Supplementary Materials and Methods

Methodological quality score

The quality score contains 19 items that cover three potential sources of errors in observational studies: selection bias (N=5 items), misclassification (N=11 items), and confounding (N=3 items), which are weighted using a proportion of 2:2:1. The maximum attainable score was 105. The following five items made up a maximal selection bias score of 42: (1) percentage loss to follow-up; percentage response; (2) absolute difference in percentage response between cases and controls; (3) percentage incident cases; (4) cases and controls from the same source population; and (5) same exclusion/inclusion criteria for cases and controls. A maximal misclassification score of 42 consisted of the following 11 items: (6) complete measure of recreational physical activities; (7) assessment of total physical activity; (8) inclusion of intensity, frequency and duration in the measure of physical activity; (9) type of administration of physical activity questionnaire; (10) understandable physical activity score; (11) inclusion of past physical activity in the physical activity measure; (12) consideration of changes over time in physical activity pattern; (13) physical activity questionnaire validated or reliable; (14) assessment of physical activity before hematologic cancer diagnosis; (15) valid cancer diagnosis; and (16) exclusion of benign cases. The following three items covering confounding amounted to a maximal score of 21: (17) statistically correct adjustment for confounders; (18) residual confounding; and (19) mutual adjustment for recreational and occupational physical activity.

Dose-response meta-analysis

As suggested by Rota and colleagues (1), dose-response relations of the form RR = exp ($a_{i1}\ dose^{p_1} + a_{i2}\ dose^{p_2}$) or, equivalently, of the fractional polynomial form

$$\log(\mathrm{RR}) = a_{i1}\ dose^{p_1} + a_{i2}\ dose^{p_2} \qquad \text{(Eq. 1)}$$

were fitted to the data of each individual study $i, i = 1, \ldots, n$. The choice of p_1 and p_2 was restricted to positive values because other choices would have violated the condition RR=1 at the reference level $dose = 0$ so that all possible combinations of $p_1, p_2 \in \{0.5, 1, 2, 3\}$ were considered. If $p_1 \neq p_2$, the fractional polynomial had degree 2, and if $p_1 = p_2$, the fractional polynomial had degree 1 for it then simplified to $\log(\mathrm{RR}) = a_{i1}\ dose^{p_1}$.

Following the model suggested by Rota and colleagues (1), it was assumed that, for a given choice of p_1 and p_2, the parameter vector $a_i = (a_{i1}, a_{i2})^T$ defining the dose-response relationship (Eq. 1) for study i came from the multivariate random-effects model

$$\alpha_i \sim N(\alpha, \Psi)\,, \qquad \text{(Eq. 2)}$$

$$a_i | \alpha_i, S_i \sim N(\alpha_i, S_i)\,, \qquad \text{(Eq. 3)}$$

under which the covariance matrix S_i was supposed to be known. Because a fractional polynomial of degree m can only be fitted if study i reported relative risks for at least m different doses (not counting the reference dose), it was additionally assumed that the number of different doses $dose_{ij},\ j = 1, \ldots, k_i$, examined in study i was greater than or equal to the degree m of the fractional polynomial. When $p_1 \neq p_2,$ the $k_i \times 2$ design matrix

$$D_i = \begin{pmatrix} {dose_{i1}}^{p_1} & {dose_{i1}}^{p_2} \\ \vdots & \vdots \\ {dose_{ik_i}}^{p_1} & {dose_{ik_i}}^{p_2} \end{pmatrix}$$

of rank 2 defined the affine transformation $D_i a_i$, a $k_i \times 1$ matrix. That transformation determined the k_i observed log relative risks

$$\log\left(RR_{ij}\right) = a_{i1}\ {dose_{ij}}^{p_1} + a_{i2}\ {dose_{ij}}^{p_2} = (D_i a_i)_j$$

