used to assign mental heath personnel.

It must be kept in mind that both the dependent variables and the covariates utilized in the analyses are local aggregates. Even though it was possible to make very accurate predictions, it should not be forgotten that what was being predicted were aggregates and not individual outcomes. Therefore, if the algorithm informs that a certain proportion of the local adult population will display PTS symptoms, individual subjects must still be screened using other tools. Estimation results contained in this document should not be used to make inferences about the predisposition of any particular individual(s) to have the condition. The fallacious nature of such inferences is extensively discussed in the literature [35].

4.4 Limitations of the study and suggestions for future research

External validity of the results of this paper should be checked. The predictive capacity of the covariates identified in this paper and the stability of the coefficients found in the estimations should be assessed across disaster contexts and in other geographies. More research must be made in order to assess whether these simple algorithms can be applied in any setting.

Also, the results of this study should not be read as identifying causality, since the estimations are only intended to reflect correlational associations between variables. Along these lines, it is important to note that only as long as pre-disaster PTS prevalence is uniform across locations it can be argued that this study of prevalence is actually informing about incidence of the symptomatology after the earthquake. The assumption seems plausible since some homogeneity (of aggregate statistics) is observed when locations not struck by the disaster are studied. This is a feature of the sample not reported in this paper due to space concerns. Detailed descriptive information of the sample is available from the corresponding author upon request. Moreover, although pre-disaster prevalence might have a role in the prediction of PTS symptoms in the aftermath, it is not likely that the high and significant coefficients for disaster-related variables in our estimations were only due to chance.

Good models to predict whether an individual will develop PTSD after a disaster are still required. Although some progress has been made towards that objective [9], this strand of research is still developing. If it was possible to predict PTSD accurately at the individual level quickly (without the need, costs or time-intensiveness of professional screening), post disaster mental health intervention could significantly improve.

Finally, PTSD is only one of several mental health problems that arise in the aftermath of a disaster [1]. The estimates derived herein only point to PTSD prevalence and symptom distribution, since no data on other anxiety disorders, depression, substance abuse, panic disorder or other mental disorders was available. Understanding the prevalence and local distribution of these other disorders is still an issue that deserves further research if we intend to achieve an optimal design of treatment services 1, 36. Nevertheless, there is evidence pointing to the fact that some disorders such as depression, dysthymia and substance abuse are frequently comorbid (and therefore highly correlated to) to PTSD [2, 36]. Although this point deserves more research, an estimation of PTSD prevalence might be a good proxy of prevalence of the main mental health problems that arise due to a disaster.

5. Conclusions

After a major earthquake, the assignment of scarce mental health emergency personnel to different geographic areas is crucial to the effective management of the crisis. The scarce information that is available in the aftermath of a disaster, may be valuable in helping predict where are the populations that are in most need.

The analyses reported in this paper show that it is possible to devise simple algorithms to predict PTS prevalence and local PTS score distribution even in a setting in which information is limited, a scenario that is likely in the immediate aftermath of a large-scale disaster. When only including peak ground acceleration (PGA), poverty rate, and household damage in linear and quadratic form, good predictive capacity was achieved. Simple algorithms to predict local prevalence and distribution of PTS symptoms using these variables were derived.

Algorithms that attain precise identification of individuals at high risk of PTSD or other mental disorders associated to disasters is one of the immediate challenges of research, not tackled in this study, which only studied local aggregates. Also, more research must be made in order to assess whether these simple algorithms can be applied in any setting.

Acknowledments

We thank Dr. José Zubizarrieta who kindly shared with us his estimates of the earthquake's peak ground acceleration for each municipality. We also thank Dr. Anthony Rosellini for his helpful comments on an initial draft of this paper.

This research used information from the *Encuesta Post Terremoto*. We thank the Chilean Ministerio de Desarrollo Social, intellectual proprietor of the survey, that made it available for this research. Results of this study are the sole responsibility of the authors and do not compromise the Ministerio de Desarrollo Social at all.

References

1. McFarlane, A. et al. (2009). Anxiety Disorders and PTSD. In: Yuval Neria et al.,eds., Mental Health and Disasters. Cambridge: Cambridge University Press. 2009:47-66

2. Ursano RJ et al. What is Psychopathology after Disasters? Considerations about the Nature of the Psychological and Behavioral Consequences of Disasters. In: Neria Y et al. (eds.) Mental Health and Disasters. Cambridge: Cambridge University Press. 2009:131-142

3. Samarasundera E, Hansell A, Leibovici D, Horwell C, Anand S, Oppenheimer C. Geological hazards: From early warning systems to public health toolkits. Health & Place 2014; 30: 116-119.

