

Original Contribution

Uncovering Heterogeneous Associations Between Disaster-Related Trauma and Subsequent Functional Limitations: A Machine-Learning Approach

Koichiro Shiba*, Adel Daoud, Hiroyuki Hikichi, Aki Yazawa, Jun Aida, Katsunori Kondo, and Ichiro Kawachi

* Correspondence to Dr. Koichiro Shiba, Department of Epidemiology, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118 (e-mail shiba_k@g.harvard.edu).

Initially submitted December 3, 2021; accepted for publication October 13, 2022.

This study examined heterogeneity in the association between disaster-related home loss and functional limitations of older adults, and identified characteristics of vulnerable subpopulations. Data were from a prospective cohort study of Japanese older survivors of the 2011 Japan Earthquake. Complete home loss was objectively assessed. Outcomes in 2013 ($n = 3,350$) and 2016 ($n = 2,664$) included certified physical disability levels, self-reported activities of daily living, and instrumental activities of daily living. We estimated population average associations between home loss and functional limitations via targeted maximum likelihood estimation with SuperLearning and its heterogeneity via the generalized random forest algorithm. We adjusted for 55 characteristics of survivors from the baseline survey conducted 7 months before the disaster. While home loss was consistently associated with increased functional limitations on average, there was evidence of effect heterogeneity for all outcomes. Comparing the most and least vulnerable groups, the most vulnerable group tended to be older, not married, living alone, and not working, with preexisting health problems before the disaster. Individuals who were less educated but had higher income also appeared vulnerable for some outcomes. Our inductive approach for effect heterogeneity using machine learning algorithm uncovered large and complex heterogeneity in postdisaster functional limitations among Japanese older survivors.

causal inference; effect heterogeneity; functional limitation; instrumental activities of daily living; machine learning; natural disaster

Abbreviations: ADL, activities of daily living; ATE, average treatment effect; BLP, best linear predictor; CATE, conditional average treatment effect; GRF, generalized random forest; IADL, instrumental activities of daily living; JAGES, Japan Gerontological Evaluation Study.

Older populations are disproportionately affected both during and after disasters (1–3). Studies have documented adverse impacts of disaster-related traumatic experiences (e.g., property damage) on older survivors' mental and cognitive health (4–7). Emerging evidence suggests that older adults may also be susceptible to postdisaster functional limitations (8, 9). Key mechanisms that aggravate functional limitations include the onset of postdisaster psychopathology (e.g., posttraumatic stress disorder and depression) (10) and displacement to unfamiliar surroundings (e.g., temporary accommodations) with the attendant disruption of social communities (11, 12). These processes may contribute to social isolation and homebound status, which in turn lead to functional decline among older adults (13, 14).

At the same time, the health impacts of disaster-related traumatic experiences are likely heterogeneous. For example, only a fraction of individuals develop mental health problems after traumatic experiences, while the majority stays “resilient” and maintains their mental health or quickly returns to normal functioning (15). Individuals from higher socioeconomic backgrounds or stronger networks of social support appear to be resilient against postdisaster psychopathology; thus, these individuals may also be resilient against functional decline following disasters (16, 17). Moreover, living in a community with strong social ties (social capital) mitigates the impact of postdisaster social isolation on functional decline (18, 19). Investigating the sources of this effect heterogeneity may provide insights into

the mechanisms underlying the disaster-health relationships and better guide resource allocation throughout the different stages of disaster mitigation, preparedness, response, and recovery (20). Moreover, evidence on effect heterogeneity will provide insights on the implications of a disaster for health disparities since disasters are likely to have greater impact among socially disadvantaged individuals (21). However, to our knowledge, no study has rigorously examined heterogeneity in functional health outcomes among disaster survivors.

In the present study, we applied a machine-learning algorithm—the generalized random forest (GRF) (22, 23)—to flexibly and inductively assess heterogeneity in the association between disaster-related trauma and functional limitations in a cohort of older survivors of the 2011 Great East Japan Earthquake and Tsunami.

METHODS

Data

We used panel data from the Iwanuma Study, which is part of a nationwide cohort study of Japanese older adults, called the Japan Gerontological Evaluation Study (JAGES) (24). Iwanuma city was one of the field sites of the JAGES located in Miyagi Prefecture (population 44,187 in 2010), approximately 80 km (128 miles) from the epicenter of the 2011 Great East Japan Earthquake. Importantly, the baseline survey of the Iwanuma Study was conducted in August 2010, 7 months before the disaster onset, which enabled us to collect rich information on characteristics of disaster survivors predating the disaster itself. JAGES conducted a census of all residents aged ≥ 65 years in Iwanuma city ($n = 8,576$) and obtained valid responses from 4,957 residents (response rate = 57.8%).

The Great East Japan Earthquake (Richter scale: 9.0) occurred on March 11, 2011. The earthquake and the subsequent tsunami caused devastating damage to the city of Iwanuma. In Iwanuma, the tsunami killed 180 residents, damaged 5,542 houses, and inundated 48% of the land area (Web Figure 1, available at <https://doi.org/10.1093/aje/kwac187>) (25).

Baseline respondents who survived the disaster were surveyed twice over the subsequent 5.5 years. The first follow-up survey was conducted in October 2013, approximately 2.5 years after the disaster. Of the eligible survivors who were healthy enough to participate and still lived in Iwanuma at the time of follow-up ($n = 4,380$), we obtained valid responses from 3,567 subjects. In November 2016, approximately 5.5 years after the disaster, JAGES conducted the second follow-up survey (valid responses: $n = 2,781$). We excluded those who had physical or cognitive disability at baseline, resulting in final analytical samples of $n = 3,350$ for the outcomes in 2013 and $n = 2,664$ for the outcomes in 2016. Figure 1 summarizes the flow of study participant selection.

Measurement

Outcome. We measured functional limitations in 2013 and 2016 using 3 different indicators: certified level of physical

disability, activities of daily living (ADL), and instrumental activities of daily living (IADL).

Level of physical disability was assessed by a standardized in-home assessment under the Japanese Long-Term Care Insurance (LTCI) scheme established in 2000 (26). All study participants were registered for the national LTCI. Every applicant who requested long-term care services was assessed for eligibility by a trained investigator. Following the assessment, the applicant's level of physical disability was classified into one of 9 levels with increasing severity. To facilitate comparison of results across outcomes, we reverse-coded the disability levels so that lower values indicate more severe levels of physical disability (8: "Can go outside using public transportations, etc." to 1: "Cannot turn over in a bed independently"; see Web Table 1 for the detailed outcome definitions). Those who were assessed as having no disability and those who did not request the care services (i.e., people not needing long-term care services) received a score of 9. We obtained physical disability information in 2013 and 2016 from their initial assessment or subsequent annual reassessments.

ADL—degree to which one is able to accomplish basic daily activities—was assessed by asking: "Do you do things like walking, bathing, and toileting by yourself?" and participants chose their response from the following 3 options: 1 = support needed completely, 2 = support needed partially, and 3 = no help needed.