from study i and the corresponding $k_i \times k_i$ covariance matrix $D_i S_i {D_i}^T$. Similarly, when $p_1 = p_2$, the $k_i \times 1$ design matrix $D_i = ({dose_{ij}}^{p_1})_{j=1}^{k_i}$, of rank 1 defined the $k_i \times 1$ matrix $D_i a_i$ representing the observed log relative risks $\log(RR_{ij}) = \mathrm{a_{i1}}\ {dose_{ij}}^{p_1} = (D_i a_i)_j$ and the corresponding $k_i \times k_i$ covariance matrix $D_i S_i {D_i}^T$. Because the full covariance matrices $D_i S_i {D_i}^T$ tended not to be reported in individual studies, Rota and colleagues (1) recommended reconstructing the covariance matrix $D_i S_i {D_i}^T$ for each study i using Greenland and Longnecker's well-known algorithm (2). Similarly, a_i and S_i from (Eq. 3) tended not to be reported. Rota and colleagues (1) advised to apply the transformation $[{D_i}^T (D_i S_i {D_i}^T)^{-1} D_i]^{-1}$ ${D_i}^T (D_i S_i {D_i}^T)^{-1} D_i a_i$ to obtain a_i and the transformation $[{D_i}^T (D_i S_i {D_i}^T)^{-1} D_i]^{-1}$ to obtain S_i. Unfortunately, in theory and practice, the $k_i \times k_i$ matrix $D_i S_i {D_i}^T$ may be singular. For instance, for $k_i > m$, defining S_i to be equal to the identity matrix would yield a $k_i \times k_i$ matrix $D_i S_i {D_i}^T = D_i {D_i}^T$ of rank $m < k_i$, in which case that matrix would be singular and not invertible so that the transformations suggested by Rota and colleagues (1) would not exist and therefore not be of any help in identifying a_i and S_i.

To overcome that problem, the novel transformations

$$({D_i}^T D_i)^{-1} {D_i}^T D_i a_i = a_i \qquad \text{(Eq. 4)}$$

$$({D_i}^T D_i)^{-1} {D_i}^T D_i S_i {D_i}^T D_i ({D_i}^T D_i)^{-1} = S_i \qquad \text{(Eq. 5)}$$

were applied to the observed $D_i a_i$, to the reconstructed covariance matrix $D_i S_i {D_i}^T$ and to the design matrix D_i determined by the model choice to identify a_i and S_i under that model choice. Those alternative transformations always exist because the $m \times m$ matrix ${D_i}^T D_i$ is symmetric, of full rank m and hence invertible. After recovering a_i and S_i with the novel method (Eq.s 4 and 5), the multivariate random-effects model (corresponding to Eq.s 2 and 3) was fitted to estimate α and Ψ, using the R package 'mvmeta' (3).

As suggested by Royston and Altman (4), the simple deviance $D = -2 \times \text{log-likelihood}$ was used to identify the best fitting model. A fractional polynomial of degree 1 with $p_1 \neq 1$ was considered a better model choice than the conventional polynomial of degree 1 defined by $\tilde{p}_1 = 1$ if its deviance was less than that of the conventional polynomial, i.e. $D(p_1) < D(\tilde{p}_1)$, and if the difference between the two deviances was statistically significant at the 10% significance level with respect to the χ^2-distribution with 1 degree of freedom, i.e. if $D(\tilde{p}_1) - D(p_1) > \chi^2_{1;0.90}$, a test that would also serve as a test of linearity. Similarly, a fractional polynomial of degree 2 with $p_1 \neq 1$ and $p_2 \neq 2$ was considered a better model choice than the conventional polynomial of degree 2 defined by $\tilde{p}_1 = 1$ and $\tilde{p}_2 = 2$ if its deviance was less than that of the conventional polynomial, i.e. $D(p_1, p_2) < D(\tilde{p}_1, \tilde{p}_2)$, and if the difference between the two deviances was statistically significant at the 10% significance level with respect to the χ^2-distribution with 2 degrees of freedom, i.e. if $D(\tilde{p}_1, \tilde{p}_2) - D(p_1, p_2) > \chi^2_{2;0.90}$. The best fitting polynomial of degree 2 was preferred to the best fitting polynomial of degree 1 if the difference between the two deviances was statistically significant at the 10% significance level with respect to the χ^2-distribution with 2 degrees of freedom, i.e. if $D(p_1) - D(p_1, p_2) > \chi^2_{2;0.90}$.

References

1. Rota M, Bellocco R, Scotti L, Tramacere I, Jenab M, Corrao G, et al. Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. Statistics in medicine. 2010;29:2679-87.

2. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. American journal of epidemiology. 1992;135:1301-9.

3. Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. Statistics in medicine. 2012;31:3821-39.

4. Royston P, Altman DG. Regression Using Fractional Polynomials Of Continuous Covariates - Parsimonious Parametric Modeling. Applied Statistics-Journal of the Royal Statistical Society Series C. 1994;43:429-67.