Joint longitudinal and survival models: associations between natural disasters exposure, disability and death

Sam Brilleman[1,2], Theodore J. Iwashyna[3],

Margarita Moreno-Betancur[1,2,4], Rory Wolfe[1,2]

International Biometric Society Australasian Conference

2nd December 2015

[1] Monash University

[2] Victorian Centre for Biostatistics (ViCBiostat)

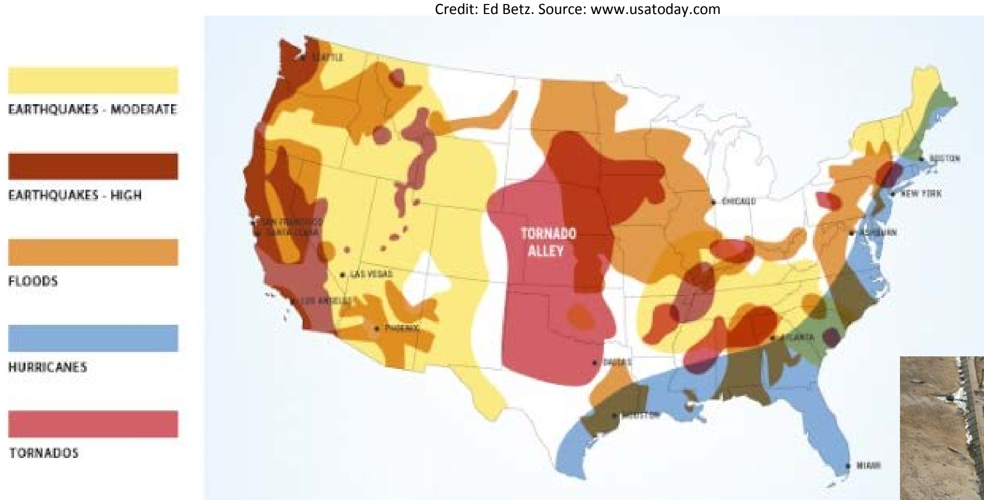
[3] University of Michigan

[4] Murdoch Childrens Research Institute

Credit: Ed Betz. Source: www.usatoday.com

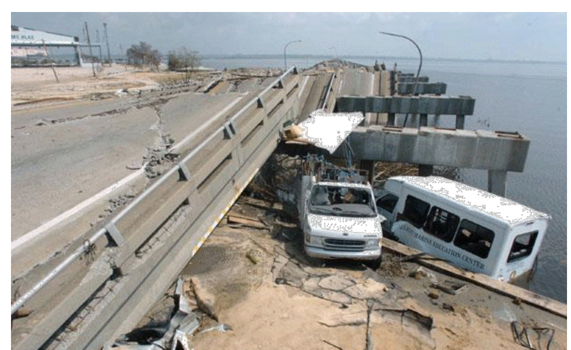
Source: www.flickr.com/photos/kevharb/4199300356/in/photostream

Credit: Steve Craven. Source: http://mercymedical.org

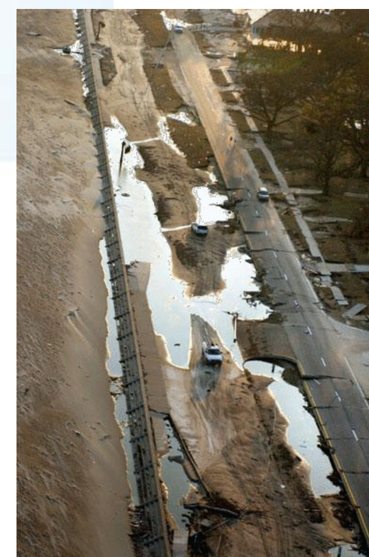


Source: www.noaa.gov

Source: http://www.theaustralian.com.au



Source: www.vanwinkle.org/biloxi.html

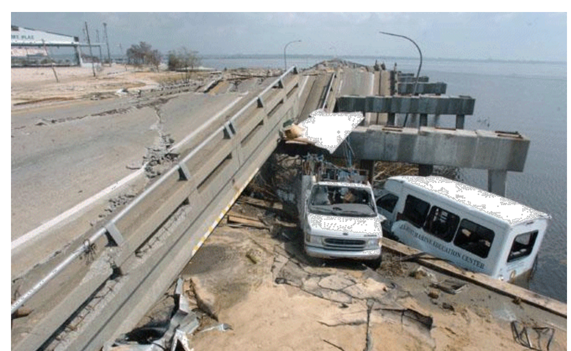


Source: www.vanwinkle.org/biloxi.html

Research question

Is natural disaster exposure associated with either individual-level changes in disability or the risk of death?