IADL, which evaluates higher-level functional capacities (e.g., going out alone using public transportation), was assessed using the 13 items from the Tokyo Metropolitan Institute of Gerontology Index of Competence (27). Total IADL Scores ranged from 0 to 13 points. Smaller scale scores indicate greater functional limitation.

Exposure. In the 2013 wave, participants retrospectively reported degree of property damage by the tsunami. Property damage in Iwanuma was independently assessed by property inspectors and classified into 5 levels: no damage, partial, minor, major, and complete destruction (28). Criteria for each level of property damage are summarized in Web Table 2. We created a binary variable representing complete home loss (1 = "complete destruction" and 0 = "no damage/less severe damage") because home loss has been shown to uniquely predict health deterioration after the disaster (29, 30).

Covariates. We selected 55 pre-disaster variables from the baseline (2010) survey wave, including 4 demographic characteristics, 3 measures of socioeconomic status, 28 health conditions, 14 psychosocial factors, and 6 behavioral factors (see Web Table 3 for the complete list of the selected variables). We selected these factors because they were likely to operate as confounders, effect modifiers, or both.

Statistical analysis

First, we estimated the population average treatment effects (ATEs) of home loss on functional limitations in 2013 and 2016. ATEs quantify the difference in the mean level of functional limitation had everyone in the

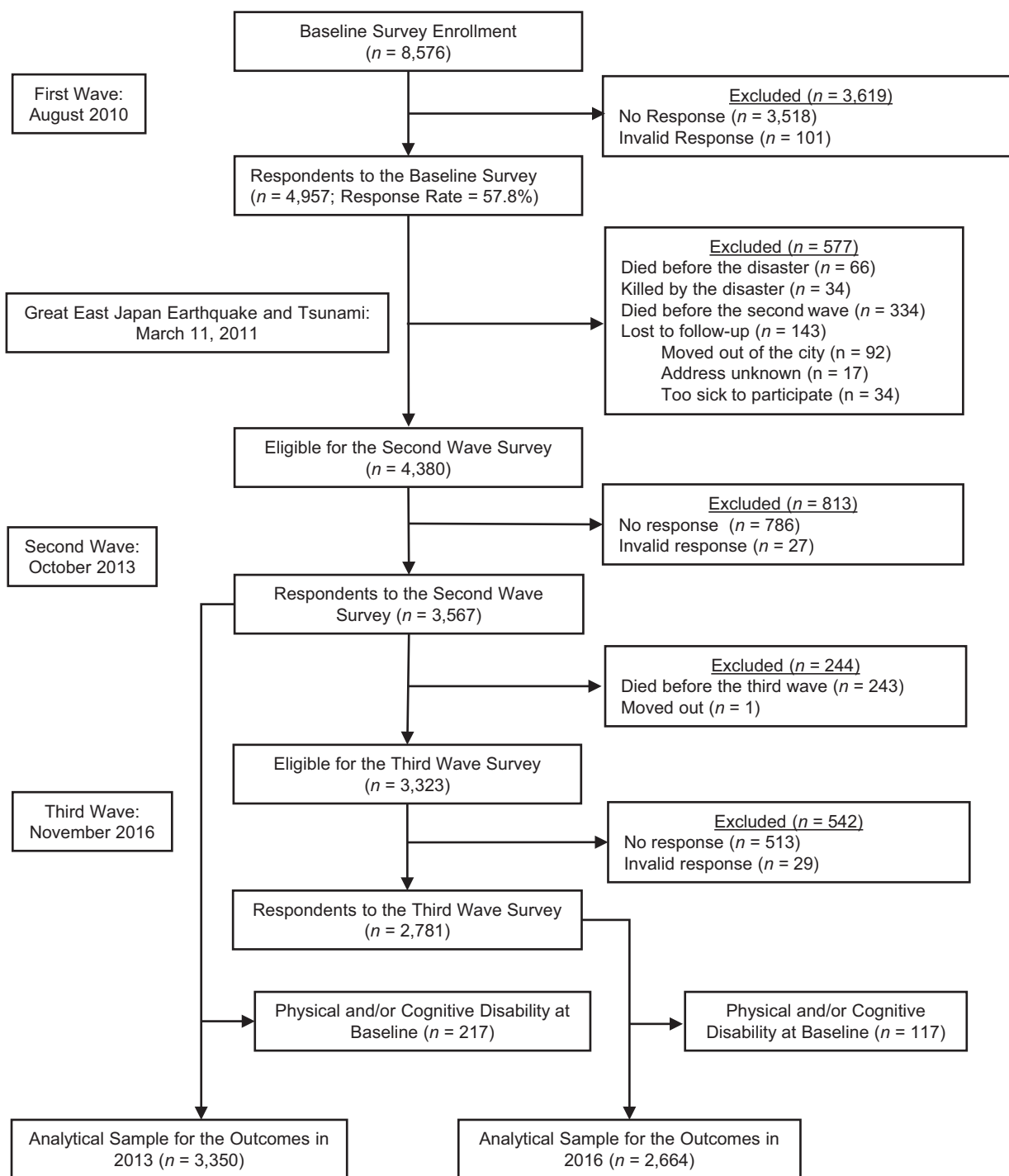


Figure 1. Sample flow chart, Iwanuma, Miyagi, Japan, 2010–2016.

population been exposed to home loss vs. nobody being exposed, $E[Y_{a=1} - Y_{a=0}]$. We estimated ATEs using the doubly robust targeted maximum likelihood estimation (TMLE). This approach estimates both the exposure model and outcome model and yields unbiased estimates for the ATEs if either of the models is consistently estimated. Both

exposure and outcome models were fitted data-adaptively via the SuperLearner, an ensemble method that uses weighted combinations of multiple candidate estimators, which included generalized linear models, gradient-boosting machine, and neural net (31–33). Targeted maximum likelihood estimation and Super Learning were performed

using the `ltmle` and `SuperLearner` R (R Foundation for Statistical Computing, Vienna, Austria) packages (34, 35).

Second, to examine heterogeneous effects, we estimated conditional average treatment effects (CATEs) of home loss on functional limitations. Formally, CATE is the effect of an exposure conditional on the values of a set of covariates ($\mathbf{L} = l$);

$$E[Y_{a=1} - Y_{a=0} | \mathbf{L} = l]$$

where Y_a is the potential outcome Y under the binary treatment $A = a$.

We estimated CATEs and identified potential sources of effect heterogeneity by using the causal forest approach from the GRF algorithm (22). Causal forest extends the random forest algorithm, a common nonparametric algorithm designed to predict conditional expectations (36). Random forest grows many regression trees, which is a type of decision tree, by partitioning bootstrapped subsamples based on the values of random subsets of covariates and computes a weighted average of outcome in each leaf (i.e., subsamples defined by the same combinations of covariate values) of a tree. Causal forest targets and assesses the contrast in the average outcome between the exposed versus unexposed individuals in each leaf (i.e., CATEs), rather than predicting the average outcome itself. Although each tree was grown using a random subset of covariates, the final CATE estimates were calculated as weighted averages of predictions from all trees and can be interpreted as being conditional on all the 55 covariates. In growing trees, causal forest randomly splits a subsample in half and uses only the first half to construct partitioning and the other half to fill in CATE estimates in each leaf—this process is often called “honesty” and reduces bias in tree predictions. Partitioning was determined such that it maximizes heterogeneity in treatment effect estimates across leaves.

