

## Appendix

For each transformed density measure,  $Y$ , we fitted a regression for the mean as a linear function of age and BMI and the other fitted covariates. For individual  $i$ ,  $E[Y_i] = b_0 + b_{age} \cdot age_i + b_{BMI} \cdot BMI_i + b_1 X_{1i} + \dots + b_n X_{ni}$ , where  $Y_i$  is their observed value;  $E[.]$  represents the expected value; and  $age_i$ ,  $BMI_i$ ,  $X_{1i}$ ,  $\dots$ ,  $X_{ni}$  are their age, BMI, and other fitted covariates; and  $b_0, \dots, b_n$  are the corresponding regression coefficients. We then divided the residuals,  $R_i = Y_i - E[Y_i]$ , by the standard deviation of the residuals,  $SD(R_i)$ , to give  $Y' = R_i / SD(R_i)$ .

The different density measures,  $Y'$ , were fitted independently and then together; the improvement in fit was assessed using the likelihood ratio test. When we fitted two density measures,  $Y_1'$  and  $Y_2'$ , into the same model, we presented the risk estimates in terms of the change in the standard deviation after adjusting also for the other measure, to be consistent with the OPERA concept. Let  $r$  be the correlation between  $Y_1'$  and  $Y_2'$ . Because the standard deviation of  $Y_j'$  adjusted for  $Y_k'$  is  $SD' = [(1-r^2)]^{0.5}$ , for  $j, k = 1, 2$ , when  $Y_j'$  is fitted with  $Y_k'$  we multiplied the  $\log(OR)$  estimate from fitting  $Y_j'$  by  $SD'$  and then exponentiated it to obtain the appropriate OPERA. [For  $r = 0.6, 0.7$ , and  $0.8$ ,  $SD' \sim 0.8, 0.7$ , and  $0.6$ , respectively.]

We also fitted weighted linear combinations,  $Y_p' = pY_1' + (1-p)Y_2'$ , and estimated the optimal weight,  $p$ , and confidence interval using maximum likelihood theory. Since the standard deviation of  $Y_p'$  is  $SD_p' = [1-2p(1-r) + 2p^2(1-r)]^{0.5}$ , we multiplied the  $\log(OR)$  estimate from fitting  $Y_p'$  by  $SD_p'$  and then exponentiated it to give the appropriate OPERA. [If  $p = 0.5$ ,  $SD_p' = (1+r)/2$ .]