Source: http://www.theaustralian.com.au

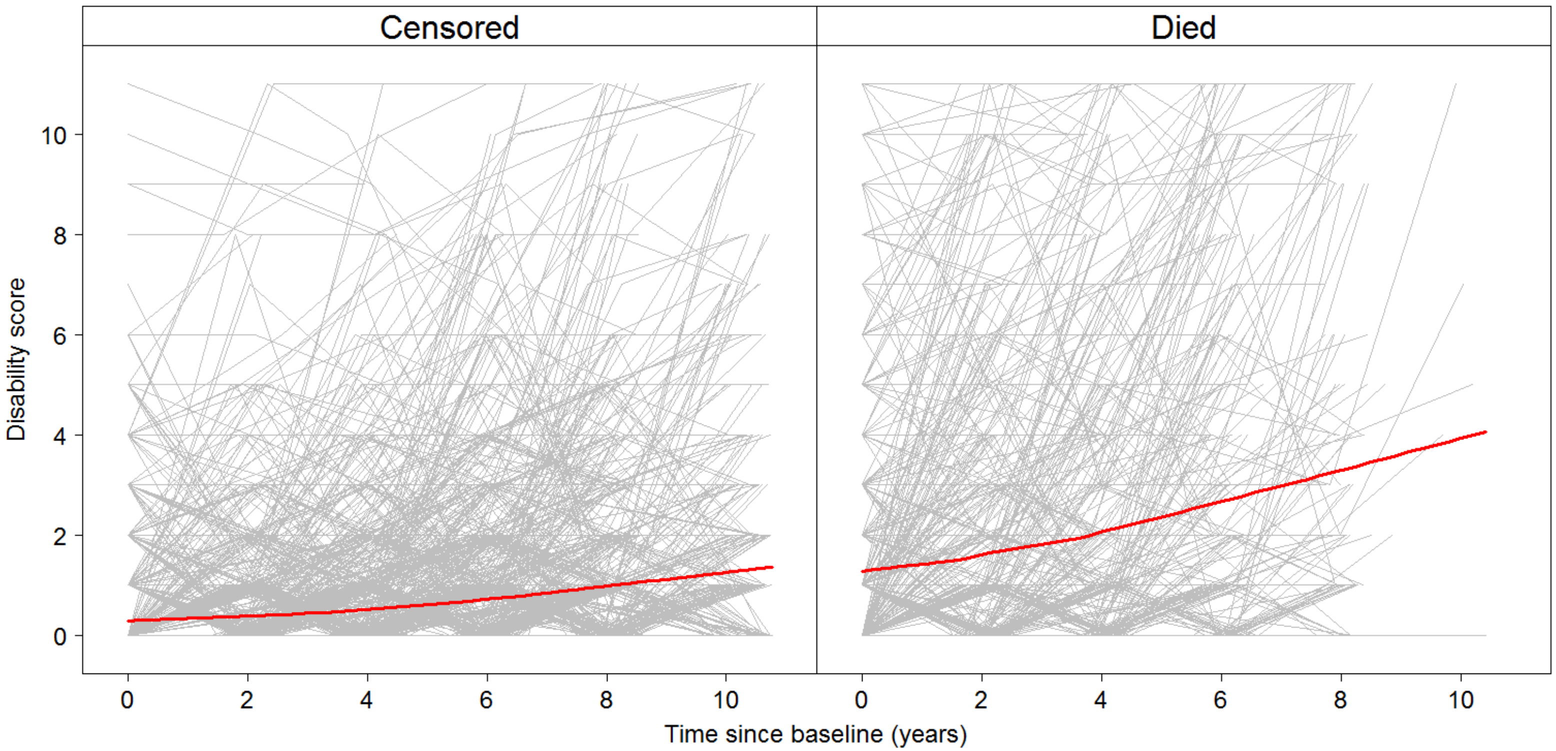


Source: www.vanwinkle.org/biloxi.html

Source: www.vanwinkle.org/biloxi.html

Data sources	U.S. Health and Retirement Study U.S. Medicare (deaths) Federal Emergency Management Agency (FEMA) database
Sample	17,559 participants, aged 50 to 90 years
Study period	1^{st} Jan 2000 – 30^{th} Nov 2010