4. Dominici F, Levy J, Louis T. Methodological Challenges and Contributions in Disaster Epidemiology. Epidemiol Rev 2005; 27(1):9-12

5. Galea S, Nandi A, Vlahov D. The epidemiology of post-traumatic stress disorder after disasters. Epidemiol Rev 2005; 27 (1): 78-91

6. Forneris CA, Gartlehner G, Brownley KA et al. Interventions to prevent post-traumatic stress disorder: a systematic review. Am J Prev Med 2013; 44:635-50.

7. Kliem S, Kroger C. Prevention of chronic PTSD with early cognitive behavioral therapy. A meta-analysis using mixed-effects modeling. Behav Res Ther 2013; 51:753-61.

8. Shalev AY, Ankri Y, Israeli-Shalev Y et al. Prevention of posttraumatic stress disorder by early treatment: results from the Jerusalem Trauma Outreach And Prevention study. Arch Gen Psychiatry 2012; 69:166-76.

9. Kessler RC, Rose S, Koenen KC, Karam EG, Stang PE, Stein DJ, … Carmen Viana M. How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. World Psychiatry 2014, *13*(3): 265–274

10. Neria Y, Nandi A, Galea S. Post-traumatic stress disorder following disasters: a systematic review. Psychol Med. 2008; 38(4):467-80

11. Neria, Y. et al. Disaster Mental Health Research: Current State, Gaps in Knowledge, and Future Directions. In: Yuval Neria et al.,eds., Mental Health and Disasters. Cambridge: Cambridge University Press. 2009:594-610

12. Norris FH, Friedman MJ, Watson PJ, Byrne CM, Diaz E, Kaniasty K. 60,000 disaster victims speak: part I. An empirical review of the empirical literature, 1981–2001. Psychiatry 2002; 65:207–39.

13. Norris FH, Wind LH. The Experience of Disaster: Trauma, Loss, Adversities, and Community Effects In: Yuval Neria et al.,eds., Mental Health and Disasters. Cambridge: Cambridge University Press. 2009:29-46

14. Zubizarreta JR, Cerdá M, Rosenbaum PR. Effect of the 2010 Chilean Earthquake on Posttraumatic Stress Reducing Sensitivity to Unmeasured Bias Through Study Design. *Epidemiology 2013; 24*(1):79–87

15. Ministerio de Desarrollo Social Encuesta Post Terremoto 2010. Chile. Available at http://observatorio.ministeriodesarrollosocial.gob.cl/enc_post_basedatos.php . Published 2010. Accessed August 19, 2016.

16. Davidson JR, Book SW, Colket JT, et al. Assessment of a new self-rating scale for post-traumatic stress disorder. Psychol Med. 1997; 27:153–160.

17. Ministerio de Planificación. Encuesta Casen 2009. Manual del Encuestador. Observatorio Social Universidad Alberto Hurtado. Chile, 2009

18. NOAA. National Geophysical Data Center / World Data Service (NGDC/WDS): Global Historical Tsunami Database. National Geophysical Data Center, NOAA. doi:10.7289/V5PN93H7. Published 2015. Accessed August 16, 2016.

19. United States Geological Survey (USGS), DYFI Archives, http://earthquake.usgs.gov/earthquakes/dyfi/archives.php [September 1st, 2015]

20. Atkinson GM, Wald DJ. "Did You Feel It?" intensity data: A surprisingly good measure of earthquake ground motion. Seism. Res. Lett. 2007; 78: 362-368

21. Wald DJ, Quitoriano V, Dengler LA, Dewey JW. Utilization of the Internet for Rapid Community Intensity Maps. Seismological Research Letters 1999, 70(6):680-697.

22. Nahuepan E,Varas J. El terremoto/tsunami en Chile. Una mirada a las estadísticas médico legales. Investigación forense 2013; 2:113-129

23. Instituto Nacional de Estadísticas. Estadísticas demográficas y vitales. Comunas: Actualización de Población 2002-2012 y Proyecciones 2012-2020. http://www.ine.cl/canales/chile_estadistico/familias/demograficas_vitales.php . Published 2015. Accessed August 19, 2016.

