Pre-Peer Review Version of Paper: Kushnick G, Behie A, Zuska F (2021). Pregnancy outcomes among evacuees of the Sinabung volcano, 2010-2018 (North Sumatra, Indonesia): A Matched Cohort Study. *American Journal of Human Biology*, e23638, https://doi.org/10.1002/ajhb.23628

Pregnancy Outcomes among Evacuees of the Sinabung Volcano, 2010-2018 (North Sumatra, Indonesia): A Matched Cohort Study

Geoff Kushnick[1], Alison Behie[1], and Fikarwin Zuska[2]

[1] School of Archaeology and Anthropology, Australian National University, Canberra, Australia
[2] Department of Anthropology, University of North Sumatra, Medan, Indonesia

ORCID: 0000-0001-9280-0213 (GK), 0000-0002-7436-3694 (AB), 0000-0001-7642-8546 (FZ)

ABSTRACT:

Objectives: Exposure to natural disasters during pregnancy is associated with adverse birth outcomes. Nonetheless, relatively little work has been done on volcanic eruptions, or on populations in developing countries that are disproportionately impacted and often less resilient and difficult to track down in the aftermath. We conducted a matched cohort study among women living in villages near the Sinabung volcano in North Sumatra, Indonesia—a long-dormant volcano, active since 2010.

Methods: We used bivariate and multivariate methods to compare an 'exposed' sample of women (n=97) who were pregnant when forced to evacuate their villages due to the volcanic eruptions and an 'unexposed' sample of non-evacuees (n=97) matched for age and year of child's birth. We collected anthropometric data and conducted structured interviews consisting of questions about maternal and birth characteristics, and stress.

Results: Evacuation during pregnancy led to an almost-five-fold increase in the adjusted odds of preterm (gestation length of 37 weeks or less) or early-term (gestation length up to 38 weeks) birth and a 1 cm decrease in birth length but had no effect on birth weight or sex of child.

Conclusions: Both adverse effects we documented have the potential to exert a negative influence on later-in-life outcomes for the children of women pregnant during evacuation. This should be considered when exploring protocols and policies around the evacuation of pregnant women following future volcanic eruptions.

KEYWORDS: Natural disasters, maternal/child health, birth length, birthweight, early born

Introduction

Exposure to natural disasters is known to cause an increase in stress, particularly in vulnerable groups such as pregnant women who already have hormonal systems under constant flux. This has negative consequences for birth outcomes and maternal/child health, such as low birthweight and higher incidence of preterm birth following fires (O'Donnell & Behie, 2013), floods (Sanguanklin et al., 2014; Tong, Zotti, & Hsia, 2011), earthquakes (Palmeiro-Silva et al., 2018; Tan et al., 2009), cyclones and hurricanes (Parayiwa & Behie, 2018; Xiong et al., 2008), and ice storms (Auger, Kuehne, Goneau, & Daniel, 2011). Many of these adverse birth outcomes are associated with increased morbidity and mortality throughout childhood and adolescence (McIntire et al, 1999) and low birthweight is a known predictor of poor infant health in the first year of life (Mace, 2000). With the projected increases in the frequency and intensity of natural disasters in coming years, and drastic increases in the global population of displaced peoples, it is essential we better understand how such changes in climate will impact pregnancy and birth outcomes.

In addition to the negative consequences reported above, stress during pregnancy has also been associated with population level changes such as a decrease in the secondary sex ratio (sex ratio at birth) with fewer males being born (Bruckner, Catalano, & Ahern, 2010; R. A. Catalano, Currier, & Steinsaltz, 2015; Navara, 2010). This has resulted in decreased secondary sex-ratio being occasionally used as an indicator of population-wide stress (R. Catalano, Bruckner, Gould, Eskenazi, & Anderson, 2005). Although other factors may influence secondary sex-ratios, and evidence from natural disasters is limited, the overall picture seems to support declines in secondary sex ratio being recorded following earthquakes further supporting the impact of natural disasters on births (Fukuda et al, 1998; Saadat, 2008; Torche & Kleinhaus, 2012).

Although developing countries are disproportionately impacted by natural disasters (Bennett & Friel, 2014; Hanna & Oliva, 2016) and are often less resilient in the aftermath of a disaster, few studies to date have explored how natural disasters impact reproduction in this context, and none have considered volcanoes. Our study focuses on pregnant women affected by the eruptions of the Sinabung volcano (Gunung Sinabung), located in the Bukit Barisan mountains of North Sumatra, Indonesia, about three hours by car from the provincial capital, Medan