Outcomes	Disability score (discrete, range from 0 to 11) Time to death or censoring
Exposure	Occurrence of a natural disaster within the previous 2 years (binary, time-varying)
Covariates	Baseline demographics (age, gender, race, wealth)

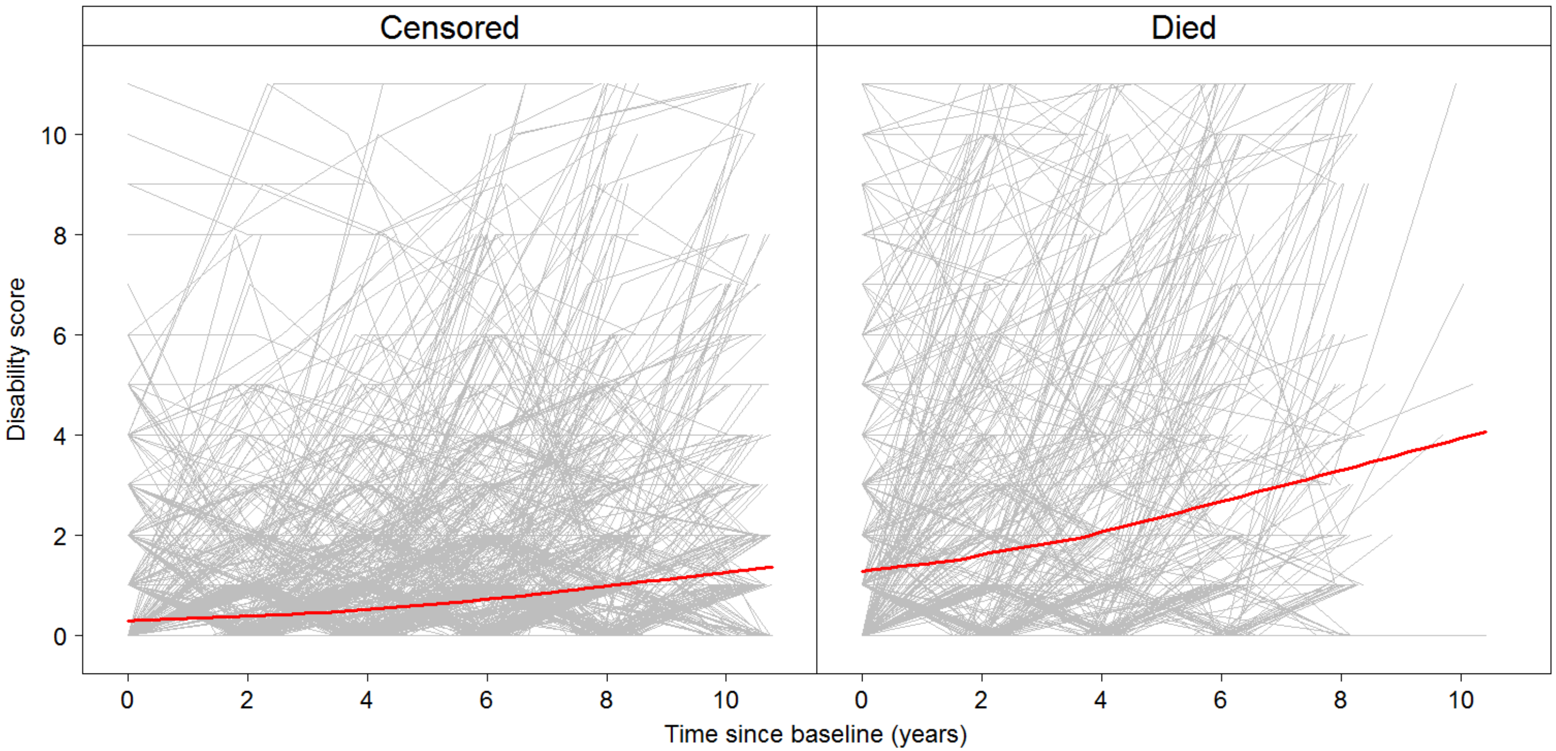
Observed disability score trajectories (and lowess smoothed average) for 2,458 individuals aged 70 to 75 years