24. Bobes J, Calcedo-Barba A, García M, François M, Rico-Villademoros E, González M, Bascarán M, Bousoño M, Grupo español de trabajo para el estudio del trastorno por estrés postraumático. Evaluación de las propiedades psicométricas de la versión española de cinco cuestionarios para la evaluación del trastorno de estrés postraumático. Actas Especialidad Psiquiatría 2000; 28(4): 207 - 218.

25. Stein MB, Walker JR, Hazen AL, Forde, DR Full and partial posttraumatic stress disorder: Findings from a community survey. American Journal of Psychiat*ry 1997*; *154*(8):1114–1119.

26. Milliken CS, Auchterlonie JL, Hoge CW. Longitudinal assessment of mental health problems among active and reserve component soldiers returning from the Iraq war. JAMA 2007; 298(18):2141–2148.

27. Koenker R. Quantile Regression. Cambridge: Cambridge University Press. 2005.

28. Efron B,Hastie T,Johnstone I, Tibshirani R. Least angle regression. Ann. Statist. 2004; 32(2):407-499.

29. O'Brien RM. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity 2007; 41(5):673.

30. Geisser S. The predictive sample reuse method with applications. J. Amer. Statist. Assoc. 1975; 70:320–328.

31. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statistics Surveys 1991; 4: 40-79.

32. North C. Epidemiology of disaster mental health. In: Robert J. Ursano et al. (eds.) Textbook of Disaster Psychiatry. Cambridge: Cambridge University Press. 2007: 29-47.

33. Schultz JM, Espinel Z, Galea S, Reissman D. Disaster ecology: Implications for disaster psychiatry. In: Ursano RJ, Fullerton CS, Weisaeth, L, Raphael B., eds., Textbook of disaster psychiatry. Cambridge: Cambridge University Press. 2007:69-96.

34. Hogg D, Kingham S, Wilson T, Griffin E, Ardagh M. Geographic variation of clinically diagnosed mood and anxiety disorders in Christchurch after the 2010/11 earthquakes. Health & Place 2014; 30: 270-278.

35. Piantadosi S, Byar D, Green S The ecological fallacy. American Journal of Epidemiology 1988; 127(5):893-904.

36. Galea S, Maxwell A. Methodological Challenges in Studying the Mental Health Consequences of Disasters. In: Yuval Neria et al, eds, Mental Health and Disasters. Cambridge: Cambridge University Press; 2009: 579-593.

Table 1. Sets of covariates used in the estimations. L=enters linearly; QP=enters as a quadratic polynomial; CP= enters as a cubic polynomial; Fn=enters as a factor variable with *n* factors. Preferred specification in grey.

Set # / **Covariate**	**0**	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**	**XIII**	**XIV**	**XV**	**XVI**	**XVII**	**XVIII**	**XIX**	**XX**
PGA	CP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	L	L	L	QP	QP	QP
Destruction	CP	QP	QP	QP																L	QP
Damage	CP	QP	QP	QP																L	QP
Destruction + Damage					QP	QP	QP	QP	QP	QP	QP	QP	QP	QP		L	L	L	QP		
Death Ratio	CP	L	L	L	L	L	L	L	L	L						L			QP		
Horiz	CP	QP	QP	QP	QP																
Horiz200						F3	F3	F3	F3		F3					F3					
Water Height	CP	QP	QP																		
Aftershocks (7 vars)	L	L																			
Aftershocks 4 to 8 MMS			L	L	L	L		L								L					
Rurality	CP	QP	QP	QP	QP	QP	QP					QP				L		L			
Poverty	CP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP	QP			L	L		QP	QP	QP
Unemployment	CP	QP	QP	QP	QP																
Gini	CP	QP	QP	QP	QP	QP															
Theil	CP	QP																			