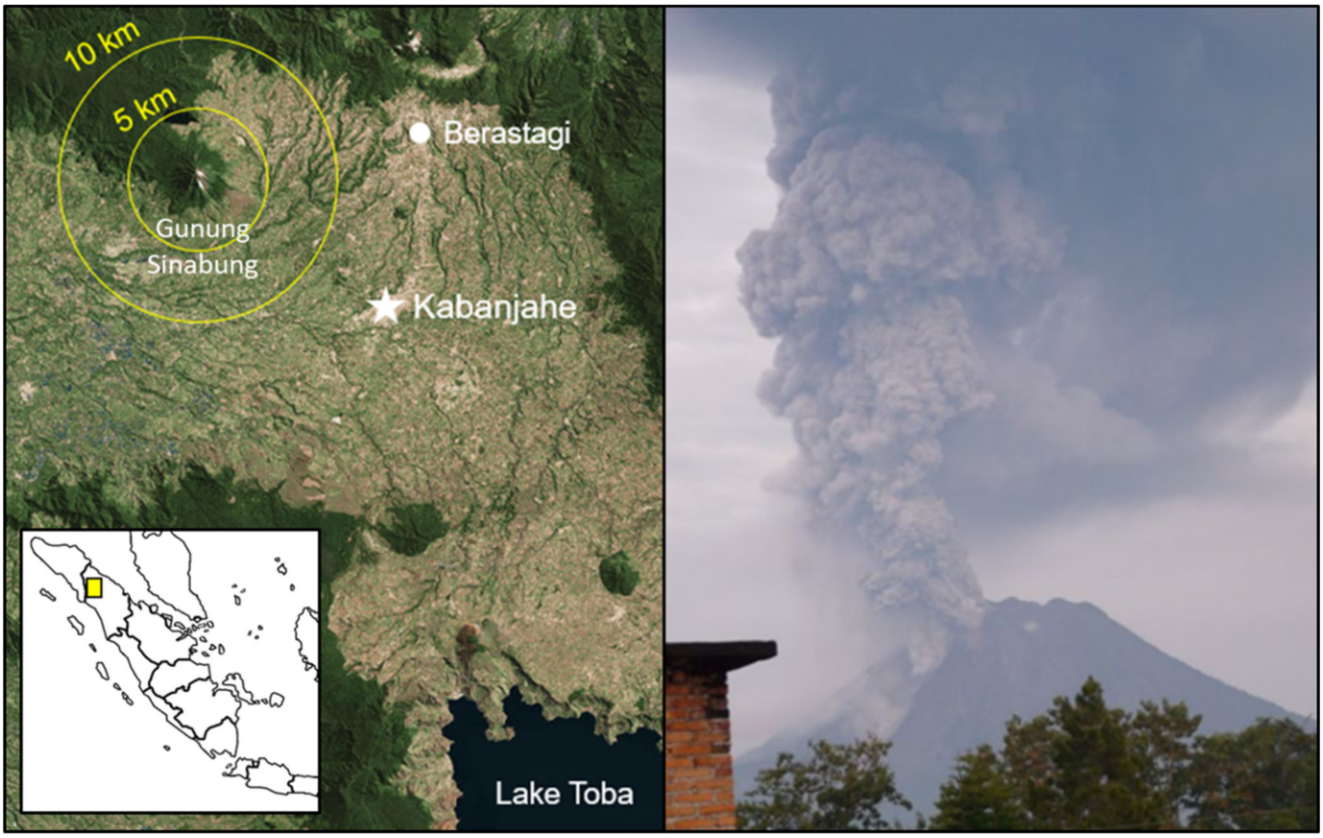

Figure 1 Map of research site (left) and picture of Gunung Sinabung erupting in 2010 (right): Participants from the evacuee sample all originated from villages (n=9) within 10 km radius of Gunung Sinabung, 6 from within 5 km radius. Participants from the non-evacuee sample originated from villages (n=10) in the Karo area (the approximate area shown on the map). Of these, 8 were outside, and 2 on the rim of or just inside, the 10 km radius. Picture taken from Kabanjahe by GK on the morning after the first eruption in 2010.

(see **Figure 1**). The volcano, which had sat dormant for hundreds of years, became active in 2010 and continues to erupt to this day. At its peak in 2014, there were 32,355 evacuees living in temporary relocation shelters. The people affected by these eruptions, by and large, belong to an ethnic group called the Karo (Kipp, 1993; Singarimbun, 1975; Steedly, 1993). This group consists mostly of rural farmers whose reproductive-related behaviour has been studied in detail by one of us, and whose fertility rates (3 to 4.5 children per woman over her lifetime) and child mortality rates (from 30 to 60 deaths by age 5 per 1,000) are higher than the national average (Kushnick, 2009, 2010; Kushnick & Fessler, 2011; Kushnick, Fessler, & Zuska, 2016).

To explore how evacuations from volcanic eruptions impact reproductive outcomes, we used a retrospective matched cohort-study design (Euser, Zoccali, Jager, & Dekker, 2009) to assess the effect of evacuation on birth outcomes, including birthweight, birth length, gestation length, and sex of offspring, in a sample of 194 participants. The exposed ('evacuee') sample consisted of women who had ever been evacuated while pregnant from their villages due to the eruptions of the Sinabung volcano that has been active since 2010. The unexposed ('non-evacuee') sample consisted of women who lived in the same region but had never been evacuated due to the volcano. We hypothesize that evacuee women will have an increased incidence of low birth weight and pre-term births compared to women who have never been evacuated. Our study illustrates the difficulties of getting an adequate sample in the context of poor or incomplete record-keeping related to the current whereabouts of long-term evacuees and relocations. For this reason, we have highlighted potential sources of bias and how we were able to deal with them in the study design and analyses. We expect results of this study to improve our understanding of the human response to disasters by providing the first information of volcanic eruption impacts as well as broadening available literature on different cultures and in developing nations.

Methods

Sampling

The participants in the evacuee sample (n=97) were recruited from: (a) evacuation centers; (b) relocation centers, found both in existing villages and new villages custom-built for relocating evacuees; and, (c) villages that had been evacuated at least once but had since had their populations return. All nine villages of origin for the participants were within a 10 km radius of the volcano, and six were within the 5 km radius (see **Figure 1**). After obtaining permission from the respective relocation center directors and village heads, we asked to be introduced to women who had ever been pregnant during an evacuation between the first eruption in 2010 and the time of study which was 2018. As we had no reliable information on how many women fell into each evacuation class, we were unable to sample from each source in a balanced way, nor were we able to use sampling weights in our analyses. Rather, we aimed to recruit all women from each location who fit the criteria to ensure we had as large a sample as possible. We realize the non-randomized sampling method may have introduced bias, which we discuss below.

The participants in the non-evacuee sample (n=97) were matched (Braga, Forrokhyar, & Bhandari, 2012) with the evacuee sample on two potential confounding variables: (a) child's year of birth, which ranged from 2010 to 2017; and, (b) the mother's age during the pregnancy broken into three age categories to make the sampling procedure tractable over the amount of time we had to recruit participants (18-25 years, 25-36 years, and 36 years or older). For a complete tally of participants in each stratum, see **Table S1** in the Supplementary Materials. Further, women in the non-evacuee sample needed to never have been evacuated from their villages due to the Gunung Sinabung eruptions. They may also have experienced some stress but, crucially, they were never evacuated. Women were recruited from 10 villages in the region, 8 outside, and 2 just on or within the rim of, the 10 km radius of the volcano (see **Figure 1**). After obtaining permission from the heads of villages, we asked the village nurse to help us identify woman who fit the matching criteria and we chose randomly from amongst them.

Data Collection

Data were collected from each participant using structured interviews and anthropometry over two months (December 2017 – January 2018). The interviews consisted of questions about: (a) basic information regarding the participant, such as date of birth, marriage status, education, religion, and occupation; (b) the context and outcome of relevant pregnancies; and, (c) amongst evacuees only, questions about the timing and context of evacuations, including subjective and objective measures of stress. The subjective measures of stress were collected using a translated version of the 'Impact of Events Scale – Revised' (IES-R), a relatively easy-to-administer test used as a screening tool for post-traumatic stress disorder (Horowitz, Wilner, & Alvarez, 1979; Weiss, 2007). We made one slight modification: In the original IES-R, participants are asked about the last 7 days; we asked about any time between the first evacuation and now. Interviews were conducted in private in Bahasa Indonesia or Bahasa Karo by two trained research assistants. Height was measured using a portable stadiometer (Charder HM200P), weight with a portable scale (Charder MS6111).