MONASH University

ViCBiostat

Observed disability score trajectories (and lowess smoothed average) for 2,458 individuals aged 70 to 75 years



Association suggests non-random dropout due to death

- i.e., disability data are likely to be missing not at random (MNAR)

MONASH University

ViCBiostat

Joint model formulation

Longitudinal submodel (for disability score)

$y_i(t_{ij})$ is disability score for individual i at time point t_{ij}

$$y_i(t_{ij}) \sim NegBin\left(\mu_i(t_{ij}), \phi\right)$$

$$\eta_i(t_{ij}) = \log\left(\mu_i(t_{ij})\right) = \boldsymbol{x}_{\boldsymbol{i}}'(t_{ij})\boldsymbol{\beta} + b_{1i} + b_{2i}t_{ij}$$

Covariates $\boldsymbol{x}_{\boldsymbol{i}}(t_{ij})$: natural disaster exposure, time (linear slope), age category, age category * time interaction, gender, race, wealth decile (categorical)

Survival submodel (for time-to-death)

$$h_i(t) = h_0(t)\exp\left(\boldsymbol{w}_{\boldsymbol{i}}'(t)\boldsymbol{\gamma} + \alpha_1\eta_i(t) + \alpha_2\frac{d\eta_i(t)}{dt}\right)$$

Covariates $\boldsymbol{w}_{\boldsymbol{i}}(t)$: natural disaster exposure, age category, gender, race, wealth decile (linear trend), age category * wealth interaction

Joint model formulation

Longitudinal submodel (for disability score)

$y_i(t_{ij})$ is disability score for individual i at time point t_{ij}

$$y_i(t_{ij}) \sim NegBin(\mu_i(t_{ij}), \phi)$$

$$\eta_i(t_{ij}) = \log\left(\mu_i(t_{ij})\right) = \boldsymbol{x}_i'(t_{ij})\boldsymbol{\beta} + b_{1i} + b_{2i}t_{ij}$$

Covariates $\boldsymbol{x}_i(t_{ij})$: natural disaster exposure, time (linear slope), age category, age category * time interaction, gender, race, wealth decile (categorical)

Survival submodel (for time-to-death)

$$h_i(t) = h_0(t)\exp\left(\boldsymbol{w}_i'(t)\boldsymbol{\gamma} + \alpha_1\eta_i(t) + \alpha_2\frac{d\eta_i(t)}{dt}\right)$$

Covariates $\boldsymbol{w}_i(t)$: natural disaster exposure, age category, gender, race, wealth decile (linear trend), age category * wealth interaction

Joint model estimation

Bayesian approach, most flexible

Various software options, e.g.

- JMbayes package in R
  - Random walk Metropolis algorithm
  - Penalised splines for baseline hazard
  - Long run times for a large dataset:
    17,559 patients → 11 hours (for 26,000 MCMC iterations)!
- Stan (called from R using RStan)
  - Hamiltonian Monte Carlo algorithm
  - Encountered problems with the sampler getting stuck when using a large dataset

	Disability score ratios
Constant	0.02 (0.02 to 0.03)
Time (years)	1.02 (1.01 to 1.04)
Age category (ref: ≥50, <60y)	
≥60, <65y	0.92 (0.81 to 1.03)
⋮	⋮
≥80, <85y	5.62 (4.89 to 6.51)
≥85, <90y	9. 51 (7.96 to 11.34)
Age category * time interaction	
≥60, <65y	1.05 (1.03 to 1.06)
⋮	⋮
≥80, <85y	1.29 (1.26 to 1.32)
≥85, <90y	1.28 (1.25 to 1.32)
Gender (ref: Male)	
Female	1.02 (0.95 to 1.09)
Race (ref: White or Caucasian)	
Black or African American	1.30 (1.17 to 1.45)
Other	1.15 (0.95 to 1.39)
Wealth (ref: Decile 1, most wealth)	
Decile 2	1.10 (0.92 to 1.29)
⋮	⋮
Decile 9	5.31 (4.54 to 6.23)
Decile 10, least wealth	9.60 (8.22 to 11.24)
Disaster exposure	
Within previous 2 years	0.99 (0.92 to 1.04)

Older age → higher baseline disability

Older age → faster rate of increase

Non-white → higher average disability

Less wealth → higher average disability

No evidence that disaster exposure is associated with disability!