PGA=peak ground acceleration
Destruction = proportion of completely destroyed households
Damage = proportion of households with severe damage (but not completely destroyed)
Rurality= proportion of the adult (>=18) population that lives in rural zones.
Death Ratio = deaths per 10000 local inhabitants
Horiz= length of entry of the sea into the land. In meters.
Hori200= simplified version of horiz. The value of horiz is rounded to the nearest multiple of 200.
Water Height= height of the highest tsunami wave recorded in the coast of the locality. In meters.
Poverty= proportion of the population 18 or older that falls below the poverty line.
Aftershocks (7 vars)= seven variables each describing the number of aftershocks of a certain intensity (1 to 2 MMS; 2 to 3 MMS; 3 to 4MMS;…;7 to 8MMS)
Aftershocks 4 to 8 MMS = total number of aftershocks from 4MMS to 8MMS between February 27 and May 1st
Unemployment= proportion of the local active population that is currently unemployed.
Gini= Gini Index of Inequality
Theil= Theil index of inequality

Table 2. Descriptive statistics of the variables used in the analyses (N=203)

	Variable		Mean	Std. Dev.	Min	Max
Dependent variables	PTS (municipality average) (PTS points)		15,858	10,142	0,525	44,947
	PTS (perc90 of municipality distribution) (PTS points)		41,079	22,684	0	99
	PTS (perc80 of municipality distribution) (PTS points)		27,507	18,415	0	78
	PTS (perc70 of municipality distribution) (PTS points)		19,867	15,190	0	67
	PTS (perc60 of municipality distribution) (PTS points)		13,956	11,994	0	55
	PTS municipality prevalence (cut-score at 20 PTS points)		0,274	0,181	0	0,727
	PTS municipality prevalence (cut-score at 30 PTS points)		0,186	0,144	0	0,630
	PTS municipality prevalence (cut-score at 40 PTS points)		0,124	0,112	0	0,484
Covariates	Peak ground acceleration (*PGA*) (g / 100)		19,286	10,099	0	32
	Proportion of destroyed households (*Destruction*)		0,031	0,061	0	0,363
	Proportion of severely damaged households (*Damage*)		0,080	0,076	0	0,318
	Deaths per 10000 inhabitants (*Death Ratio*)		0,008	0,036	0	0,356
	Horizontal Inundation from Tsunami (*Horiz*) (meters)		0,008	0,036	0	0,356
	Tsunami maximum wave height (*Water Ht*) (meters)		29,216	69,966	0	376,542
	Number of aftershocks (February 27 to May 1st)	Intensity 1 to 2 (in MMS[1])	1,828	2,678	0	14
		Intensity 2 to 3 (in MMS)	17,778	3,081	0	99
		Intensity 3 to 4 (in MMS)	11,458	22,053	0	33
		Intensity 4 to 5 (in MMS)	3,123	9,712	0	16
		Intensity 5 to 6 (in MMS)	0,867	3,494	0	4
		Intensity 6 to 7 (in MMS)	0,187	0,984	0	2
		Intensity 7 to 8 (in MMS)	0,010	0,450	0	1
	RURALITY (proportion of the adult population)		0,406	0,435	0	1
	POVERTY (proportion of the adult population)		0,147	0,094	0	0,541
	UNEMPLOYMENT (proportion of the active population)		0,095	0,056	0	0,403
	GINI Inequality Coefficient		0,376	0,056	0,237	0,571
	THEIL Inequality Coefficient		0,261	0,103	0,092	0,854

[1]MMS: Modified Mercalli Scale

Table 3. Variance explained (adjusted R^2) for models aiming to predict the mean and upper centiles of the PTS score distribution, and PTS prevalence using 20, 30 and 40 points on the Davidson's scale as cut scores. Preferred specifications in grey (N=203)

