International Journal of Epidemiology 2013;**42**:849–859
doi:10.1093/ije/dyt077

METHODOLOGY

Within-person variability in calculated risk factors: Comparing the aetiological association of adiposity ratios with risk of coronary heart disease

David Wormser,[1]* Ian R White,[2] Simon G Thompson[1] and Angela M Wood[1]

[1]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK and [2]MRC Biostatistics Unit, Cambridge, UK

*Corresponding author. Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge CB1 8RN, UK. E-mail: david.wormser@phpc.cam.ac.uk

Accepted 8 April 2013

Background Within-person variability in measured values of a risk factor can bias its association with disease. We investigated the extent of regression dilution bias in calculated variables and its implications for comparing the aetiological associations of risk factors.

Methods Using a numerical illustration and repeats from 42 300 individuals (12 cohorts), we estimated regression dilution ratios (RDRs) in calculated risk factors [body-mass index (BMI), waist-to-hip ratio (WHR), and waist-to-height ratio (WHtR)] and in their components (height, weight, waist circumference, and hip circumference), assuming the long-term average exposure to be of interest. Error-corrected hazard ratios (HRs) for risk of coronary heart disease (CHD) were compared across adiposity measures per standard-deviation (SD) change in: (i) baseline and (ii) error-corrected levels.

Results RDRs in calculated risk factors depend strongly on the RDRs, correlation, and comparative distributions of the components of these risk factors. For measures of adiposity, the RDR was lower for WHR [RDR: 0.72 (95% confidence interval 0.65–0.80)] than for either of its components [waist circumference: 0.87 (0.85–0.90); hip circumference: 0.90 (0.86–0.93) or for BMI: 0.96 (0.93–0.98) and WHtR: 0.87 (0.85–0.90)], predominantly because of the stronger correlation and more similar distributions observed between waist circumference and hip circumference than between height and weight or between waist circumference and height. Error-corrected HRs for BMI, waist circumference, WHR, and WHtR, were respectively 1.24, 1.30, 1.44, and 1.32 per SD change in baseline levels of these variables, and 1.24, 1.27, 1.35, and 1.30 per SD change in error-corrected levels.

Conclusions The extent of within-person variability relative to between-person variability in calculated risk factors can be considerably larger (or smaller) than in its components. Aetiological associations of risk factors should be compared through the use of error-corrected HRs per SD change in error-corrected levels of these risk factors.

Keywords Regression dilution bias, measurement error, within-person variation, adiposity measures

Introduction

Epidemiological analyses often aim to estimate the association between underlying levels of risk factors and the likelihood of a disease. Because most risk factors are measured with error and are subject to fluctuations within individuals, analyses that use only a single measurement of a risk factor may produce biased estimates of its associations with a disease.[1] This bias can be caused by: (i) errors in technical measurement; (ii) short-term within-person variation; and/or (iii) long-term within-person variation. In this paper, these sources of variability are collectively classed as 'within-person variability'. In regression analyses with only a single risk factor, within-person variability leads to underestimation of the true magnitude of the association between long-term average levels of the risk factor and disease (regression dilution bias),[2–8] whereas in analyses with multiple error-prone risk factors the associations may be biased either toward under- or toward overestimation.[9] Various methods have been proposed to quantify and correct the effect of within-person variability in the associations estimated from a single measurement of a risk factor.[1,10]

Although within-person variability in directly measured risk factors (e.g. blood pressure[11] or fibrinogen[12]) has been extensively studied, less is known about within-person variability in calculated risk factors, such as sums and differences (e.g. change in body weight) or ratios[13] [e.g. body-mass index (BMI = weight/height^2) or waist-to-hip ratio (WHR)] of measured variables. Indeed, the extent of within-person variability in calculated risk factors can often appear to be greater or smaller than expected in comparison to the within-person variability in its component measurements.[13] This is important in direct comparisons of the strength of disease-associations of risk factors, including calculated risk factors, that have different degrees of within-person variability. For instance, there has been considerable interest in comparing the magnitude of associations of various cardiovascular risk factors, such as different types of lipid markers or measures of adiposity, with the risk of coronary heart disease (CHD).[14,15] Such comparisons are straightforward when the effect of within-person variability is ignored (i.e. in analyses using measured values). Assuming log-linear relationships of risk factors with CHD risk, their associations are often compared per standard-deviation (SD) change in the baseline levels of these risk factors. Because of different degrees of within-person variability in some risk factors, however, the interpretation of these findings becomes more complicated when their associations are also corrected for regression dilution bias, and the use of baseline SD as the unit with which to compare associations of risk factors with the disease may be inappropriate.