	Hazard ratios	
Age category (ref: ≥50, <60y)		
≥60, <65y	2.54 (1.05 to 6.16)	
⋮	⋮	
≥80, <85y	7.76 (3.31 to 17.03)	Older age → higher hazard
≥85, <90y	10.08 (3.81 to 23.71)	
Gender (ref: Male)		
Female	0.61 (0.53 to 0.68)	Males → higher hazard
Race (ref: White or Caucasian)		
Black or African American	0.90 (0.72 to 1.11)	White/Caucasian → higher hazard
Other	0.75 (0.46 to 1.15)	
Wealth trend across deciles		
Linear trend (0 = Decile 1; 9 = Decile 10)	1.15 (1.01 to 1.28)	Less wealth → higher hazard
Age category * wealth trend interaction		
≥60, <65y	0.92 (0.81 to 1.06)	
⋮	⋮	
≥80, <85y	0.89 (0.78 to 1.01)	But effect of wealth diminishes with age
≥85, <90y	0.87 (0.76 to 1.00)	
Disaster exposure		
Within previous 21 days	0.94 (0.56 to 1.43)	
Within previous 2 years, but not 21 days	1.02 (0.87 to 1.18)	← No evidence that disaster exposure is associated with death!
Association parameter		
Current value of linear predictor	1.54 (1.41 to 1.66)	
Current slope of linear predictor	1.62 (0.93 to 2.81)	

Natural disasters are common!

Disaster type	Number of individuals experiencing this disaster type at least once (%)	Number of person-disaster events (%)
Storm	12944 (74%)	28894 (45.2%)
Hurricane	6415 (37%)	16090 (25.2%)
Snow	5496 (31%)	10436 (16.3%)
Fire	3229 (18%)	4291 (6.7%)
Flood	1083 (6%)	1294 (2.0%)
Tornado	662 (4%)	662 (1.0%)
Earthquake	259 (1%)	259 (0.4%)
Other	1943 (11%)	1943 (3.0%)
All disasters	16075 (92%)	63869 (100%)

Natural disasters are common!

Disaster type	Number of individuals experiencing this disaster type at least once (%)	Number of person-disaster events (%)
Storm	12944 (74%)	28894 (45.2%)
Hurricane	6415 (37%)	16090 (25.2%)
Snow	5496 (31%)	10436 (16.3%)
Fire	3229 (18%)	4291 (6.7%)
Flood	1083 (6%)	1294 (2.0%)
Tornado	662 (4%)	662 (1.0%)
Earthquake	259 (1%)	259 (0.4%)
Other	1943 (11%)	1943 (3.0%)
All disasters	16075 (92%)	63869 (100%)

MONASH University

ViCBiostat

Natural disasters are common!

Disaster type	Number of individuals experiencing this disaster type at least once (%)	Number of person-disaster events (%)
Storm	12944 (74%)	28894 (45.2%)
Hurricane	6415 (37%)	16090 (25.2%)
Snow	5496 (31%)	10436 (16.3%)
Fire	3229 (18%)	4291 (6.7%)
Flood	1083 (6%)	1294 (2.0%)
Tornado	662 (4%)	662 (1.0%)
Earthquake	259 (1%)	259 (0.4%)
Other	1943 (11%)	1943 (3.0%)
All disasters	16075 (92%)	63869 (100%)