Set of covariates (see table 1)	**Dependent variable**							
	PTS average[1]	60th percentile[2]	70th percentile[2]	80th percentile[2]	90th percentile[2]	Prevalence[3] (cut score20)	Prevalence[3] (cut score30)	Prevalence[3] (cut score40)
0	0,737	0,624	0,646	0,660	0,657	0,694	0,679	0,625
I	0,721	0,617	0,627	0,651	0,642	0,686	0,683	0,633
II	0,683	0,566	0,58	0,607	0,609	0,642	0,659	0,609
III	0,682	0,569	0,578	0,606	0,608	0,645	0,663	0,610
IV	0,671	0,574	0,575	0,604	0,597	0,641	0,658	0,606
V	0,670	0,569	0,577	0,604	0,606	0,642	0,658	0,607
VI	0,670	0,574	0,606	0,627	0,64	0,643	0,661	0,613
VII	0,670	0,584	0,605	0,623	0,631	0,647	0,660	0,610
VIII	0,669	0,577	0,611	0,631	0,64	0,645	0,660	0,610
IX	0,670	0,592	0,621	0,634	0,64	0,648	0,664	0,614
X	0,667	0,574	0,608	0,631	0,64	0,646	0,659	0,607
XI	0,669	0,575	0,611	0,63	0,641	0,645	0,662	0,615
XII	0,668	0,586	0,618	0,633	0,64	0,647	0,660	0,610
XIII	0,633	0,556	0,584	0,607	0,609	0,621	0,636	0,594
XIV	0,463	0,402	0,411	0,444	0,448	0,464	0,458	0,417
XV	0,637	0,534	0,546	0,56	0,576	0,629	0,643	0,589
XVI	0,628	0,542	0,559	0,573	0,587	0,627	0,634	0,575
XVII	0,615	0,496	0,532	0,555	0,563	0,606	0,620	0,576
XVIII	0,675	0,591	0,617	0,636	0,641	0,646	0,663	0,617
XIX	0,652	0,569	0,589	0,602	0,617	0,636	0,646	0,591
XX	0,676	0,576	0,609	0,62	0,635	0,649	0,663	0,611

[1] On a first stage Weighted OLS regressions with individual PTS score data were estimated (N=23907). On a second stage aggregate predicted values were compared to empirical aggregates through OLS regression at the municipality level. R^2 reported are for second stage.
[2] On a first stage weighted quantile regressions with individual PTS score data were estimated (N=23907). On a second stage aggregate predicted values were compared to empirical aggregates through OLS regression at the municipality level. R^2 reported are for second stage.
[3] OLS regressions with empirical local prevalence as dependent variable.

Table 4. Root Mean Square Error (RMSE) for models aiming to predict the mean and upper centiles of the PTS score distribution, and PTS prevalence using 20, 30 and 40 points in the Davidson's scale as cut scores. Preferred specifications in grey. (N=203)

Set of covariates (see table 1)	**Dependent variable**							
	PTS average[1]	60th percentile[2]	70th percentile[2]	80th percentile[2]	90th percentile[2]	Prevalence[3] (cut score20)	Prevalence[3] (cut score30)	Prevalence[3] (cut score40)
0	5,210	7,363	9,053	10,751	13,285	0,10009	0,08183	0,06879
I	5,366	7,432	9,287	10,895	13,580	0,10144	0,08124	0,06808
II	5,714	7,910	9,864	11,560	14,200	0,10838	0,08424	0,07024
III	5,724	7,888	9,886	11,569	14,208	0,10783	0,08382	0,07017
IV	5,825	7,839	9,913	11,602	14,402	0,10841	0,08436	0,07053
V	5,835	7,884	9,888	11,608	14,252	0,10837	0,08441	0,07043
VI	5,830	7,831	9,533	11,241	13,608	0,10800	0,08401	0,06987
VII	5,830	7,740	9,546	11,313	13,778	0,10750	0,08412	0,07017
VIII	5,838	7,800	9,471	11,183	13,617	0,10773	0,08411	0,07018
IX	5,823	7,662	9,348	11,147	13,616	0,10731	0,08370	0,06983
X	5,852	7,824	9,514	11,193	13,607	0,10763	0,08431	0,07041
XI	5,832	7,817	9,479	11,207	13,596	0,10772	0,08390	0,06970
XII	5,842	7,714	9,385	11,162	13,614	0,10747	0,08415	0,07016
XIII	6,141	7,990	9,800	11,538	14,184	0,11132	0,08705	0,07155
XIV	7,429	9,272	11,660	13,729	16,860	0,13236	0,10622	0,08581
XV	6,110	8,188	10,232	12,220	14,771	0,11010	0,08626	0,07205
XVI	6,185	8,119	10,083	12,036	14,580	0,11036	0,08735	0,07324
XVII	6,292	8,512	10,387	12,284	15,003	0,11356	0,08890	0,07316
XVIII	5,783	7,673	9,400	11,113	13,590	0,10753	0,08375	0,06954
XIX	5,986	7,873	9,737	11,623	14,040	0,10906	0,08583	0,07183
XX	5,771	7,807	9,493	11,356	13,697	0,10710	0,08374	0,07008

[1] On a first stage Weighted OLS regressions with individual PTS score data were estimated (N=57531). On a second stage aggregate predicted values were compared to empirical aggregates through OLS regression at the municipality level. R^2 reported are for second stage.
[2] On a first stage weighted quantile regressions with individual PTS score data were estimated (N=25949). On a second stage aggregate predicted values were compared to empirical aggregates through OLS regression at the municipality level. R^2 reported are for second stage.
[3] OLS regressions with empirical local prevalence as dependent variable.

