Supporting information for 'What the future held: Early-life psychosocial adversity is associated with health deterioration through adulthood in a cohort of British women'

Daniel Nettle

1. Sources of variables and descriptive statistics

Table S1 gives the original NCDS variable names for sources of the data used in the study, along with descriptive statistics for each variable.

Variable	NCDS variable name(s)	Valid *n*	Descriptives	Notes
HEALTH	n5739 (age 23), n503913 (age 33), hlthgen (age 42)	17766 (Age 23: 6266, Age 33: 5727, Age 42: 5773	Mean 3.18 , s.d. 0.72	Scale inverted to: 1 Poor 2 Fair 3 Good 4 Excellent
CRP	crp	3836	Mean 2.38, s.d. 4.16	
lnCRP	crp	3836	Mean 0.06, s.d. 1.28	Logarithm of CRP
CPA	n95, n183, n222, n658	6814	0: 1793 1: 3136 2: 1493 3: 345 4: 47	Derived as described in Nettle et al. (2011). Missing values of component variables are treated as zero; that is, missing cases are assumed not to have experienced that adversity
SEP	curroscr, n540033, sc, hqual23, hqual33, nd7iamt	8959	Mean 0, s.d. 0. 89	See section S2 for details
AFP	n502023	5785	Mean 25.86, s.d. 5.31	Women with no pregnancy by 33 (22%) given AFP=33.
SMOKE	smoking	5773	1: 2625 2: 1644 3: 1504	Recoded to: 1 Never smoked 2 Occasional or former smoker 3 Regular smoker
BMI	dvht23, dvwt23	6148	Mean 22.12, s.d. 3.25	BMI = dvwt23/dvht23^2

Table S1. Information and descriptive statistics on study variables.

2. Indices of adult socioeconomic position

The index SEP used in the main paper was the first principal component from a principal components analysis using social class at ages 23, 33 and 42 (4-point scale at age 23, 6-point scale at ages 33 and 42), highest educational qualifications at ages 23 and 33 (6-point scale), and household annual gross income at age 42. Scores were saved using the regression method, and missing values for individual variables were replaced with mean. The SEP factor accounted for 49.12% of variance in the measures and haf an eigenvalue of 2.95. The six component variables were correlated with the SEP factors at 0.67, 0.74, 0.68, 0.89, 0.89, and 0.57 respectively.

The SEP variable is a summary over the whole of adulthood. Thus, individuals have the same value of SEP at age 23, age 33 and age 42. We also calculated a separate SEP principal component at each age. These separate measures were based on two variables each (social class and qualifications at ages 23 and 33, social class and income at age 42). Because of missing data and the unbalanced distribution of individual variables, the separate age-specific socioeconomic measures had less normal and smooth distributions, and more individuals assigned a value of zero, than the summary SEP variable. The age-specific measures also correlated strongly with the summary SEP variable (r=0.86, r=0.92, r=0.65 respectively) and moderately with one another (rs 0.41-0.74).

We repeated the analyses of HEALTH in the main paper using the age-specific socioeconomic variables instead of the summary one. Results were qualitatively unchanged and no conclusions were affected. However, the effect size for socioeconomic position itself was attenuated. For example, the parameter estimate for SEP in model 2 of table 1 would be 0.085 (s.e. 0.0063) instead of 0.13 (s.e. 0.0072).

3. Model diagnostics for HEALTH

The models for the HEALTH variable treat it as continuous and employ a Gaussian error structure, although the dependent variable can only take a limited number of values. However, for all models, the distribution of the residuals is acceptably normal. For example, figure S1 plots the residuals for model 4. Other models show a similar pattern. In addition, the variance of the residuals appears to be reasonably homogenous. For model 4, the variance of the model residuals at the different levels of HEALTH are as shown in table S2. Other models show a similar degree of homogeneity.

HEALTH value	Variance of residuals
1	0.227
2	0.241
3	0.233
4	0.248

Table S2. Variance of the residuals of model 4 for health by level of health

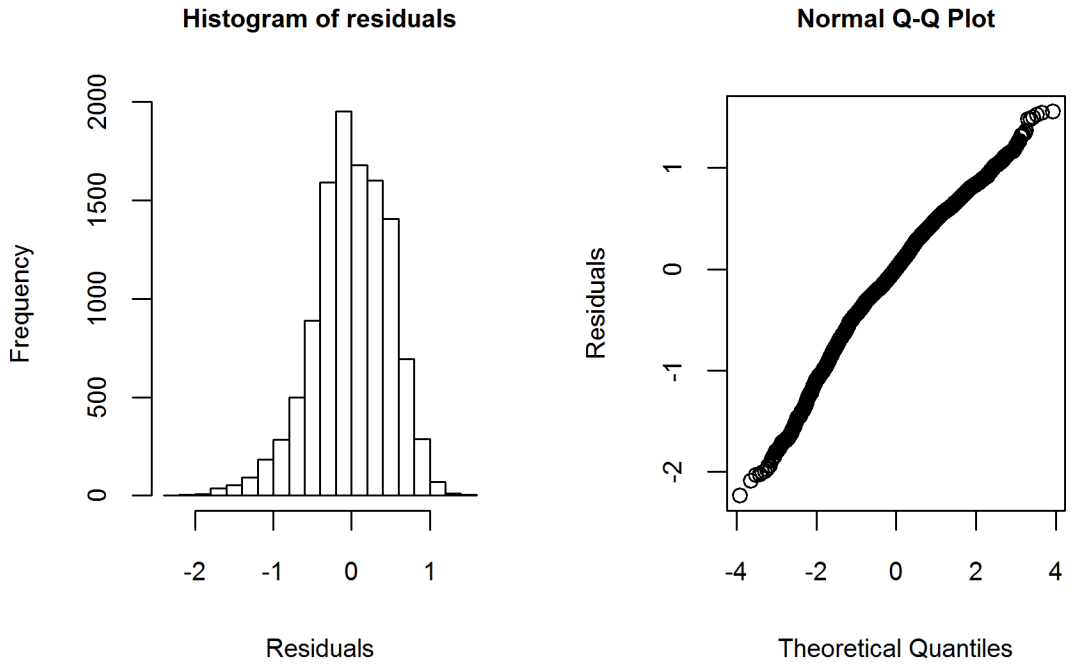

Figure S1. Histogram of residuals and normal Q-Q plot for the residuals from model 4 for HEALTH in the main paper. A straight line in the normal Q-Q plot indicates normally distributed residuals.

References

Nettle, D., Coall, D. A., & Dickins, T. E. (2011). Early-life conditions and age at first pregnancy in British women. *Proceedings of the Royal Society B: Biological Sciences, 278*, 1721-1727. doi: 10.1098/rspb.2010.1726