Variables

The data consisted of six outcome variables, the exposure variable, nine potential confounders, and a calculated measure of total subjective stress. Outcome variables were: (1)(a) *Birthweight (kg)*; (b) *Low Birthweight*: whether the child weighed less than 2500g (2.5kg) at birth, which is the recognized standard (UNICEF & Organization, 2004); and, (c) *Low Birthweight 2*: whether the child weighed 2500kg (2.5kg) or less at birth, an adjustment for the 'heaping' of reported birthweights at the threshold (Blanc & Wardlaw, 2005; Channon, Padmadas, & McDonald, 2011); (2) *Birth Length (cm)*; (3) *Gestation Length*: whether gestation length was less than 9 months (i.e., fell into the preterm or early-term categories); and, (4) *Sex of Offspring*: whether the child was female. See **Table S2** in the Supplementary Materials for more information about outcome variables and their clinical definitions.

The exposure variable was whether the participant was evacuated during pregnancy (see *Sampling* above for more details). Potential confounders consisted of the following: (1) *Child's Year of Birth*; (2) *Mother's Age (years)*; (3) *Mother's Height (cm)*; (4) *Mother's Body Mass Index (BMI):* kg/m^2, a risk factor for adverse birth outcomes (Cedergren, 2007); (5) *Mother's Religion*: whether the woman was Christian, included because the local population, while majority Christian, has sizable Muslim minority; (6) *Mother's Education*: whether the woman has at least some high school education, a cut off chosen because it is the first level that is not universally available and free in Indonesia; (7) *Mother's Occupation*: Whether the woman was a farmer—in the rural villages around 95% of women are farmers (Kushnick, 2010); (8) *Betel Nut*: whether the woman is a betel nut (*Areca catechu*) chewer, for which there is mixed evidence for being associated with adverse birth outcomes (Grant & Cardarelli, 2014); and, (9) *Smoking*: whether the woman is a smoker.

The measure of subjective stress was collected amongst the evacuee sample using the 'Impact of Events Survey – Revised' (IES-R) questionnaire (Horowitz et al., 1979; Weiss, 2007). Here, we provide an initial exploration of these data. The IES-R includes a standardized battery of 22 statements followed by a response on the following scale: not at all (0); a little bit (1); moderately (2); quite a bit (3); or, extremely (4). The statements are listed in **Table S3** in the Supplementary Materials but included the following for example: "Any reminder brought back feelings about it" and "Pictures about it popped into my mind". The questions were translated to Bahasa Indonesia and prefaced with instructions to answer with reference to any time since the first eruption. The IES-R responses are then transformed into three sub-scores: Avoidance (mean of items 5, 7, 8, 11, 12, 12, 17, and 22); Intrusion (mean of items 1, 2, 3, 6, 9, 14, 16, and 20); and, Hyperarousal (mean of 4, 10, 15, 18, 19, and 21). The Total Mean Score is the sum of the three means. To facilitate our initial exploration of these data, we categorized this into Low (0 to 5), Moderate (5.001 to 7.5), and High (7.501 and higher).

Analyses

We conducted both bivariate and multivariate analyses to estimate the effect of evacuation on various pregnancy outcomes. For the bivariate analyses, we used standard statistical tests—t-tests for continuous outcomes and contingency tables, with chi-square and Fisher's exact tests, for binary outcomes. For the multivariate analyses, which allowed us to adjust for potential confounding variables, we used logistic regression (Hosmer & Lemeshow, 2000) for binary outcomes and linear regression (Weisberg, 2005) for continuous outcomes. The multivariate analyses were conducted over two stages. The first stage analyses aimed at identifying which variables confounded the relationship between exposure and the outcome of interest (McNamee, 2005). To do so, we compared the beta coefficients in the baseline (bivariate) model with one that included each potential confounder. Any added variable that changed the beta estimate (i.e., $\%\Delta\beta$) was 10% or greater was treated as a confounder. We included the matched variables in this scheme, as matching often but not always alleviates confounding (Sjölander & Greenland, 2013). In the second stage of analyses, we built multivariate versions of the models with identified confounders included to estimate adjusted effects of exposure. For hypothesis testing, α was set to 0.05. Because there were missing data points, we conducted all analyses twice—once with casewise deletion and once with imputed datasets. The two-stage multivariate analyses let us eliminate covariates from Stage 2 models that showed no evidence of confounding the relationship between exposure and outcome in the Stage 1 models. This was necessary to best account for small sample size and a small number of events for some outcomes.

To impute missing data, we used multiple imputation with chained equations (Liublinska & Rubin, 2012; Mckinnon, 2010; White, Royston, & Wood, 2010). We created 20 imputed datasets twice, once including the continuous versions of the outcome variables, once including the binary versions. The imputations were done using the *mi* routine in Stata 15.1. The imputation models were linear for continuous variables and logit with augmentation in the case of perfect prediction for binary variables. All models included the exposure variable (i.e., whether the individual was part of the evacuee or non-evacuee sample). Missing values for both independent and dependent variables were imputed. Missingness, for the most part, was the result of the inability of the respondents to recall the answers to some questions, which was likely due to a mix of random and non-random factors. For instance, birth length was missing for 39% of births in the evacuee sample but only 22% in the non-evacuee sample (χ^2=7.04, *df*=2, *p*=0.008). This was not the case for all variables. For instance, whether it was preterm or not was missing for 8% of births in the evacuee sample and only 6% in the non-evacuee sample (Fisher's exact test: *p*=0.783). All analyses and imputations were done using Stata 15.1.

Biases

We identified four potential sources of bias, and dealt with them as follows: First, retrospective cohort studies, unlike prospective ones, cannot provide definitive evidence for the temporal sequence of events necessary for causation (Euser et al., 2009). We feel, however, that our design established convincingly the temporal precedence of exposure to outcome as the information about thing timing of evacuations was a community-level phenomenon.

Second, sampling bias may have arisen due to our recruitment strategy. We had no way of building a sampling frame from which to draw a random sample from the overall population of woman who were pregnant during evacuation. We found it challenging to identify the current whereabouts of some long-term evacuees and relocations. Further, we were unable to find a large enough sample of women with the requisite exposure to conduct systematic sampling. Rather, we relied on village nurses and evacuation center directors to introduce us to potential respondents and included all women with the requisite exposure. This may have led to some women to be over- or under-represented. Children that died, for instance, may have been missed due to cultural norms in the communities studied—similar issues arose in earlier research conducted by one of us amongst the Karo (Kushnick, 2010).

Third, as with all retrospective studies, recall bias is possible. One way to deal with this is to ask respondents to produce official birth records. In this case, we did ask to view records if they were available, but we also allowed for recalls for the following reasons. We were concerned that the non-evacuee sample would be more likely than the evacuee sample to have records available. This ended up being the case, with 16% of non-evacuees compared to 1% of evacuees able to produce a record (Fisher's exact: *p*=0.0003). Further, previous studies have shown that recall bias is more problematic for mundane events than important ones like the ones we were studying (Casey, Rieckhoff, & Beebe, 1992; Shenkin et al., 2017). Relatedly, some respondents could not recall important information, so we were left with missing data. Rather than drop those cases, we used multiple imputation (White et al., 2010) in the multivariate analyses and compared these results with estimates from casewise-deleted data.