Table 5. OLS regression results, average PTS score and Prevalences

	Average PTS score		**Prevalence (cut score 20)**		**Prevalence (cut score 20)**		**Prevalence (cut score 20)**	
PGA	0.087	0.072	0.002	0.002	-0.001	-0.002	-0.002	-0.002
	(0.047)	(0.046)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
PGA^2	0.006***	0.006***	0.000	0.000	0.000*	0.000*	0.000*	0.000*
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Destruction + Damage	67.969***	69.297***	1.191***	1.230***	1.019***	1.060***	0.854***	0.887***
	(4.891)	(4.862)	(0.263)	(0.262)	(0.201)	(0.203)	(0.179)	(0.180)
$(Destruction + Damage)^2$	-64.538***	-65.658***	-1.082*	-1.123*	-0.944*	-0.989**	-0.858*	-0.894*
	(10.139)	(10.134)	(0.482)	(0.477)	(0.377)	(0.374)	(0.369)	(0.365)
Poverty	47.089***	48.070***	0.775**	0.797**	0.555*	0.578**	0.413*	0.432**
	(4.383)	(4.378)	(0.260)	(0.264)	(0.216)	(0.222)	(0.161)	(0.164)
$Poverty^2$	-86.325***	-87.580***	-1.269*	-1.290*	-0.853	-0.877	-0.733*	-0.751*
	(9.870)	(9.877)	(0.643)	(0.653)	(0.554)	(0.569)	(0.365)	(0.375)
Death Ratio	9.659		0.287		0.311		0.249	
	(5.003)		(0.200)		(0.208)		(0.173)	
Constant	0.947*	0.872*	0.010	0.009	-0.005	-0.006	-0.011	-0.012
	(0.390)	(0.390)	(0.024)	(0.024)	(0.018)	(0.018)	(0.013)	(0.014)

* p<0.05, ** p<0.01, *** p<0.001
Standard errors between parentheses

Table 6. Quantile regression results, Quantiles 60, 70, 80 and 90

	60th percentile		**70th percentile**		**80th percentile**		**90th percentile**	
PGA	-0.030***	-0.024***	-0.011	-0.025***	0.066***	0.065***	0.077***	0.056***
	(0.005)	(0.004)	(0.006)	(0.006)	(0.007)	(0.007)	(0.014)	(0.014)
PGA²	0.006***	0.005***	0.007***	0.008***	0.008***	0.008***	0.013***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Destruction + Damage	98.234***	99.495***	117.479***	118.049***	142.546***	143.254***	191.285***	191.337***
	(0.450)	(0.412)	(0.567)	(0.592)	(0.630)	(0.621)	(1.328)	(1.323)
(Destruction + Damage)²	-93.827***	-92.791***	-121.463***	-118.386***	-139.744***	-141.871***	-230.830***	-231.021***
	(0.958)	(0.886)	(1.207)	(1.271)	(1.343)	(1.334)	(2.828)	(2.841)
Poverty	47.563***	48.137***	87.265***	87.279***	142.496***	142.263***	164.752***	166.804***
	(0.449)	(0.414)	(0.565)	(0.594)	(0.629)	(0.624)	(1.325)	(1.329)
Poverty²	-67.704***	-69.806***	-161.844***	-161.529***	-265.332***	-264.664***	-310.499***	-314.720***
	(1.222)	(1.128)	(1.539)	(1.618)	(1.713)	(1.699)	(3.606)	(3.618)
Death Ratio	22.246***				28.365***		9.558***	
	(0.587)		(0.739)		(0.822)		(1.732)	
Constant	-2.372***	-2.423***	-1.703***	-1.694***	-1.942***	-1.922***	6.417***	6.299***
	(0.047)	(0.043)	(0.059)	(0.062)	(0.066)	(0.065)	(0.139)	(0.139)

* p<0.05, ** p<0.01, *** p<0.001
Standard errors between parentheses