We implemented GRF using the R (R Foundation for Statistical Computing) package `grf` (37). We used cross-fitting with 20 folds so that predictions for each fold were made based on trees trained without using observations from the fold. Before growing 2,000 regression trees to construct a causal forest for each outcome, we used out-of-bag samples and tuned the following parameters via cross-validation: fraction of the data used to build each tree, number of variables tried for each split, a target for the minimum number of observations in each tree leaf, the fraction of data to be used for determining splits, whether the estimation sample tree should be pruned such that no leaves are empty, maximum imbalance of a split, and penalty for imbalanced splits. We provide R code in Web Appendix 1 to facilitate reproducing our analysis.

To evaluate the model performance and formally test for the presence of effect heterogeneity, we conducted the following analyses. First, we performed “best linear predictor” (BLP) analysis (38). The BLP analysis fits the following model:

$$Y_i - \hat{Y}_i = \alpha(A_i - p(L_i)) * \bar{\tau} + \beta(A_i - p(L_i)) (\hat{\tau}_i - \bar{\tau}) + \varepsilon,$$

where $\hat{\tau}_i$ is the predicted CATE for individual i , $\bar{\tau}$ is the mean of the CATE estimates, \hat{Y}_i is the predicted outcome,

and $p(L) = \Pr[A = 1 | L]$ is the probability of home loss conditional on the covariates. If the coefficient α (the mean forest) is close to 1, then it indicates that the average forest prediction is well calibrated. If the coefficient β (the differential forest) is close to 1, then it indicates the forest prediction adequately captures the underlying heterogeneity. The single-sided test of $\beta > 0$ is used as an omnibus test for effect heterogeneity. Second, we ranked each individual in each fold based on their CATE, grouped these ranks into quintiles (because the ranks were specific to each of the 20 folds, there were 20 ties for each rank), and re-estimated the ATE via targeted maximum likelihood estimation, stratifying by the quintile of the CATE. We expect the ATE estimate to monotonically increase across quintiles defined by the CATE ranks. Third, as an alternative way of testing presence of effect heterogeneity, we followed the recommendation by Athey and Wager (2019) (23). Specifically, we grouped the analytical sample into 2 groups according to whether their estimated CATEs were above or below the median CATE estimate. We then estimated group-specific ATE in each of the 2 groups and compared the 2 ATE estimates. The difference in the 2 ATE estimates was tested via bootstrapping.

Last, we investigated sources of effect heterogeneity. To identify the most salient sources of effect heterogeneity, we compared characteristics of the top 10% and the bottom 10% of the CATE distributions. Because lower values for our continuous outcomes suggest greater functional limitations, negative CATE values indicate functional impairment following home loss. The top 10% (i.e., those who were predicted to experience the smallest functional impairment following home loss) was labeled as the “resilient” group, while the bottom 10% (i.e., those who were predicted to experience the largest functional impairment following home loss) was labeled as the “vulnerable” group. Moreover, we used the “variable importance” feature of GRF and, for each outcome, identified the top 3 pre-disaster characteristics that were most often used in growing the trees. We then created heatmaps showing how estimated CATEs were distributed according to the levels of the 3 pre-disaster characteristics.

As a sensitivity analysis, we examined whether results were robust to using different cutoffs to create binary housing damage variables. We performed imputation of missing data using random forest via the R (R Foundation for Statistical Computing) package “`missforest`” (39) (more details are available in Web Appendix 2). All analyses were performed using R, version 3.6.0.

RESULTS

Table 1 summarizes the pre-disaster demographic characteristics of the analytical sample ($n = 3,350$ for the 2013 outcomes and $n = 2,664$ for the 2016 outcomes) according to home loss status. Individuals who experienced home loss (vs. no home loss) tended to be from lower socioeconomic status backgrounds (fewer years of schooling and lower household income) and to report depressive symptoms before the disaster (e.g., 37% vs. 29% among those without home loss exposure for the 2013 sample). In both 2013 and 2016, those who experienced home loss (vs. no home loss)

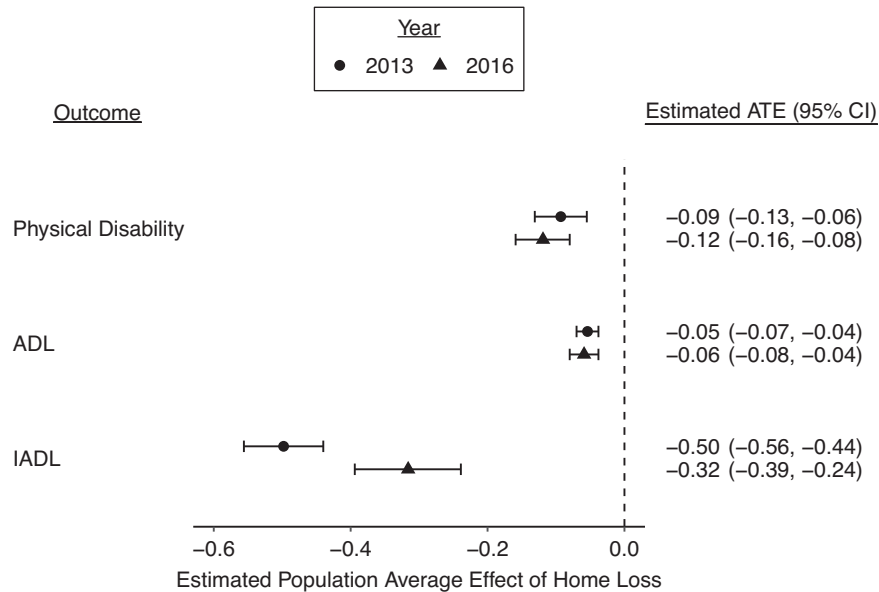


Figure 2. Estimated population average treatment effect (ATE) of home loss on functional limitation indicators in 2013 (2.5 years after the disaster) and 2016 (5.5 years after), Iwanuma, Miyagi, Japan, 2010–2016. Population average effects (i.e., ATEs and confidence intervals (CIs)) of the exposures were estimated via the doubly robust targeted maximum likelihood estimation. Models were estimated data-adaptively via the SuperLearner using generalized linear models, gradient-boosting machine, and neural net as candidate estimators. All models adjusted for the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Instrumental activities of daily living (IADL) status was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Instrumental activities of daily living scores ranged from 0 to 13 points, where smaller scores indicate lower functional independence. Activities of daily living (ADL) had 3 levels (1 = support needed completely, 2 = support needed partially, and 3 = no help needed). Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, decrease in these outcomes indicate increased functional limitation.

had greater functional limitations across all indicators (Web Table 4).