Fourth, because those exposed to the ill effects of natural disasters may differ in important ways from those unexposed, one must be cautious treating natural disasters as natural experiments (Remler & Van Ryzin, 2015). Our design allowed for the control of potential confounders via sampling and statistical modeling (McNamee, 2005; Pourhoseingholi et al, 2012).

Results

Bivariate Analyses

BINARY VARIABLES: Bivariate comparison of the evacuee and non-evacuee samples for the binary variables is illustrated in the top panel of **Table 1**. We used chi-square tests, except in the case of small sample sizes where we used Fisher's exact tests. Only 1 of 4 binary outcomes showed a statistically significant difference. Compared to non-evacuees, evacuees were more likely to have either preterm or early-term births (exposed 12%, unexposed 4%). There was no difference in low birthweight in the normal coding (exposed 6%, unexposed 5%) nor the alternative coding (exposed 11%, unexposed 11%). There was a difference in sex of child (exposed 47% female, unexposed 40% female), but it was not a statistically significant difference. Of the binary potential confounding variables, 4 of 5 were significantly different. Compared to non-evacuees, evacuees were more likely to be farmers (exposed 92%, unexposed 68%) and chew betel nut (exposed 64%, unexposed 37%), and were less likely to be Christian (exposed 54%, unexposed 78%) and to have at least a high school education (exposed 59%, unexposed 76%). There was no statistical difference in smoking.

CONTINUOUS VARIABLES: Bivariate comparisons of the evacuee and non-evacuee samples for the continuous variables are presented in the bottom panel of **Table 1** (and illustrated in **Figure S1** in the Supplementary Materials). All of the variables were reasonably normally distributed, so we used t-tests. The sampling strategy was effective in leveling the samples on the matched variables. Child's year of birth was identical in the two samples; participant age differed

Table 1 Bivariate analyses for binary (top) and continuous (bottom) variables.

Binary Variable		Count		Missing	χ^2	*P*	
		Evac	Non				
OUTCOMES							
Low BW:	*Yes*	6	5	6	-- [A]	0.485	
	No	87	90				
Low BW2:	*Yes*	9	9	6	0.002	0.962	
	No	84	86				
Gestation Length:	*Early/Preterm*	11	4	14	-- [A]	0.047	*
	Normal	78	87				
Sex of Child:	*Female*	45	39	1	0.87	0.350	
	Male	51	58				
CONFOUNDERS							
Religion:	*Christian*	52	76	0	13.23	0.0003	***
	Muslim	45	21				
Education:	*HS or higher*	57	74	0	6.79	0.009	**
	Up to Jr HS	40	23				
Occupation:	*Farmer*	89	66	0	16.15	0.00006	***
	Other	8	30				
Betel Nut:	*Yes*	62	36	0	13.94	0.0002	***
	No	35	61				
Smoking:	*Yes*	1	0	0	-- [A]	0.500	
	No	96	97				
Continuous Variable		**Mean**		**Missing**	***t***	***p***	
		Evac	Non				
OUTCOMES							
Birth weight (kg)		3.3	3.2	6	-0.694	0.244	
Birth length (cm)		48.9	50.3	59	3.321	0.0012	**
CONFOUNDERS							
Age (yrs)		30.4	30.7	0	0.305	0.761	
Child's year of birth		2013.9	2013.9	0	--	--	
Height (cm)		150.1	150.7	0	0.938	0.342	
BMI		26.2	26.9	1	0.929	0.354	

[A] Fisher's exact test
Significance: *p<0.05, **p<0.01, ***p<0.001

Table 2 Confounding analyses using logistic regression for binary outcomes and linear regression for continuous outcomes: Coefficient estimates for exposure (β_E) and their change ($\%\Delta_\beta$) when potential confounders are added.

	Binary Outcomes								Continuous Outcomes			
	Pre/Early		LowBW		LowBW2		SEX		BW		BLength	
	β_E	$\%\Delta_\beta$	β_E	$\%\Delta_\beta$	β_E	$\%\Delta_\beta$	β_E	$\%\Delta_\beta$	β_E	$\%\Delta_\beta$	β_E	$\%\Delta_\beta$
MISSING DATA:												
Exposure (Baseline)	1.12	--	0.22	--	0.02	--	0.27	--	0.05	--	-1.39	--
Matched												
+ Age (Child)	1.14	2	0.22	0	0.03	50	0.27	0	0.05	0	-1.39	0
+ Age (Mother)	1.12	0	0.20	-9	0.02	0	0.26	-4	0.05	0	-1.31	-6
Unmatched												
+ Sex	1.12	0	0.20	-9	-0.01	-150	--	--	0.07	40	-1.37	-1
+ Religion	1.05	-6	0.09	-59	-0.18	-1000	0.31	15	0.05	0	-1.10	-21
+ Occupation	1.43	28	0.03	-86	-0.14	-800	0.24	-11	0.07	40	-1.41	1
+ Education	1.03	-8	0.41	86	0.20	900	0.33	22	0.04	-20	-1.33	4
+ Betel Nut	1.42	27	0.29	32	0.10	400	0.18	-33	0.03	-40	-1.39	0
+ BMI	1.08	-4	0.18	18	-0.02	-200	0.28	4	0.06	20	-1.42	2
IMPUTED DATA:												
Exposure (Baseline)	1.11	--	0.29	--	0.16	--	0.27	--	0.05	--	-1.39	--
Matched												
+ Age (Child)	1.11	0	0.29	0	0.16	0	0.27	0	0.05	0	-1.40	1
+ Age (Mother)	1.12	1	0.30	3	0.16	0	0.26	-4	0.05	0	-1.39	0
Unmatched												
+ Sex	1.10	-1	0.27	-7	0.11	-62	--	--	0.06	20	-1.39	0
+ Religion	1.06	-5	0.20	-31	-0.02	-107	0.30	11	0.06	20	-1.17	-16
+ Occupation	1.42	28	0.12	-59	-0.00	-100	0.21	-22	0.07	40	-1.38	-1
+ Education	1.01	-9	0.43	48	0.28	-3	0.32	19	0.04	-20	-1.34	-4
+ Betel Nut	1.39	25	0.36	24	0.23	-21	0.17	-37	0.03	-40	-1.40	1
+ BMI	1.08	-3	0.26	-10	0.12	-59	0.25	-7	0.06	20	-1.44	4

slightly because matching was done using age-class categories, but the difference was not substantive nor statistically significant. There was no substantive or statistically significant difference between evacuee and non-evacuee samples for the other two confounding variables, woman's height (cm) and BMI. Only one of the two continuous outcomes differed. Children who were *in utero* during their mother's evacuations were 1.4 cm shorter at birth than children born to non-evacuee mothers, and the difference was statistically significant (p=0.0095). There was no substantive or significant difference for birth weight (kg).