	Hazard ratios
Age category (ref: ≥50, <60y)	
≥60, <65y	2.54 (1.05 to 6.16)
⋮	⋮
≥80, <85y	7.76 (3.31 to 17.03)
≥85, <90y	10.08 (3.81 to 23.71)
Gender (ref: Male)	
Female	0.61 (0.53 to 0.68)
Race (ref: White or Caucasian)	
Black or African American	0.90 (0.72 to 1.11)
Other	0.75 (0.46 to 1.15)
Wealth trend across deciles	
Linear trend (0 = Decile 1; 9 = Decile 10)	1.15 (1.01 to 1.28)
Age category * wealth trend interaction	
≥60, <65y	0.92 (0.81 to 1.06)
⋮	⋮
≥80, <85y	0.89 (0.78 to 1.01)
≥85, <90y	0.87 (0.76 to 1.00)
Disaster exposure	
Within previous 21 days	0.94 (0.56 to 1.43)
Within previous 2 years, but not 21 days	1.02 (0.87 to 1.18)
Association parameter	
Current value of linear predictor	1.54 (1.41 to 1.66)
Current slope of linear predictor	1.62 (0.93 to 2.81)

Older age → higher hazard

Female → smaller hazard

Non-white → smaller hazard

Less wealth → higher hazard

But effect of wealth diminishes with age

No evidence that disaster exposure is associated with death!

"A one unit increase in the estimated log disability score is associated with a 54% increase in the hazard of death"

or

"A doubling in the estimated disability score is associated with a 35% increase in the hazard of death‡"

‡ Since a doubling in disability score is equivalent to a 0.693 unit increase in log disability score (i.e., log(2) = 0.693)

Association parameter	
Current value of linear predictor	1.54 (1.41 to 1.66)
Current slope of linear predictor	1.62 (0.93 to 2.81)

"A one unit increase in the estimated log disability score is associated with a 54% increase in the hazard of death"

or

"A doubling in the estimated disability score is associated with a 35% increase in the hazard of death‡"

‡ Since a doubling in disability score is equivalent to a 0.693 unit increase in log disability score (i.e., log(2) = 0.693)

"A one unit per year increase in the rate of change in estimated log disability score is associated with a 62% increase in the hazard of death"

or

"A doubling in the rate of change in estimated disability score is associated with a 40% increase in the hazard of death"

Association parameter	
Current value of linear predictor	1.54 (1.41 to 1.66)
Current slope of linear predictor	1.62 (0.93 to 2.81)

Conclusions

Able to estimate the effect of disaster exposure on disability, **even in the presence of non-random dropout** due to death

- i.e., disability data which was missing not at random (MNAR)

Able to estimate the effect of disaster exposure on death, **conditional on an individual's underlying level of disability**

- conditional on **both level and rate of change** in disability
- allowing for measurement error in the observed disability scores

Able to quantify the association between disability and death in a (hopefully!) meaningful way

What is a joint model?

The simultaneous estimation of two distinct “submodels” which traditionally would have been separately estimated

1. The **longitudinal submodel**:
   - A mixed effects regression model for a repeatedly measured marker (e.g., disability score)
2. The **event submodel**:
   - A proportional hazards regression model for a time-to-event outcome (e.g., time-to-death)

The two submodels are linked via shared parameters, and they are estimated under a single joint likelihood function

MONASH University

ViCBiostat

Benefits of joint modelling

Able to adjust for **longitudinal data which is MNAR**, by jointly modelling the longitudinal data and the dropout process

Able to include the longitudinal outcome as a time-varying covariate in the survival model, even when it is subject to **measurement error and measured intermittently**

Can investigate **associations between any aspect of the longitudinal trajectory and the event risk**

Can be used for “dynamic” risk prediction:

- fit joint model to available data
- predict event risk
- update event risk as new longitudinal data becomes available

Comparison with separate models

- Covariates associated with increased risk of death (e.g. less wealth) had disability score ratios which were attenuated under separate modelling

Increased risk of death → Disability score ratio **attenuated** under separate modelling

- Covariates protective against death (e.g. female) had disability score ratios which were over-estimated under separate modelling

Decreased risk of death → Disability score ratio **over-estimated** under separate modelling