Figure 2 depicts the estimated ATEs of home loss on the functional limitation outcomes. In 2013, home loss was consistently associated with increased functional limitations across indicators (e.g., estimate for certified physical disability = -0.09 ; 95% confidence interval: $-0.13, -0.06$). The population-average associations between home loss and increased functional limitations persisted when we examined the outcomes in 2016. Sensitivity analysis using different cutoffs for property damage (Web Figure 2) showed robust associations.

The distributions of the CATEs estimated via GRF are shown in Figure 3. Web Table 5 shows the summary statistics for the estimated CATEs: The CATE estimates were below zero for most individuals (e.g., CATEs for certified physical disability level in 2013: minimum = -1.95 , maximum = 0.06) and showed consistent trends with the ATE estimates (i.e., home loss was associated with functional impairment). However, the variation in the estimated CATEs indicates effect heterogeneity (e.g., standard deviation = 0.27 for certified physical disability level in 2013). As shown in Web Table 6, there was evidence of effect heterogeneity for all functional limitation outcomes across the 2 approaches for testing the presence of heterogeneity although the evidence was more modest for

ADL in 2016 when using the omnibus test of heterogeneity based on the BLP analysis ($P = 0.10$). The estimated coefficients from the BLP analyses ranged from 1.04 to 1.28 for the mean forest prediction and 0.91 to 2.01 for the differential forest prediction and tended to be larger (away from 1) for the outcomes assessed in 2016 (Web Table 6). In the calibration plot (Web Figure 3), we observed monotonic increase in the rank-specific ATE estimate with increasing CATE ranking. When assessing outcomes in 2016, the calibration plot showed some deviations from the perfect monotonically increasing trend (e.g., for IADL in 2016, the ATE estimate for the CATE quintile 3 was greater than the ATE estimates for quintiles 4 and 5). Sensitivity analysis using different cutoffs for property damage similarly indicated effect heterogeneity (Web Figure 4).

Table 2 compares the key pre-disaster characteristics of the resilient group (top decile of the CATE distribution) with those of the vulnerable group (bottom decile) for the certified physical disability outcome in 2013 and 2016. Comparisons of all 55 covariates are available in Web Table 7. We identified common pre-disaster characteristics of the vulnerable group (i.e., individuals for whom home loss was more strongly associated with greater levels of physical disability). When assessing the outcome in 2013, the people in the vulnerable group (vs. the resilient group) were more

Table 1. Baseline Sociodemographic Characteristics of Analytical Samples in 2010 According to Home Loss Status, Iwanuma, Miyagi, Japan, 2010–2016^a

Baseline Characteristic	Sample for the 2013 Outcomes (n = 3,350)						Sample for the 2016 Outcomes (n = 2,664)					
	Home Loss (n = 148)			No Home Loss (n = 3,112)			Home Loss (n = 107)			No Home Loss (n = 2,496)		
	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)
Age, years			73.6 (6.4)			73.2 (6.0)			71.9 (5.2)			72.5 (5.5)
Gender												
Men	87	59		1,709	55		58	54		1,383	55	
Women	61	41		1,403	45		49	46		1,113	45	
Marital status												
Married	98	74		2,216	73		77	79		1,834	75	
Widowed	30	23		679	22		20	20		494	20	
Divorced	2	1.5		77	2.5		0	0		65	2.7	
Single	0	0		36	1.2		0	0		27	1.1	
Others	2	1.5		15	0.5		1	1.0		13	0.5	
Living alone												
No	134	97		2,772	91		99	96		2,242	92	
Yes	4	2.9		269	8.8		4	3.9		204	8.3	
Education, years												
<6 years	0	0		33	1.1		0	0		21	0.9	
6–9	92	67		969	32		66	65		732	30	
10–12	35	26		1,346	44		28	28		1,116	46	
≥13	8	5.8		660	22		7	6.9		559	23	
Other	2	1.5		23	0.8		0	0		17	0.7	
Job												
Working	23	20		514	19		21	23		433	19	
Retired	66	56		1,782	64		51	57		1,452	65	
Never worked	28	24		479	17		18	20		348	16	

Table continues

Table 1. Continued

Baseline Characteristic	Sample for the 2013 Outcomes (n = 3,350)				Sample for the 2016 Outcomes (n = 2,664)							
	Home Loss (n = 148)		No Home Loss (n = 3,112)		Home Loss (n = 107)		No Home Loss (n = 2,496)					
	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)			
Household income per 10,000 yen ^b			170 (127)			234 (141)			184 (134)			235 (137)
Depressive symptoms ^c												
Mild/severe	45	37		778	29		35	38		604	27	
None	77	63		1,931	71		56	62		1,595	73	
Self-rated health												
Bad	24	17		384	13		19	19		318	13	
Not good	93	65		2,183	71		66	65		1,780	73	
Good	15	11		430	14		12	12		316	13	
Very good	10	7.0		61	2.0		5	4.9		41	1.7	
Body mass index ^d			23.9 (2.80)			23.5 (3.10)			24.2 (2.80)			23.5 (2.99)
Total IADL ^e			11.4 (2.50)			11.9 (1.76)			11.9 (1.59)			12.0 (1.59)
Instrumental IADL ^e			4.55 (1.09)			4.74 (0.73)			4.73 (0.70)			4.78 (0.65)
Intellectual IADL ^e			3.40 (0.97)			3.67 (0.67)			3.51 (0.81)			3.72 (0.61)
Social IADL ^e			3.48 (0.93)			3.47 (0.91)			3.62 (0.71)			3.51 (0.87)
ADL ^f			2.98 (0.14)			2.99 (0.13)			2.99 (0.10)			2.99 (0.10)
No. of treatments for major diseases ^g			1.40 (1.28)			1.46 (1.35)			1.36 (1.25)			1.41 (1.29)

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living; SD, standard deviation.

^a The sample sizes for the subgroups do not add up to the size of the total samples (n = 3,350 for 2013 and n = 2,664 for 2016) due to missing data for home loss status.

^b Annual household income was divided by the square root of the number of household members to account for household size.

^c We used the Geriatric Depression Scale (range: 0–15 points; higher scores indicate more depressive symptoms) to assess depressive symptoms. We used a cutoff of ≥5 points to define moderate/severe depressive symptoms.

^d Weight (kg)/height (m)².

^e IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0 to 13 points for total IADL, 0 to 5 points for instrumental IADL, 0 to 4 points for intellectual IADL, and 0 to 4 points for social IADL, where smaller scores indicate lower functional independence.

^f ADL had 3 levels (1 = support needed completely, 2 = support needed partially, and 3 = no help needed).

^g We calculated counts of current treatment for major diseases, including cancer, heart diseases, stroke, hypertension, diabetes, obesity, hyperlipidemia, osteoporosis, arthritis, fracture, respiratory diseases, gastrointestinal diseases, liver diseases, psychiatric diseases, dysphagia, visual impairment, hearing loss, dysuria, and insomnia.