Multivariate Analyses

STAGE 1 (CONFOUNDING): To assess which of the potential confounding variable s would need to be included in the models used to estimate adjusted effects of exposure, we built baseline models with exposure as the sole independent variable and each pregnancy outcome as the dependent variable. For binary outcomes, we used logistic regression; for continuous, we used linear regression. We then estimate percent change ($\%\Delta_\beta$) in the beta coefficients when each potential confounder was added using 10% as the threshold change necessary for a variable to be considered a bonafide confounder to be included in the models in Stage 2. The results of the confounding analysis are shown in **Table 2**. For the most part, the sampling scheme alleviated the confounding effects of the matched variable with only one exception—its inclusion resulted in a 50% increase in the beta coefficient for exposure in the low birthweight analysis.

STAGE 2 (ADJUSTED ESTIMATES): To assess the adjusted effect of exposure on pregnancy outcomes, we used multiple logistic regression for binary outcomes and multiple linear regression for continuous outcomes (see **Table 3**). As explained above, all continuous outcomes were reasonably normally distributed. Amongst the binary outcomes, the only substantive and statistically significant effect of exposure was an almost-five-fold increase in the odds of being born before 37 weeks gestation (pre- or early-term births) even when controlling for the variables identified as confounders. Exposure was not associated with substantive or statistically significant changes in the other binary outcomes, such as low birthweight or sex of offspring. Amongst the continuous outcomes, the only substantive and statistically significant effect of exposure was around a 1 cm decrease in birth length (cm) when controlling for variables identified as confounders. Exposure was not associated with substantive or significant changes in birthweight (kg).

Table 3 Adjusted estimates of the effect of evacuation ('exposure') using multiple logistic regression for binary outcomes and multiple linear regression for continuous outcomes.

Binary Outcome	Controls[A]	Data	n	Cov/n[B]	OR	95% C.I. Lower	95% C.I. Upper	*p*	
Pre/Early	O/B	Missing	179	59.7	4.84	1.31	17.92	0.018	*
	O/BN	Imputed	194	64.7	4.84	1.29	19.19	0.020	*
Low BW	R/O/E/BN/B	Missing	187	31.2	0.99	0.25	3.92	0.987	
	R/O/E/BN/B	Imputed	194	32.3	1.08	0.28	4.21	0.907	
Low BW2	CA/R/O/E/BN/B/S	Missing	186	23.3	0.67	0.21	2.16	0.508	
	R/O/BN/B/S	Imputed	194	32.3	0.79	0.26	2.38	0.673	
GIRL	R/O/E/BN	Missing	192	38.4	1.30	0.68	2.47	0.428	
	R/O/E/BN	Imputed	194	38.8	1.24	0.65	2.35	0.516	
Continuous Outcome	**Controls[1]**	**Data**	**n**	**Cov/n[B]**	**β**	**95% C.I. Lower**	**95% C.I. Upper**	***p***	
BW (kg)	O/E/BN/B/S	Missing	186	31.0	0.05	-0.11	0.21	0.510	
	R/O/E/BN/B/S	Imputed	194	27.7	0.06	-0.11	0.22	0.484	
BL (cm)	R	Missing	135	67.5	-1.10	-1.96	-0.24	0.012	*
	R	Imputed	194	97.0	-1.17	-1.20	-0.36	0.005	**

[A] CA=Child's Age; R=Religion; O=Occupation; E=Education; BN=Betel Nut; B=BMI; S=Sex
[B] Cov/n = Covariates per observation used to estimate model
Significance: *$p<0.05$; **$p<0.01$

Importantly, although there were some minor changes in effect sizes between the missing-data and imputed-data estimates, they did not change the overall conclusion that there were substantive and statistically significant effects of exposure on being born before 37 weeks gestation (pre- or early-term birth) and birth length (cm), but not on the other outcomes investigated. As shown in **Table 3**, there was variation in the number of covariates per observation in the models, with a range of 23.3 to 97.0, but none was sufficiently low to warrant caution. The number of events per covariate, on the other hand, was low for some outcomes.

Discussion

We used a matched cohort-study design and multivariate modelling to compare birth outcomes for women pregnant during an evacuation from the eruptions of Gunung Sinabung (North Sumatra, Indonesia) between 2010 and 2018 (the 'exposed' group) and non-evacuees living in the same region (the 'unexposed' group). We found that evacuees fared worse for some outcomes with an almost-five-fold increase in the adjusted odds of pregnancies ending pre-term, and that the children of evacuees were around 1 cm shorter at birth than the children of non-evacuees. We found no statistical difference between evacuee and non-evacuee children in birthweight or sex of offspring.

In our study, evacuee pregnancies ended earlier, which supports previous studies that show an increased incidence of early deliveries, including preterm and early-term births, following natural disasters (Harville, Xiong, & Buekens, 2010; Zotti, Williams, Robertson, Horney, & Hsia, 2013). Going into labor early is thought to result from many interconnected variables that can affect the mother and/or fetus ranging from genetic factors to disease and infection to increased stress on the body (Romero et al., 2006). While in some situations, early delivery may be a result of medical intervention, it is important to note that in rural Indonesia in general, antenatal care is viewed as expensive and, thus, most births are attended by nurses, midwives, and traditional birth attendants (Ansariadi & Manderson, 2015). There is an extremely low incidence of birth intervention, induced labor and cesarean section (Blencowe et al., 2013), thus, these are very unlikely to have impacted our results. The almost-five-fold increase in the odds of early or preterm birth for evacuee mothers is substantial and seems to be related to the evacuation of the mother. Given the fact that preterm birth is a risk factor for morbidity and mortality in early childhood, this relationship needs to be explored further in other disaster studies.

In addition to being born early, newborns born to evacuees were shorter, but not lighter, than those born to non-evacuees. This is an unexpected result as the common finding following disaster is for babies to be born significantly lighter if exposed to natural disasters *in utero*. It is possible that because we were reliant on introductions by village nurses and other leaders to

Figure 2 Outcomes (estimates with standard errors) in the evacuee sample by gradients of total mean scores on the stress instrument, the 'Impact of Events Scale – Revised' (Horowitz, Wilner, & Alvarez, 1979; Weiss 2007).

identify potential respondents, especially in the evacuee ('exposed') sample, we may have missed mothers whose babies died (which is consistent with previous work on the Karo [Kushnick, 2010]) and who coincidentally may have also been low birthweight. This could have led to a small number of low birthweight observations, impacting results.

