Supplementary Online Content

Schoenbaum M, Kessler RC, Gilman SE, et al; for the Army STARRS Collaborators. Predictors of suicide and accident death in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). Published online March 3, 2014. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2013.4417

eAppendix. Detailed Methods for Predictive Associations

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

eAppendix. Detailed Methods for Predictive Associations

Discrete-time survival analysis with person-month the unit of analysis[11] was used to study associations of predictors with suicides and accident deaths. The general model form was:

$$Y_t = g(\alpha X_t + \beta Z_{s<t} + \gamma T + \varepsilon), \varepsilon \sim F(\varepsilon), \tag{1}$$

where Y_t represents a dichotomous outcome (suicide or accident death) in person-month t, X_t represents a vector of characteristics (e.g., age, sex, rank) true for the soldier in month t, some of which can vary over time, $Z_{s<t}$ represents a vector of Army entry characteristics (AFQT score, whether the soldiers entered with an accession waiver); T represents time (person − month); α, β, and γ represent parameters to be estimated, and ε represents the error term. In these analyses, $F(\varepsilon)$, the error distribution, is the logistic distribution, and $g(\cdot)$ is the indicator function for positive values.

As some potentially important predictors of suicide and accident deaths are highly inter-related (e.g., soldier age, rank, and length of Army service; education and rank), the analysis of predictors is based on bivariate versions of the above model. That is, we consider only one substantive predictor at a time along with a control for person-month. χ^2 tests were used in each such model to evaluate whether the suicide rate was uniform across the values of the predictor variables for soldiers overall. Each bivariate analysis was also stratified by the three broad deployment status categories of never, currently, and previously deployed. Suicide rates were calculated per 100,000 person-years of active service for soldiers with each value of each predictor. We also evaluated whether the associations of predictors with outcomes varied significantly across deployment status categories

To evaluate the possibility that the rise in the Army suicide rate since 2004 is related to changes in composition of the Army, we estimated the full multivariate version of Equation (1)

separately using data from 2004-2005, 2006-2007, and 2008-2009 and evaluated whether

parameters varied significantly over these different time periods. We then simulated what the

Army suicide and accident death rates would have been in 2006-2007 and 2008-2009 if Army

composition in those years had been the same as in 2004-2005. We did this, using the above

notation, by letting $\hat{Y}_i = g\big(\hat{\alpha}_{[2004-2005]}X_i + \hat{\beta}_{[2004-2005]}Z_i\big)$ represent the predicted probability

of mortality for the ith person-month in 2004-2005, $\hat{Y}_i = g\big(\hat{\alpha}_{[2006-2007]}X_i + \hat{\beta}_{[2006-2007]}Z_i\big)$ the

predicted probability of mortality for the ith person-month in 2006-2007, and so on. For the

simulation, we then calculated $\hat{Y}_i = g\big(\hat{\alpha}_{[2006-2007]}X_i + \hat{\beta}_{[2006-2007]}Z_i\big)$ and

$\hat{Y}_i = g\big(\hat{\alpha}_{[2008-2009]}X_i + \hat{\beta}_{[2008-2009]}Z_i\big)$, respectively, for each person-month in our 2004-2005

sample, and then averaged the predicted probabilities of suicide death across the respective

simulations. The resulting numbers represent the predicted mortality rates for 2006-2007 and

2008-2009, respectively, under two assumptions: that the distributions of characteristics in those

periods were the same as in 2004-2005, but that the relationships between the covariates and

mortality (i.e., α and β) in a given period remains as estimated for that period (e.g., that 2006-

2007 parameter estimates would continue to apply in that period, even if covariates were at

2004-2005 levels).