Supplementary Online Content

Eraly SA, Nievergelt CM, Maihofer AX, et al; for the Marine Resiliency Study Team**.** Assessment of plasma C-reactive protein as a biomarker of posttraumatic stress disorder risk. *JAMA Psychiatry*. Published online February 26, 2014. doi: 10.1001/jamapsychiatry.2013.4374.

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Zero-inflated negative binomial regression

CAPS scores solely comprise non-negative integers, and their distribution in our study was positively skewed but with a large excess of zeros. Therefore, we assessed whether models using a regression procedure specifically designed to account for positively skewed integer-valued distributions with a high incidence of zeros – zero-inflated negative binomial regression (ZINBR) and zero-inflated Poisson regression (ZIPR) [1] – provided a superior fit in comparison to models using negative binomial regression (NBR), Poisson regression (PR), or ordinary linear regression (OLR). These zero-inflated models involve mixture distributions with probability *p* of being a degenerate distribution at 0 (i.e., the value is always 0) and probability 1–*p* of being a negative binomial or Poisson distribution (ZINBR and ZIPR, respectively). In regression using these distributions, the probability *p* of being in the degenerate distribution is assessed using logistic regression (the "zero model"), and the negative binomial or Poisson portion is assessed using a generalized linear model for count data (the "count model").

We found ZINBR to be greatly superior to the other models. For instance, when comparing models with the same outcome (visit 2 CAPS score) and predictor variables (baseline CAPS, cohort, CES score, PBE score, log plasma CRP concentration) as in our final regression model (Table 2), we found that ZINBR provided a vastly superior fit in comparison to each of NBR (*p* value by the Vuong test [2] = 3.78×10^{-43}); ZIPR ($p = 7.46 \times 10^{-69}$); PR ($p = 9.15 \times 10^{-122}$); OLR with log-transformation of CAPS scores ($p = 2.42 \times 10^{-65}$); and OLR without log-transformation ($p = 5.95 \times 10^{-90}$) (in the two latter cases, the Vuong test was modified to allow for comparisons involving linear regression). This superiority of ZINBR was retained at similar significance levels even after compensating for the additional parameters in zero-inflated models by considering the Akaike or Bayesian information criteria (AIC or BIC). Accordingly, we determined the association of our predictors of interest with CAPS using ZINBR.

Zeros in the zero-inflated negative binomial mixture distribution can arise in two ways: from being in the degenerate part, with probability *p*, or from being in the negative binomial part and happening to be 0, with probability (1–p)×p', where p' is the probability of a zero in the negative binomial distribution. However, in our analyses p' but not p was generally negligible, so that p was much greater than (1–p)×p' (i.e., the great majority of zero outcomes were accounted for by the zero model) and 1-p' was much greater than p' (i.e., the great majority of outcomes in the count model were non-zero). Specifically, in the case of our final regression model (Table 2), ~92.01% of zero outcomes in our study population would be expected to arise in the zero model (the median probability of a zero outcome being from the zero model would be 95.14%), and ~99.05% of outcomes in the count model would be expected to be non-zero (the median probability of a non-zero outcome in the count model would be 99.13%). Given these characteristics of the data, for ease of wording the zero model is referred to in this paper as approximately denoting the odds of a zero vs. non-zero outcome, and the count model as approximately denoting the extent of the outcome when it is non-zero.

Measures for variables not included in the final regression model: Comorbid depression and anxiety were assessed by the Beck Depression Inventory (BDI[3]) and the Beck Anxiety Inventory (BAI[4]), respectively (these instruments do not directly yield diagnoses of depression or anxiety but provide numerical scores of the severity of the symptoms of these disorders). Alcohol use was quantified as the sum of the numerical responses to the items related to consumption in the Alcohol Use Disorders Identification Test questionnaire[5-6], and tobacco use by self-report on a four-point Likert scale (0 = none; 1 = non-daily light use; 2 = non-daily heavy use; 3 = daily light use; 4 = daily heavy use). Height, weight, waist circumference, and resting blood pressure were measured at baseline, and body mass index (BMI) was calculated as weight (kg)/height $(m)^2$. Race and ethnicity were obtained by self-report using options defined by the investigators (see Table 1). None of these variables met criteria for inclusion in our regression models as confounders, namely association with both the outcome, post-deployment (visit 2) CAPS, and the predictor of interest, baseline plasma CRP concentration (please see the *Methods* in the main body of the paper for details). Analysis with all available subjects, with only those having visit 2 data, or with only those having data for each of the final model variables yielded identical potential confounders.

eResults. Ordinal logistic regression:

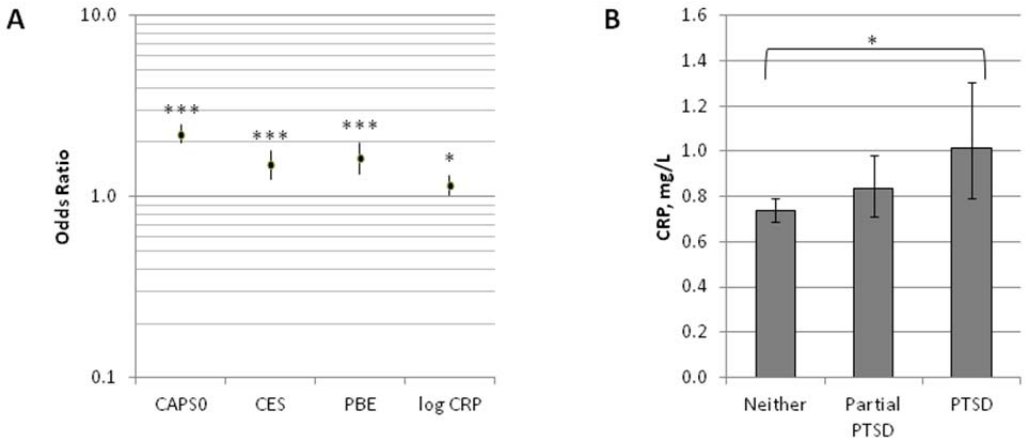
CRP was a significant predictor in ordinal logistic regression of the categorical outcome of PTSD, partial PTSD [7-9], or neither: each 10-fold increment in CRP concentration was associated with an OR of worse PTSD diagnostic category of 1.299 (1.027 – 1.643; p = 0.029) (Table 4). Comparing the predictors, one standard deviation increases in log10 CRP, CAPS0, CES, and PBE were associated with ORs of worse PTSD diagnostic category of 1.156, 2.211, 1.496, and 1.621, respectively (eTable and eFigure A). Adjusted baseline CRP values for subjects with visit 2 PTSD, partial PTSD, or neither were 1.015, 0.836, and 0.740 mg/l, respectively (eFigure B).

eReferences

1. Atkins DC, Gallop RJ. Rethinking how family researchers model infrequent outcomes: a tutorial on count regression and zero-inflated models. *J Fam Psychol.* 2007;21(4):726-735. doi: 710.1037/0893-3200.1021.1034.1726.
2. Vuong QH. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica.* 1989).57:307–333.
3. Beck AT, Steer RA, MG. G. Psychometric properties of the Beck depression inventory: twenty-five years of evaluation. Clin Psychol Rev. *Clin Psychol Rev.* 1998;8:77-100.
4. Beck AT, Epstein N, Brown G, RA S. An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology.* 1988;56:893-897.
5. Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction.* 1993;88(6):791-804.
6. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med.* 1998;158(16):1789-1795.
7. Schnyder U, Moergeli H, Klaghofer R, Buddeberg C. Incidence and prediction of posttraumatic stress disorder symptoms in severely injured accident victims. *Am J Psychiatry.* Apr 2001;158(4):594-599.
8. Pietrzak RH, Goldstein MB, Malley JC, Johnson DC, Southwick SM. Subsyndromal posttraumatic stress disorder is associated with health and psychosocial difficulties in veterans of Operations Enduring Freedom and Iraqi Freedom. *Depress Anxiety.* 2009;26(8):739-744.
9. Adams RE, Boscarino JA, Galea S. Alcohol use, mental health status and psychological well-being 2 years after the World Trade Center attacks in New York City. *Am J Drug Alcohol Abuse.* 2006;32(2):203-224.

eFigure

Ordinal logistic regression of PTSD, partial PTSD, or neither at visit 2. *A*, Adjusted odds ratios of worse PTSD diagnostic category associated with one standard deviation increases in the indicated variables. Data are from the ordinal logistic regression model summarized in eTable. *B*, Baseline plasma C-reactive protein concentration of subjects with PTSD, partial PTSD, or neither at visit 2, adjusted for the same covariates as the regression model in eTable. CAPS0, Clinician Administered PTSD Scale score at visit 0 (baseline); CES, Combat Exposure Scale score; PBE, post-battle experience score; CRP, baseline plasma C-reactive protein concentration. Error bars delineate 95% CIs; ***, p<0.001; *, p<0.05 (all values two-tailed). Please refer to the text for details.



eTable

Ordinal logistic regression of PTSD, partial PTSD, or neither at visit 2. The number of subjects in this model was 1719. OR, odds ratio; CI, confidence interval; CAPS0, visit 0 (baseline) Clinician Administered PTSD Scale score; CES, Combat Exposure Scale score; PBE, post-battle experience score; CRP, baseline plasma C-reactive protein concentration.

Variable	OR[a,b]	95% CI	*P* value
Threshold: No vs. Partial PTSD[b]	22.694	14.782 - 34.840	<0.001
Threshold: Partial PTSD vs. PTSD[b]	115.758	71.531 - 187.331	<0.001
Cohort 1	1.108	0.665 - 1.846	0.145
Cohort 2	0.935	0.602 - 1.45	0.613
Cohort 3	0.564	0.395 - 0.805	<0.001
Cohort 4	0[c]		
CAPS0	1.053	1.045 - 1.061	<0.001
CES	1.036	1.020 - 1.052	<0.001
PBE	1.106	1.061 - 1.154	<0.001
log CRP	1.299	1.027 - 1.643	0.029

[a]Odds ratio of worse PTSD diagnostic category (computed by exponentiating the corresponding coefficient in the regression model and adjusted for the variables listed in the table).

[b]Threshold values indicate the odds of the indicated outcomes at baseline (cohort equals 4 and all other parameters have zero values).

[c]This parameter is set to zero because it is redundant.