That said, birth length reduction has been reported following ice storms in Canada (Dancause et al., 2011) and a Chilean earthquake (Palmeiro-Silva et al., 2018). It is thought that approximately 50% of variation in birth length is due to genetics (Lunde, Melve, Gjessing, Skjærven, & Irgens, 2007), however we found no difference in maternal height between evacuee and non-evacuee samples, suggesting exposure in utero was having some impact on birth length. We did a preliminary exploration of total mean stress based on administration of the IES-R questionnaire (Horowitz et al., 1979; Weiss, 2007) and found a gradient of decreasing birth length, but not birth weight or proportion born early, with increasing subjective stress (**Figure 2**). This supports the idea that reduced birth length was at least partially due to increasing stress among mothers. This may have future implications for these children as retardation in linear growth can result in developmental and cognitive delays later in life (Grantham-McGregor et al., 2007). It is thus important for future studies to consider birth length in addition to birth weight and gestation in order to better understand how all forms of growth retardation are impacted by maternal stress caused by natural disasters.

Missing data and confounding factors were two potential sources of bias in this study. **Table 1** shows that 8 of 15 variables used in our study had anywhere from 1 to 59 missing data points. Analyzing only cases with complete records can lead to biased results (White et al., 2010). We conducted the multivariate versions of our analyses using both casewise-deleted and imputed (using multiple imputation by chained equations) versions of the data. The results differed to a very small degree quantitatively but were identical qualitatively.

It is important to note that despite efforts to match the evacuee and non-evacuee samples, they differed in several factors that could not have been caused exposed or unexposed status. Evacuees were much more likely to be farmers and chew betel nut, and less likely to be Christian and have at least some high school education. When this is the case, it is misleading for natural disasters to be treated as natural experiments, though they are often touted as such (Auger et al., 2011; King & Laplante, 2015). The argument for treating them as natural experiments is that 'assignment' to the exposed and unexposed groups is random, but that is not always the case. Those living in villages at the base of the Sinabung volcano, which was active in the past, may be a more marginalized and disadvantaged group. The effects of the volcano, thus, may have had a disproportionate effect on a non-random sample of Karo people living in the area. There are many other examples in the literature. Hurricane Katrina, for example, had a disproportionate effect on low socioeconomic-status African American communities (Fussell, Sastry, & Vanlandingham, 2010). The effects of natural disasters are better viewed as observational studies, and researchers should always be concerned about potential confounders as we have here.

As our results show that mothers pregnant during evacuations from volcanic eruptions in Indonesia have adverse birth outcomes including preterm birth and shorter babies. As both factors are associated with negative health consequences later in life, it is important that it is noted in protocols and policies for the evacuation of pregnant women. We are in the process of exploring the association between objective and subjective stress in evacuee mothers in relation to birth outcomes in more detail. This should help to further understand the relationship between evacuation and negative birth outcomes for Indonesian mothers.

Data Availability: Open Science Framework: http://dx.doi.org/10.17605/OSF.IO/47SJG

Acknowledgments: Warm thanks are extended to the participants and the two research assistants, Karmila Karo and Lasma Sinaga. The study was conducted with a research permit issued by the Republic of Indonesia's Ministry of Research, Technology, and Higher Education. Local permission was granted by the relevant authorities, including respective village heads and evacuation center officials. Ethics approval was granted by the Ethics Committee at The Australian National University (ANU). The study was conducted with funds granted by the ANU College of Arts and Social Sciences.

Author Contributions: GK and AB designed and procured funding for the study; FZ facilitated the permit for foreign researchers and provided in-country support; GK conducted the fieldwork and analyses and wrote the first draft of the manuscript; GK, AB, and FZ helped revise the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

Ansariadi, A., & Manderson, L. (2015). Antenatal care and women's birthing decisions in an Indonesian setting: does location matter? *Rural Remote Health, 15*(2), 2959.

Auger, N., Kuehne, E., Goneau, M., & Daniel, M. (2011). Preterm birth during an extreme weather event in Quebec, Canada: a "natural experiment". *Matern Child Health J, 15*(7), 1088-1096.

Bennett, C. M., & Friel, S. (2014). Impacts of Climate Change on Inequities in Child Health. *Children (Basel, Switzerland), 1*(3), 461-473.

Blanc, A. K., & Wardlaw, T. (2005). Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure. *Bulletin of the World Health Organization, 83*, 178-185.

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A.-B., . . . the Born Too Soon Preterm Birth Action, G. (2013). Born Too Soon: The global epidemiology of 15 million preterm births. *Reproductive Health, 10*(1), S2.

Braga, L. H. P., Forrokhyar, F., & Bhandari, M. (2012). Confounding: What it is and how do we deal with it? *Canadian Journal of Surgery, 55*(2), 132-138.

Bruckner, T. A., Catalano, R., & Ahern, J. (2010). Male fetal loss in the U.S. following the terrorist attacks of September 11, 2001. *BMC public health, 10*(1), 273-273.

Casey, R., Rieckhoff, M., & Beebe, S. (1992). The accuracy of maternal report. *Clin Pediatr 31*, 200-5.

Catalano, R., Bruckner, T., Gould, J., Eskenazi, B., & Anderson, E. (2005). Sex ratios in California following the terrorist attacks of September 11, 2001. *Human Reproduction, 20*(5), 1221-1227.

Catalano, R. A., Currier, R. J., & Steinsaltz, D. (2015). Hormonal evidence of selection in utero revisited. *American Journal of Human Biology, 27*(3), 426-431.

Cedergren, M. I. (2007). Optimal Gestational Weight Gain for Body Mass Index Categories. *Obstetrics & Gynecology, 110*(4), 759-764.

Channon, A., Padmadas, S., & McDonald, J. (2011). Measuring birth weight in developing countries: does the method of reporting in retrospective surveys matter? *Matern Child Health J, 15*(1), 12-18.

Dancause, K. N., Laplante, D. P., Oremus, C., Fraser, S., Brunet, A., & King, S. (2011). Disaster-related prenatal maternal stress influences birth outcomes: Project Ice Storm. *Early Human Development, 87*(12), 813-820.

Euser, A. M., Zoccali, C., Jager, K. J., & Dekker, F. W. (2009). Cohort studies: prospective versus retrospective. *Nephron: Clinical Practice, 113*, 214-217.

Fukuda, M., Fukuda, K., Shimizu, T., & Møller, H. (1998). Decline in sex ratio at birth after Kobe earthquake. *Human Reproduction, 13*(8), 2321-2322.

Fussell, E., Sastry, N., & Vanlandingham, M. (2010). Race, socioeconomic status, and return migration to New Orleans after Hurricane Katrina. *Population And Environment, 31*(1-3), 20-42.

