Supplemental Material 3.

Statistical Methods

Modelling framework

For each type of system damage examined, the i^{th} study in the set of studies with relevant data, provides data on the number of patients with that type of damage, y_i, the total number of patients in the study, n_i, and the average duration of follow-up, fu_i. Models were developed for the rate of damage, r_i, for the i^{th} study. The estimate of r_i was taken to be:

$$(y_i + 0.5)/([n_i – 0.5y_i] \text{ X } fu_i) = (y_i + 0.5)/ E_i,$$

where E_i denotes "exposure" to risk in the i^{th} cohort. The 0.5 added to each count is to avoid numerical issues with zero counts and an adjustment to the person-years at risk of subtracting half the follow-up time for each patient experiencing damage is used as the times of damage occurrence are not known. The rates are scaled to represent rates in units of *patients with damage per 100 patient years*. For the change in SDI, the rate examined was the rate of increase in the SDI for patients measured as *rate of increase per 100 patient years.*

For each of the items selected as endpoints, the estimate of $z_i = \log(r_i)$ was calculated and taken as the "outcome" variable. The asymptotic variance of this estimate, assuming that the count of patients with damage followed a Poisson distribution, is $1/(r_i E_i)$ and can be estimated by using the estimated value of r_i in the variance formula. It is convenient to denote the variance estimate simply as v_i .

The rate of damage is expected to vary across study cohorts. A simple random effects model for this variation, that assumes there is variation around some common "average" value r, is used. However, this is modelled in terms of the logarithm of rates as

$$z_i = a + d_i + e_i ,$$

where a is a common population mean log rate, the d_is are assumed to have a normal distribution with mean 0 and variance τ^2 and e_i is a normal sampling error with variance ν_I . The variance τ^2 is termed the "absolute heterogeneity" and

represents the variation in the rates of damage across studies. A relative heterogeneity measure, denoted I^2, defined as

$$I^2 = [(Q - df)/Q] \text{ X } 100\%,$$

where $Q = \Sigma_i (z_i-\mu)^2/ \nu_I$ is the Cochrane measure of the variability in the rates of damage across studies (on the log scale), with *df* representing the degrees of freedom for the model being estimated and μ being a weighted average of the study estimates, $\Sigma_i w_i z_i/ \Sigma_i w_i$, with $w_i=1/v_i$. I^2, which is taken to be zero if it is negative, is designed to measure the percentage of total variability across studies due to true heterogeneity and not to sampling variability. This framework can be extended to incorporate explanatory variables into a regression model for the z_is, written as,

$z_i = a + d_i + b_1x_1 + \ldots + b_kx_k + e_i$, where the b coefficients associated with the k explanatory variables are log relative risks. In a meta-analysis the x variables are generally termed “moderators” and the model is termed a meta-regression model. The x variables are study level variables that may explain some of the heterogeneity in the study rates. Note that when moderators are included, the variance τ^2 is generally referred to as “residual heterogeneity” (*i.e.* after adjusting for moderators) in comparison to “absolute heterogeneity” derived from a simple model without moderators. Restricted maximum likelihood (REML) estimation of this regression model is implemented via weighted least squares in the R package, ‘metafor’ as described by Viechtbauer (ref).

Moderators

If a single moderator, say the proportion of patients in a study cohort having used GC, is in the model, the question is whether this variable is significantly related to the rate of damage and, if so, how much of the variability across studies in the rate of damage can be explained by this variable. Because of the limited data available for analysis, results presented in this paper were primarily from univariate meta-regression models.