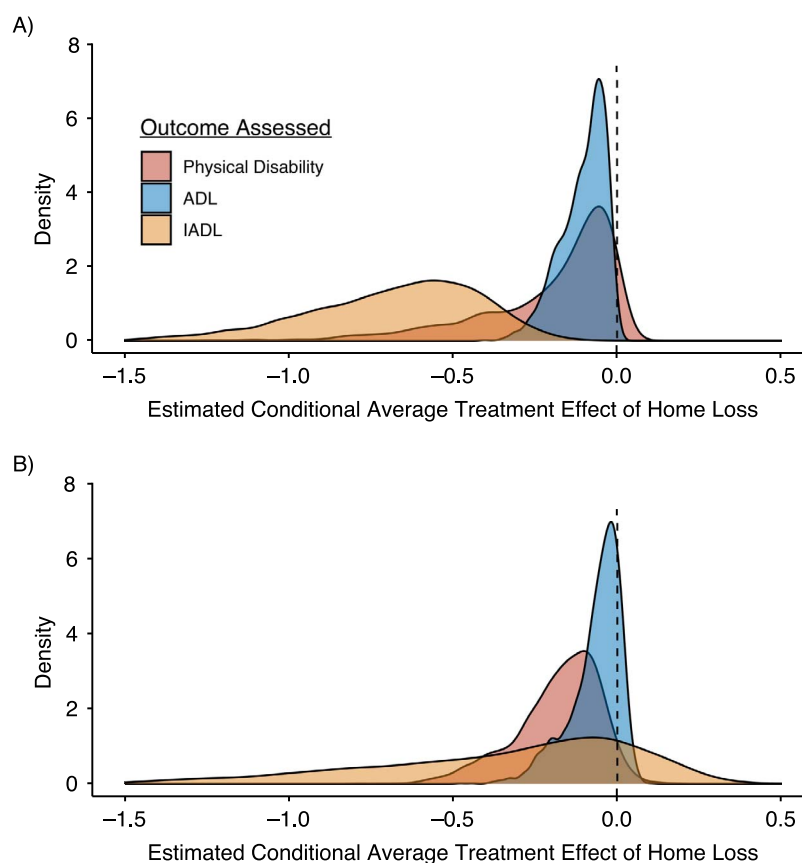


Figure 3. Distributions of estimated conditional average treatment effect (ATE) of home loss on functional limitation indicators 2.5 years after the disaster (year 2013) (A) and 5.5 years after the disaster (year 2016) (B), Iwanuma, Miyagi, Japan, 2010–2016. Heterogeneous effects (i.e., conditional ATEs) were estimated using generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Activity of daily living (ADL) had 3 levels (1 = support needed completely, 2 = support needed partially, and 3 = no help needed). Instrumental activities of daily living (IADL) status was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Instrumental activities of daily living scores ranged from 0 to 13 points, where smaller scores indicate lower functional independence. Levels of certified physical disability ranged from 1 (“Cannot roll over in a bed independently”) to 9 (“no physical disability/not requesting care services”), where smaller values indicate greater levels of disability. Thus, decrease in these outcomes indicate increased functional limitation.

likely to be older (mean age = 83.1 vs. 68.6 in the resilient group), men (67% vs. 36%), not married (56% vs. 2%), living alone (10% vs. 0.9%), less educated (e.g., ≥ 13 years schooling: 8.5% vs. 25%), and not working (95% vs. 71%) before the earthquake. Members of the vulnerable group also tended to have baseline health problems, such as more mild/severe depressive symptoms (38% vs. 18%), poor self-rated health (e.g., “Very good”: 6.8% vs. 19%), lower total IADL scores (mean = 9.60 vs. 12.5), and more major diseases (mean count of treatments for major conditions = 2.07 vs. 1.02). This pattern generally persisted when assessing the physical disability level in 2016 and the other outcome indicators—total IADL scores (Web Table 8) and ADL (Web Table 9)—with some exceptions (e.g., living alone: 8.2% vs. 5.0%, $P = 0.20$ for the physical disability outcome in 2016).

We also found mixed patterns for pre-disaster household income. The vulnerable group and the resilient group had comparable pre-disaster household incomes when assessing physical disability level outcome in 2013 and ADL in 2016.

The vulnerable (vs. resilient) group was characterized by lower pre-disaster income when assessing physical disability level in 2016 (mean (10,000 yen) = 242.7 vs. 284.2; P value < 0.001) and ADL in 2013 (mean = 235.7 vs. 270.7; P value = 0.003; Web Table 9). However, members of the vulnerable group, who were less educated and had baseline health problems (e.g., depressive symptoms) as described above, also had higher pre-disaster income when we assessed IADL in 2013 and 2016 (e.g., mean in 2013 = 254.9 vs. 233.9; P value < 0.01; Web Table 8). We conducted an ad-hoc sensitivity analysis examining the associations between home loss and the outcomes in 2013 among those who participated in both follow-up waves in 2013 and 2016 ($n = 2,664$) and obtained comparable ATE and CATE estimates (Web Table 10).

Web Table 11 shows 3 variables that GRF used most often in splitting trees for each outcome. For instance, 15% of all trees for disability level in 2013 used age in sample splitting. The relationships between each of

Table 2. Pre-Disaster Sociodemographic Characteristics of People at Top 10% vs. Bottom 10% of the Estimated Conditional Average Treatment Effect of Home Loss on Certified Physical Disability Level in 2013 and 2016, Iwanuma, Miyagi, Japan, 2010–2016^a

Baseline Characteristic	2013 (n = 3,350)						2016 (n = 2,664)						
	Resilient (n = 340)			Vulnerable (n = 340)			Resilient (n = 280)			Vulnerable (n = 280)			
	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	P Value ^b
CATE estimates			0.00 (0.02)			-0.69 (0.19)			-0.01 (0.04)			-0.43 (0.08)	<0.001
Age			68.6 (2.69)			83.1 (5.04)			68.6 (3.14)			80.0 (3.41)	<0.001
Gender													<0.001
Men	121	36		229	67		109	39		179	64		
Women	219	64		111	33		171	61		101	36		
Marital status													<0.001
Married	334	98		148	44		244	87		178	64		
Widowed	6	1.8		180	53		27	9.6		93	33		
Divorced	0	0		3	0.9		7	2.5		5	1.8		
Single	0	0		7	2.1		1	0.4		3	1.1		
Others	0	0		2	0.6		1	0.4		1	0.4		
Living alone													0.2
No	337	99		306	90		266	95		257	92		
Yes	3	0.9		34	10		14	5.0		23	8.2		
Education, years													<0.001
<6 years	1	0.3		17	5.0		1	0.4		3	1.1		
6–9	74	22		175	51		60	21		113	40		
10–12	174	51		114	34		148	53		105	38		
≥13	86	25		29	8.5		71	25		55	20		
Others	5	1.5		5	1.5		0	0		4	1.4		
Job													<0.001
Working	100	29		17	5.0		90	32		15	5.4		
Retired	216	64		174	51		173	62		177	63		
Never worked	24	7.1		149	44		17	6.1		88	31		
Household income, per 10,000 yen ^c			235 (142)			232 (125)			284 (130)			243 (108)	<0.001