Grant, J., & Cardarelli, K. (2014). Evaluation of the Effect of Areca Nut Use During Pregnancy in the Commonwealth of the Northern Mariana Islands. *Obstetrics & Gynecology, 123*, 54S-55S.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., & Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet, 369*, 60-70.

Hanna, R., & Oliva, P. (2016). Implications of climate change for children in developing countries. *The Future of Children, 26*(1), 115-132.

Harville, E., Xiong, X., & Buekens, P. (2010). Disasters and perinatal health: a systematic review. *Obstetrical and gynecological survey, 65*(11), 713-728.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of event scale: a measure of subjective stress. *Psychosomatic Medicine, 41*, 209-218.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (Second Edition ed.). NY: Wiley.

King, S., & Laplante, D. P. (2015). Using natural disasters to study prenatal maternal stress in humans. In M. C. Antonelli (Ed.), *Perinatal programming of neurodevelopment* (pp. 285-313). NY: Springer.

Kipp, R. S. (1993). *Dissociated identities: Ethnicity, religion, and class in an Indonesian Society*. Ann Arbor: University of Michigan Press.

Kushnick, G. (2009). Parental supply and offspring demand amongst karo batak mothers and children. *Journal of Biosocial Science, 14*(2), 183-193.

Kushnick, G. (2010). Resource competition and reproduction in karo batak villages. *Human Nature, 21*(1), 62-81.

Kushnick, G., & Fessler, D. M. T. (2011). Karo Batak cousin marriage, cosocialization, and the Westermarck hypothesis. *Current Anthropology, 52*(3), 443-448.

Kushnick, G., Fessler, D. M. T., & Zuska, F. (2016). Disgust, gender, and social change: Testing alternative explanations for the decline of cousin marriage in Karo society. *Human Nature, 27*(4), 533-555.

Liublinska, V., & Rubin, D. B. (2012). Re: "Dealing With Missing Outcome Data in Randomized Trials and Observational Studies". *Am J Epidem, 176*(4), 357-358.

Lunde, A., Melve, K. K., Gjessing, H. K., Skjærven, R., & Irgens, L. M. (2007). Genetic and Environmental Influences on Birth Weight, Birth Length, Head Circumference, and Gestational Age by Use of Population-based Parent-Offspring Data. *American Journal of Epidemiology, 165*(7), 734-741.

Mace, R. (2000). Evolutionary ecology of human life history. *Animal Behaviour, 59*, 1-10.

McIntire, D. D., Bloom, S. L., Casey, B. M., & Leveno, K. J. (1999). Birth Weight in Relation to Morbidity and Mortality among Newborn Infants. *New England Journal of Medicine, 340*(16), 1234-1238.

Mckinnon, A. (2010). The use and reporting of multiple imputation in medical research -- a review. *Journal of Internal Medicine, 268*, 586-593.

McNamee, R. (2005). Regression modelling and other methods to control confounding. *Occupational and Environmental Medicine, 62*, 500-506.

Navara, K. J. (2010). Programming of offspring sex ratios by maternal stress in humans: assessment of physiological mechanisms using a comparative approach. *J Comp Physiol B, 180*(6), 785-796.

O'Donnell, M. H., & Behie, A. M. (2013). Effects of bushfire stress on birth outcomes: a cohort study of the 2009 Victorian Black Saturday bushfires. *International Journal of Disaster Risk Reduction, 5*, 98-106.

Palmeiro-Silva, Y. K., Orellana, P., Venegas, P., Monteiro, L., Varas-Godoy, M., Norwitz, E., . . . Illanes, S. E. (2018). Effects of earthquake on perinatal outcomes: A Chilean register-based study. *PLoS ONE, 13*(2), e0191340.

Parayiwa, C., & Behie, A. M. (2018). Effects of prenatal maternal stress on birth outcomes following tropical cyclone Yasi in Queensland, Australia (2011). *International Journal of Disaster Risk Reduction, 28*, 768-775.

Pourhoseingholi, M. A., Baghestani, A. R., & Vahedi, M. (2012). How to control confounding effects by statistical analysis. *Gastroenterology and Hepatology: From Bed to Bench, 5*(2), 79-83.

Remler, D., & Van Ryzin, G. (2015). Ch 13: Natural and quasi experiments. In *Research methods in practice: strategies for description and causation* (pp. 427-464). Thousand Oaks, CA: Sage Publications.

Romero, R., Espinoza, J., Kusanovic, J., Gotsch, F., Hassan, S., Erez, O., . . . Mazor, M. (2006). The preterm parturition syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology, 113*(s3), 17-42.

Saadat, M. (2008). Decline in sex rato at birth after Bam (Kerman Province, Southern Iran) earthquake. *Journal of Biosocial Science, 40*(06), 935-937.

Sanguanklin, N., McFarlin, B. L., Park, C. G., Giurgescu, C., Finnegan, L., White-Traut, R., & Engstrom, J. L. (2014). Effects of the 2011 Flood in Thailand on Birth Outcomes and Perceived Social Support. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 43*(4), 435-44.

Shenkin, S. D., Zhang, M. G., Der, G., Mathur, S., Mina, T. H., & Reynolds, R. M. (2017). Validity of recalled v. recorded birth weight: a systematic review and meta-analysis. *Journal of Developmental Origins of Health and Disease, 8*(2), 137-148.

Singarimbun, M. (1975). *Kinship, descent, and alliance among the Karo Batak*. Berkeley: University of California.

Sjölander, A., & Greenland, S. (2013). Ignoring the matching variables in cohort studies - when is it valid and why? *Stat Med, 32*(27), 4696-4708.

Steedly, M. M. (1993). *Hanging without a rope: Narrative experience in colonial and postcolonial Karoland*. Princeton, NJ: Princeton University Press.

Tan, C. E., Li, H. J., Zhang, X. G., Zhang, H., Han, P. Y., An, Q., . . . Wang, M. Q. (2009). The Impact of the Wenchuan Earthquake on Birth Outcomes. *PLoS ONE, 4*(12), e8200.

Tong, V. T., Zotti, M. E., & Hsia, J. (2011). Impact of the Red River Catastrophic Flood on Women Giving Birth in North Dakota, 1994–2000. *Matern Child Health J, 15*(3), 281-288.

Torche, F., & Kleinhaus, K. (2012). Prenatal stress, gestational age and secondary sex ratio: the sex-specific effects of exposure to a natural disaster in early pregnancy. *Human Reproduction, 27*(2), 558-567.

UNICEF, & Organization, W. H. (2004). *Low birthweight: Country, regional and global estimates*. New York: UNICEF.