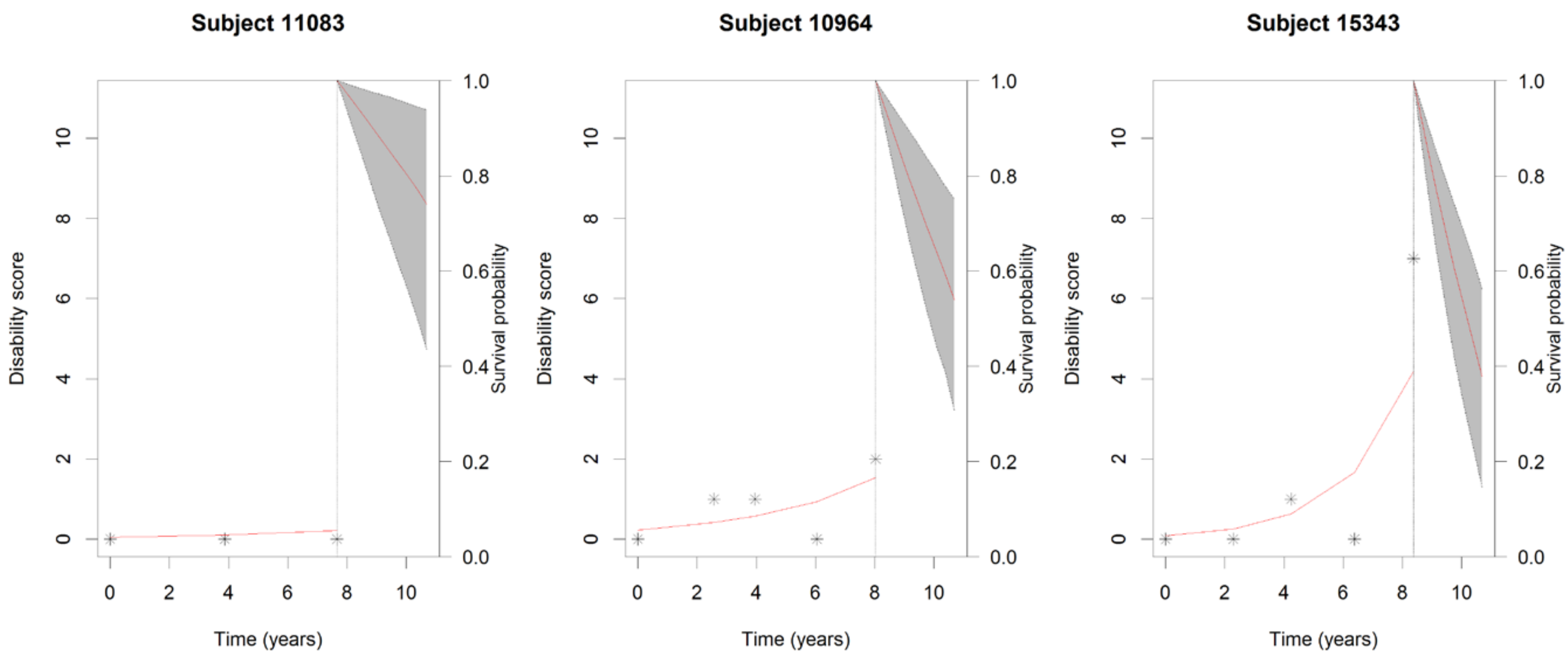
MONASH University

ViCBiostat

Joint models for each disaster type

Disaster exposure type	Longitudinal submodel: disability score ratios	Survival submodel: hazard ratios
Storm	0.99 (0.93 to 1.05)	1.03 (0.94 to 1.12)
Hurricane	0.99 (0.92 to 1.06)	1.03 (0.92 to 1.15)
Snow	1.00 (0.93 to 1.07)	1.03 (0.89 to 1.18)
Fire	1.00 (0.91 to 1.09)	0.99 (0.84 to 1.15)
Flood	0.86 (0.72 to 1.03)	0.72 (0.44 to 1.09)
Tornado	1.18 (0.87 to 1.59)	1.67 (1.12 to 2.38)
Earthquake	0.96 (0.72 to 1.30)	1.34 (0.64 to 2.43)

Results are from 7 separate joint models (one for each disaster type) and each model if adjusted for the same baseline covariates as for previous models (age, gender, race and wealth)



Subject 11083
Subject 10964
Subject 15343
Disability score
Survival probability
Time (years)
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0