Table continues

Table 2. Continued

Baseline Characteristic	2013 (n = 3,350)						2016 (n = 2,664)						
	Resilient (n = 340)			Vulnerable (n = 340)			Resilient (n = 280)			Vulnerable (n = 280)			
	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	P Value ^b
Depressive symptoms ^d													<0.001
Mild/severe	61	18		128	38		22	7.9		94	34		
None	279	82		212	62		258	92		186	66		
Self-rated health													<0.001
Very good	63	19		23	6.8		55	20		28	10		
Good	260	76		212	62		215	77		196	70		
Not good	17	5.0		85	25		8	2.9		49	18		
Bad	0	0		20	5.9		2	0.7		7	2.5		
Body mass index ^e			25.6 (2.51)			22.8 (3.15)			24.4 (2.84)			22.7 (2.43)	<0.001
Total IADL ^f			12.5 (0.78)			9.60 (3.32)			12.2 (1.40)			12.3 (1.06)	0.8
Instrumental IADL ^f			4.96 (0.20)			3.82 (1.50)			4.66 (0.78)			4.94 (0.27)	<0.001
Intellectual IADL ^f			3.80 (0.45)			3.02 (1.18)			3.87 (0.46)			3.79 (0.47)	0.001
Social IADL ^f			3.73 (0.53)			2.76 (1.36)			3.65 (0.63)			3.58 (0.82)	0.7
ADL ^g			3.00 (0.05)			2.93 (0.33)			2.99 (0.08)			3.00 (0.06)	0.6
Counts of treatment for major diseases ^h			1.02 (1.02)			2.07 (1.60)			1.16 (1.04)			1.70 (1.37)	<0.001

Abbreviations: ADL, activities of daily living; CATE, conditional average treatment effect; IADL, instrumental activities of daily living; SD, standard deviation.

^a Heterogeneous effects (i.e., CATEs) were estimated via the generalized random forest algorithm, using the 55 pre-disaster demographic and socioeconomic factors, health conditions, psychosocial variables, and behaviors from the 2010 wave. Levels of certified physical disability ranged from 1 ("Cannot roll over in a bed independently") to 9 ("no physical disability/not requesting care services"), where smaller values indicate greater levels of disability. Thus, decrease in this outcome indicate increased functional limitation. Top 10% of the distributions of CATEs were labeled as a "resilient" group because they showed weaker associations between home loss and decreased outcome. Bottom 10% of the distributions of CATEs were labeled as a "vulnerable" group because they showed greater associations between home loss and decreased outcome.

^b P values for between-group differences. We used Wilcoxon rank sum test for continuous variables and Fisher's exact test for categorical variables.

^c Annual household income was divided by the square root of the number of household members to account for household size.

^d We used the Geriatric Depression Scale (range: 0–15 points; higher scores indicate more depressive symptoms). We used a cutoff of ≥5 points to define moderate/severe depressive symptoms.

^e Weight (kg)/height (m)².

^f IADL was measured by the 13-item Tokyo Metropolitan Institute of Gerontology Index of Competence. Scores ranged from 0 to 13 points for total IADL, 0 to 5 points for instrumental IADL, 0 to 4 points for intellectual IADL, and 0 to 4 points for social IADL, where smaller scores indicate lower functional independence.

^g ADL had 3 levels (1 = support needed completely, 2 = support needed partially, and 3 = no help needed).

^h We calculated counts of current treatment for major diseases, including cancer, heart diseases, stroke, hypertension, diabetes, obesity, hyperlipidemia, osteoporosis, arthritis, fracture, respiratory diseases, gastrointestinal diseases, liver diseases, psychiatric diseases, dysphagia, visual impairment, hearing loss, dysuria, and insomnia.

these variables and estimated CATEs are shown in Web Figure 5 (for outcomes in 2013) and Web Figure 6 (for outcomes in 2016). For example, CATEs for disability level in 2013 showed monotonic decrease with increasing age (Spearman's correlation = -0.71 as shown in Web Table 11). Distributions of estimated CATEs across combinations of these 3 variables are shown in Web Figure 7 (for disability level), Web Figure 8 (for IADL), and Web Figure 9 (for ADL). These heatmaps identified some complex high-dimensional heterogeneity. For example, when assessing physical disability level in 2016, higher pre-disaster income was a source of resilience (i.e., CATEs closer to 0) among individuals who were 65–68 years old; on the other hand, higher income was a source of vulnerability among people >76 years old.

DISCUSSION

In this prospective study of older survivors from the 2011 Great East Japan Earthquake, we estimated the heterogeneous impacts of disaster-induced home loss on subsequent functional limitations. Our main findings are 3-fold. First, after adjusting for a set of 55 pre-disaster characteristics, there was strong evidence of population average effects of home loss on increased functional limitations across all indicators 2.5 years and 5.5 years after the disaster. Second, we identified evidence of heterogeneity in the associations between home loss and functional limitation as demonstrated by the spread in the distributions of the estimated CATEs as well as the results of the 2 tests for effect heterogeneity that we implemented; however, the results concerning effect heterogeneity need to be interpreted with caution for the outcomes in 2016 because causal forest fit for these outcomes appeared to be less than perfect. Third, we inductively identified patterns in pre-disaster characteristics of subgroups particularly vulnerable to functional impairment following home loss. Vulnerable individuals tended to have lower educational attainment and pre-disaster health problems (e.g., depressive symptoms), but these individuals also had higher pre-disaster income.

Our finding for the population average effects of home loss on increased functional limitations is consistent with what has been reported previously (8, 9). The link between home loss and functional limitations might be explained by postdisaster residential displacement. Hikichi et al. (12) reported decreased IADL among people who experienced individual relocation (i.e., individually relocating to public temporary housing via lottery, moving to rental housing on the open market, or building new homes) while no association with IADL was observed for relocating to public temporary housing together with other members of a community. Hence, home loss and subsequent individual relocation may have caused disruption of pre-existing social relationships and social isolation, ultimately resulting in functional limitations of older survivors (12, 40).

Our study also identified subgroups for whom home loss was associated with substantially greater functional impairment compared with the population average. Most studies assessing effect heterogeneity rely on a deductive

approach in which researchers select a limited set of predictors—typically just one or a few—a priori as sources of effect heterogeneity and statistically test interactions, one variable at a time. The inductive approach for assessing effect heterogeneity using GRF is advantageous because it does not require investigators to specify effect modifiers a priori but rather finds them from numerous candidates in a data-driven way. Vulnerable individuals tended to be older, not married, living alone, less educated, and not working and to have baseline health problems. These pre-disaster factors were mostly consistent with what prior trauma research has identified as sources of vulnerability (16, 17, 41). These findings support our hypothesis that heterogeneity in functional limitations following disasters arises from individual variations in the likelihood of experiencing postdisaster mental health problems and social isolation and homebound status. For instance, pre-disaster depression, a risk factor for posttraumatic stress disorder (16), was more prevalent among the vulnerable group in this study (e.g., 38% vs. 18% in the resilient group for physical disability level in 2013). Factors such as old age, nonmarried status, and living alone might have accelerated social isolation among individuals who experienced residential relocation into temporary housing (42).