Weisberg, S. (2005). *Applied linear regression* (Third ed.). Hoboken, NJ: Wiley.

Weiss, D. S. (2007). The impact of event scale: revised. In J. P. Wilson & C. S. Tang (Eds.), *Cross-cultural assessment of psychological trauma and PTSD* (pp. 219-238). NY: Springer.

White, I., Royston, P., & Wood, A. M. (2010). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine, 30*, 377-399.

Xiong, X., Harville, E. W., Buekens, P., Mattison, D. R., Elkind-Hirsch, K., & Pridjian, G. (2008). Exposure to Hurricane Katrina, Post-traumatic Stress Disorder and Birth Outcomes. *The American Journal of the Medical Sciences, 336*(2), 111-115.

Zotti, M. E., Williams, A. M., Robertson, M., Horney, J., & Hsia, J. (2013). Post-disaster reproductive health outcomes. *Matern Child Health J, 17*(5), 783-796.

Supplementary Materials

Pregnancy Outcomes among Evacuees of the Sinabung Volcano, 2010-2018 (North Sumatra, Indonesia): A Matched Cohort Study

Geoff Kushnick [1] ✉, Alison Behie [1], and Fikarwin Zuska [2]

[1] The Australian National University, Canberra, Australia
[2] University of North Sumatra, Medan, Indonesia
✉ *Corresponding Author:* Dr Geoff Kushnick, geoff.kushnick@anu.edu.au

iD 0000-0001-9280-0213 (GK); 0000-0002-7436-3694 (AB); 0000-0001-7642-8546 (FZ)

Table 1 Non-evacuee (n=97) and non-evacuee (n=97) samples matched on child's birth year and mother's age class.

Child's YOB	18-25 y.o.		25-36 y.o.		36+ y.o.		Total		
	Evac.	Non	Evac.	Non	Evac.	Non	Evac.	Non	Tot
2010	0	0	5	5	1	1	6	6	*12*
2011	0	0	6	6	2	2	8	8	*16*
2012	3	3	10	10	2	2	15	15	*30*
2013	4	4	5	5	3	3	12	12	*24*
2014	3	3	6	6	4	4	15	15	*30*
2015	5	5	7	7	3	3	15	15	*30*
2016	4	4	7	7	1	1	12	12	*24*
2017	7	7	4	4	3	3	14	14	*28*
Total	*26*	*26*	*52*	*52*	*19*	*19*	***97***	***97***	***194***

Table S2 Outcomes used in the study and their clinical definitions.

Outcome	Variables Used in Study		Clinical
	Measured	Categorical	
Low birth weight	BW (kg)	Low birth weight (LBW) 0 = 2,500g or more 1 = < 2,500g Low birth weight 2 (LBW2) 0 = > 2,500g 1 = 2,500g or less	Low birth weight (LBW) < 2,500g Very low birth weight (VLBW) < 1,500g Extremely low birth weight (ELBW) < 1,000g
Short birth length	BL (cm)	None	n/a
Low gestational age	None	Early or preterm (EPT) 0 = 9 mo or more 1 = < 9 mo	Full term (39 – 41 weeks) ---------------- 9 months ---------------- Early term (37 – 38 weeks) Moderate to late preterm (32 – 37 weeks) Very preterm (28 – 32 weeks) Extremely preterm (Less than 28 weeks)
Sex	None	Sex 0 = Male 1 = Female	n/a

Table S3 'Impact of Events – Revised' questions (Horowitz, et al., 1979; Weiss, 2007).

Questions
1. Any reminder brought back feelings about it.
2. I had trouble staying asleep.
3. Other things kept making me think about it.
4. I felt irritable or angry.
5. I avoided letting myself get upset when I thought it or was reminded of it.
6. I thought about it when I didn't mean to.
7. I felt as if it hadn't happened or wasn't real.
8. I stayed away from reminders about it.
9. Pictures about it popped into my head.
10. I was jumpy and easily startled.
11 I tried not to think about it.
12. I was aware that I still had a lot of feelings about it, but I didn't deal with them.
13. My feelings about it were kind of numb.
14. I found myself acting or feeling like I was back at that time.
15. I had trouble falling asleep.
16. I had waves of strong feelings about it.
17. I tried to remove it from my memory.
18. I had trouble concentrating.
19. Reminders of it caused me to have physical reactions, such as sweating, trouble breathing, nausea, or a pounding heart.
20. I had dreams about it.
21. I felt watchful and on guard.
22. I tried not to talk about it.

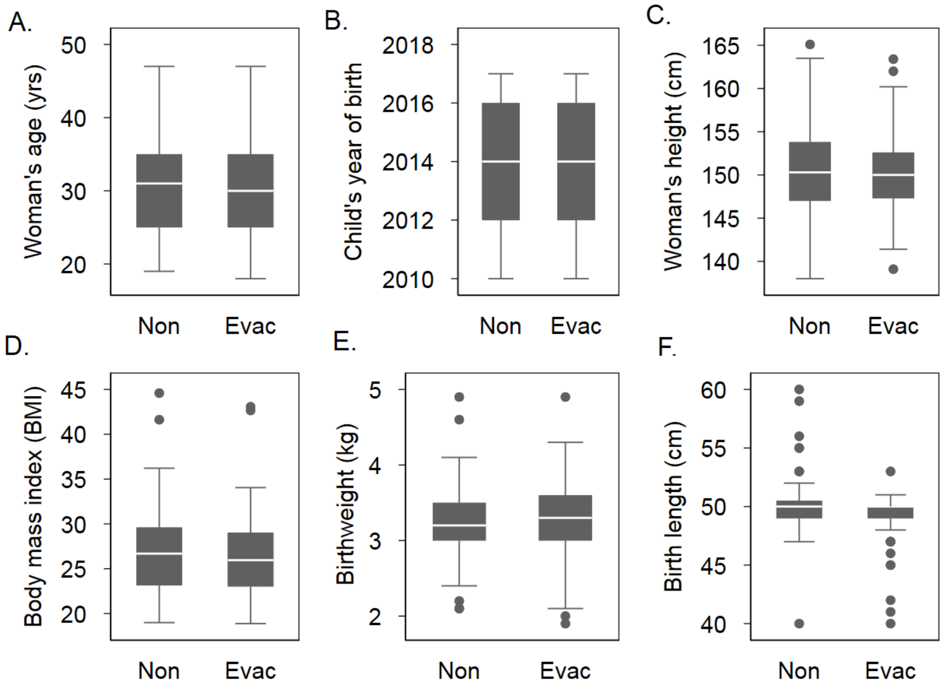

Figure S1 Continuous measures in non-exposed and exposed samples, including potential confounders (woman's age, child's year of birth, woman's height, and woman's BMI); and outcomes (birthweight and birth length).