Our inductive approach for assessing effect heterogeneity provided potentially new insights that could have been missed with a deductive approach. In prior work examining effect heterogeneity deductively, higher socioeconomic status such as higher income alone has been linked to disaster resilience (4). In contrast, our inductive approach captured complex heterogeneity resulting from interactions between multiple characteristics. We demonstrated that the most vulnerable individuals for the IADL outcomes tended to have lower educational attainment and more health problems—each of these being established risk factors for vulnerability—but also higher household income before the disaster. Some prior evidence suggests that individuals with status inconsistency (e.g., discordance between educational attainment versus earned income) are more likely to engage in unhealthy behaviors, such as excess drinking, which may contribute to morbidity and functional limitations (43, 44); hence, higher income when coupled with low education may produce greater vulnerability. Alternatively, it is possible that individuals with high income were more likely to purchase new homes privately (i.e., individual relocation) and suffered from the loss of pre-existing social capital (12). However, the mechanisms for the complex heterogeneity remain unclear, and future research is warranted.

Seven limitations should be noted. First, our ATE and CATE estimates are based on the assumption that the 55 observed covariates sufficed to adjust for confounding of the exposure-outcome associations, which may not hold. Moreover, the covariates do not necessarily suffice to adjust for confounding between each pre-disaster characteristic and the outcomes; hence, the results do not tell us whether intervening in those pre-disaster characteristics can change the impact of home loss (20). Second, our binary assessment of home loss was relatively crude and does not capture nuanced differences in value of the damaged properties and duration of occupancy. Third, selective attrition due to

loss-to-follow-up may have induced selection bias (for more discussion on the potential selection bias, see Web Appendix 3, Web Figure 10, and Web Table 12) (45). Fourth, while physical disability was objectively assessed in this study, we relied on self-reporting for the other outcomes (i.e., IADL and ADL), which may be biased. Fifth, our study design does not allow us to distinguish between persistence of disability that emerged soon after the disaster and delayed onset of disability that emerged during the interval between 2.5 to 5.5 years after the disaster. Sixth, the results concerning effect heterogeneity need to be interpreted with caution, particularly for the outcomes in 2016, because the BLP analysis and the calibration plot indicating causal forest fit for these outcomes may be less than perfect. Seventh, some of the observed patterns from our inductive approach (e.g., complex heterogeneity involving income) are inconclusive and may be due to chance. Future studies should test for a specific hypothesis with a deductive approach.

In conclusion, our study provided evidence for considerable heterogeneity in the adverse impacts of disaster-induced home loss on subsequent functional limitations among older survivors. We identified subpopulations at exceptionally high risk of postdisaster functional impairment and sources of the complex heterogeneity, which would have likely been overlooked had we estimated only the population average effects or used the conventional deductive approach for assessing effect heterogeneity. Assessing such effect heterogeneity can contribute to more targeted postdisaster public health interventions to preserve survivors' functional health status.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, United States (Koichiro Shiba); Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, United States (Adel Daoud, Aki Yazawa, Ichiro Kawachi); Institute for Analytical Sociology, Linköping University, Linköping, Sweden (Adel Daoud); The Division of Data Science and Artificial Intelligence, the Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, Sweden (Adel Daoud); Division of Public Health, Kitasato University School of Medicine, Kanagawa, Japan (Hiroyuki Hikichi); Department of Epidemiology and Prevention, Center for Clinical Sciences, National Center for Global Health and Medicine, Tokyo, Japan (Aki Yazawa); Department of Oral Health Promotion, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan (Jun Aida); Center for Preventive Medical Sciences, Chiba University, Chiba, Japan (Katsunori Kondo); and Center for Gerontology and Social Science, National Center for Geriatrics and Gerontology, Aichi, Japan (Katsunori Kondo).

This work was supported by the National Institutes of Health (grant R01 AG042463 to I.K.); Grants-in-Aid for Scientific Research from the Japan Society for the

Promotion of Science (grants KAKENHI 23243070, KAKENHI 22390400, KAKENHI 20H00557, and KAKENHI 24390469 to K.K. and KAKENHI JP21J01171 to A.Y.); Health Labour Sciences Research Grant from the Japanese Ministry of Health, Labour, and Welfare (grant H24-Choju-Wakate-009 to K.K.); the Strategic Research Foundation Grant-Aided Project for Private Universities from the Japanese Ministry of Education, Culture, Sports, Science, and Technology (grant S0991035 to K.K.); and Japan Agency for Medical Research and Development (grants JP17dk0110017, JP18dk0110027, JP18ls0110002, JP18le0110009, JP19dk0110034, JP19dk0110037, JP20dk0110034). A.Y. is financially supported by JSPS Research Fellowship for Young Scientists.

The JAGES data used in this study will be made available upon request, as per National Institutes of Health data access policies. Confidentiality concerns prevent us from depositing our data in a public repository. Authors requesting access to the Iwanuma data need to contact the principal investigator of the parent cohort (K.K.: kkondo@chiba-u.jp) and the Iwanuma substudy principal investigator (I.K.: ikawachi@hsph.harvard.edu) in writing. Proposals for using the Iwanuma data will be discussed among researchers in the project team to ensure that the necessary confidentiality of the data is respected with ongoing analyses.

Presented at the 2022 Society for Epidemiologic Research annual meeting, June 14–17, 2022, Chicago, Illinois.

The views expressed in this article are those of the authors and do not reflect those of the funders.

Conflict of interest: none declared.

REFERENCES

- Hutton D. *Older People in Emergencies: Considerations for Action and Policy Development*. Geneva, Switzerland: World Health Organization; 2008.
- Fernandez LS, Byard D, Lin C-C, et al. Frail elderly as disaster victims: emergency management strategies. *Prehosp Disaster Med*. 2002;17(2):67–74.
- Ngo EB. When disasters and age collide: reviewing vulnerability of the elderly. *Nat Hazards Rev*. 2001;2(2):80–89.
- Goldmann E, Galea S. Mental health consequences of disasters. *Annu Rev Public Health*. 2014;35(1):169–183.
- Hikichi H, Aida J, Kondo K, et al. Increased risk of dementia in the aftermath of the 2011 Great East Japan Earthquake and Tsunami. *Proc Natl Acad Sci U S A*. 2016;113(45):E6911–E6918.
- Hikichi H, Aida J, Kondo K, et al. Persistent impact of housing loss on cognitive decline after the 2011 Great East Japan earthquake and tsunami: evidence from a 6-year longitudinal study. *Alzheimers Dement*. 2019;15(8):1009–1018.
- Cherry KE, Su LJ, Welsh DA, et al. Cognitive and psychosocial consequences of Hurricanes Katrina and Rita among middle-aged, older, and oldest-old adults in the Louisiana Healthy Aging Study (LHAS). *J Appl Soc Psychol*. 2010;40(10):2463–2487.

8. Tsuboya T, Aida J, Hikichi H, et al. Predictors of decline in IADL functioning among older survivors following the great East Japan earthquake: a prospective study. *Soc Sci Med*. 2017;176:34–41.
9. Pruchno R, Wilson-Genderson M, Heid AR, et al. Type of disaster exposure affects functional limitations of older people 6 years later. *J Gerontol A Biol Sci Med Sci*. 2020;75(11):2139–2146.
10. Clouston SAP, Guralnik JM, Kotov R, et al. Functional limitations among responders to the World Trade Center attacks 14 years after the disaster: implications of chronic posttraumatic stress disorder. *J Trauma Stress*. 2017;30(5):443–452.
11. Aldrich DP. *Building Resilience: Social Capital in Post-Disaster Recovery*. Chicago, IL: University of Chicago Press; 2012.
12. Hikichi H, Sawada Y, Tsuboya T, et al. Residential relocation and change in social capital: a natural experiment from the 2011 Great East Japan Earthquake and Tsunami. *Sci Adv*. 2017;3(7):e1700426.
13. Perissinotto CM, Stijacic Cenzer I, Covinsky KE. Loneliness in older persons: a predictor of functional decline and death. *Arch Intern Med*. 2012;172(14):1078–1083.
14. Cohen-Mansfield J, Shmotkin D, Hazan H. The effect of homebound status on older persons. *J Am Geriatr Soc*. 2010;58(12):2358–2362.
15. Bonanno GA, Mancini AD. Beyond resilience and PTSD: mapping the heterogeneity of responses to potential trauma. *Psychol. Dent Traumatol*. 2012;4(1):74.
16. Neria Y, Nandi A, Galea S. Post-traumatic stress disorder following disasters: a systematic review. *Psychol Med*. 2008;38(4):467–480.
17. Bonanno GA, Galea S, Bucchiarelli A, et al. What predicts psychological resilience after disaster? The role of demographics, resources, and life stress. *J Consult Clin Psychol*. 2007;75(5):671–682.
18. Hikichi H, Aida J, Tsuboya T, et al. Can community social cohesion prevent posttraumatic stress disorder in the aftermath of a disaster? A natural experiment from the 2011 Tohoku earthquake and tsunami. *Am J Epidemiol*. 2016;183(10):902–910.
19. Gero K, Hikichi H, Aida J, et al. Associations between community social capital and preservation of functional capacity in the aftermath of a major disaster. *Am J Epidemiol*. 2020;189(11):1369–1378.
20. VanderWeele TJ, Knol MJ. A tutorial on interaction. *Epidemiol Method*. 2014;3(1):33–72.
21. Raker EJ, Arcaya MC, Lowe SR, et al. Mitigating health disparities after natural disasters: lessons from the RISK project. *Health Aff*. 2020;39(12):2128–2135.
22. Athey S, Tibshirani J, Wager S. Generalized random forests. *Ann Stat*. 2019;47(2):1148–1178.
23. Athey S, Wager S. Estimating treatment effects with causal forests: an application [preprint]. *arXiv*. 2019. <https://doi.org/10.48550/arXiv.1902.07409>. Accessed August 29, 2022.
24. Kondo K. Progress in aging epidemiology in Japan: the JAGES project. *J Epidemiol*. 2016;26(7):331–336.
25. Miyagi Prefecture. *Current Situations of Damage and Evacuation* [in Japanese]. <https://www.city.iwanuma.miyagi.jp/bosai/fukko/shinsaiaplli/documents/kirokushi.pdf>. Published 2016. Accessed August 29, 2022.
26. Tamiya N, Noguchi H, Nishi A, et al. Population ageing and wellbeing: lessons from Japan's long-term care insurance policy. *Lancet*. 2011;378(9797):1183–1192.
27. Koyano W, Shibata H, Nakazato K, et al. Measurement of competence: reliability and validity of the TMIG Index of Competence. *Arch Gerontol Geriatr*. 1991;13(2):103–116.
28. Cabinet Office of Japan. Certification of housing damage from disaster [in Japanese]. <http://www.bousai.go.jp/taisaku/unyou.html>. Accessed October 17, 2020.
29. Tsuboya T, Aida J, Hikichi H, et al. Predictors of depressive symptoms following the Great East Japan earthquake: a prospective study. *Soc Sci Med*. 2016;161:47–54.
30. Shiba K, Hikichi H, Aida J, et al. Long-term associations between disaster experiences and cardiometabolic risk: a natural experiment from the 2011 Great East Japan Earthquake and Tsunami. *Am J Epidemiol*. 2019;188(6):1109–1119.
31. Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. *Am J Epidemiol*. 2017;185(1):65–73.
32. Ripley B, Venables W, Ripley MB. Package 'nnet'. *R package version* 2016;7:3–12. <https://cran.microsoft.com/web/packages/nnet/nnet.pdf>. Accessed August 29, 2022.
33. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: Association for Computing Machinery; 2016:785–794.
34. Lendle SD, Schwab J, Petersen ML, et al. Ltmle: an R package implementing targeted minimum loss-based estimation for longitudinal data. *J Stat Softw*. 2017;81(1):1–21.
35. Polley E, LeDell E, Kennedy C, et al. Package 'SuperLearner'. 2019. <https://mran.microsoft.com/snapshot/2018-03-03/web/packages/SuperLearner/SuperLearner.pdf>. Accessed August 29, 2022.
36. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.
37. Tibshirani J, Athey S, Wager S, et al. Package 'grf'. 2018. <https://cran.r-project.org/web/packages/grf/grf.pdf>. Accessed August 29, 2022.
38. Chernozhukov V, Demirer M, Duflo E, et al. *Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India*. 2018. <https://doi.org/10.3386/w24678>. Accessed August 29, 2022.
39. Stekhoven DJ, Stekhoven MDJ. Package 'missForest'. 2012. <https://cran.r-project.org/web/packages/missForest/missForest.pdf>. Accessed August 29, 2022.
40. Hikichi H, Aida J, Kondo K, et al. Six-year follow-up study of residential displacement and health outcomes following the 2011 Japan earthquake and tsunami. *Proc Natl Acad Sci U S A*. 2021;118(2):e2014226118.
41. Norris FH, Friedman MJ, Watson PJ, et al. 60,000 disaster victims speak: part I. An empirical review of the empirical literature, 1981–2001. *Psychiatry*. 2002;65(3):207–239.
42. Nicholson NR. A review of social isolation: an important but underassessed condition in older adults. *J Prim Prev*. 2012;33(2–3):137–152.
43. Parker DA, Parker ES, Harford TC, et al. Status inconsistency and drinking patterns among working men and women. *Alcohol Clin Exp Res*. 1978;2(2):101–105.
44. Lin JC, Guerrieri JG, Moore AA. Drinking patterns and the development of functional limitations in older adults: longitudinal analyses of the health and retirement survey. *J Aging Health*. 2011;23(5):806–821.
45. Shiba K, Kawahara T, Aida J, et al. Causal inference in studying the long-term health effects of disasters: challenges and potential solutions. *Am J Epidemiol*. 2021;190(9):1867–1881.