The current paper has two objectives. Its primary objective is to illustrate the extent of within-person variability in calculated risk factors through a numerical example and the use of data on measures of adiposity from 12 different studies in the Emerging Risk Factors Collaboration.[16] The secondary objective of the paper is to demonstrate how to compare the magnitudes of disease-association of risk factors that have different degrees of within-person variability, with regression dilution bias taken into account.

Numerical example

Regression dilution ratios in calculated variables

The extent of within-person variability in a risk factor can be quantified through the regression dilution ratio (RDR), which is appealing because of its simplicity and its familiarity in the epidemiological literature.[7,11,12] The RDR can be defined as the ratio of the between-person variance to the total variance (i.e. between-person variance + within-person variance) in a risk factor. Values of the RDR that are close to 1 indicate little within-person variability, whereas values closer to 0 imply greater levels of within-person variability.

For illustration and algebraic simplicity, suppose we are interested in the relationship between risk and $T_2 - T_1$, where the components T_1 and T_2 represent two error-free variables with a bivariate normal distribution (*BVN*). For example, T_1 and T_2 may be underlying body weight on two occasions or waist and hip circumference at the same point in time. Hence, the exposures of interest are the change in weight or the difference in waist and hip circumference, respectively. Let Q_{1i} and Q_{2i} represent the observed variables measured with error for individual i. It is assumed that the classical additive measurement error model[17] applies to T_1 and T_2:

$$Q_{1i} = T_{1i} + e_{1i}$$
$$Q_{2i} = T_{2i} + e_{2i}$$
$$\text{where} \begin{bmatrix} T_{1i} \\ T_{2i} \end{bmatrix} \sim BVN\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\right)$$
$$\text{and} \begin{bmatrix} e_{1i} \\ e_{2i} \end{bmatrix} \sim BVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \upsilon_1^2 & \tau\upsilon_1\upsilon_2 \\ \tau\upsilon_1\upsilon_2 & \upsilon_2^2 \end{bmatrix}\right).$$

The within- and between-person variances for Q_{1i} are υ_1^2 and σ_1^2 respectively, and likewise for Q_{2i}.

ρ is the correlation between the error-free values T_1 and T_2, whereas τ is the correlation between the within-person errors in these values. The RDRs for Q_1 and Q_2 are therefore:

$$\text{RDR}(Q_1) = \frac{\sigma_1^2}{\sigma_1^2 + \upsilon_1^2} \text{ and } \text{RDR}(Q_2) = \frac{\sigma_2^2}{\sigma_2^2 + \upsilon_2^2}.$$

The within- and between-person variances, and the RDR for the observed difference $Q_2 - Q_1$ are:

$$\text{within-person variance } (Q_2 - Q_1) = \upsilon_1^2 + \upsilon_2^2 - 2\tau\upsilon_1\upsilon_2$$

$$\text{between-person variance } (Q_2 - Q_1) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$$

$$\text{RDR}(Q_2 - Q_1) = \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) + (\upsilon_1^2 + \upsilon_2^2 - 2\tau\upsilon_1\upsilon_2)}.$$

Consider the value of the between-person variance. If T_1 and T_2 are similarly distributed with equal variances (e.g. σ^2), then the between-person variance for $Q_2 - Q_1$ is simply $2\sigma^2 - 2\rho\sigma^2$, which approaches zero as the correlation, ρ approaches 1. For $\tau < \rho$ the between-person variance will be relatively greater than the between-person variance for $Q_2 - Q_1$, which will yield a smaller RDR than the RDR that would result from the use of Q_1 and Q_2 alone, as illustrated below.

The algebraic forms for the within- and between-person variances for the summation $T_1 + T_2$ can be derived in exactly the same way as shown above. For ratios of variables, such as T_2/T_1, it is easier to consider the log-transformed equivalent, $\log T_2 - \log T_1$, to which the equations given above can be applied. Usually, the RDR of a log-transformed ratio approximates the RDR of the untransformed ratio,[12] and variances of log-transformed variables are similar to the squared coefficients of variation of the untransformed variables.

Empirical results

Figure 1 shows the calculated RDRs for $Q_1 + Q_2$ and $Q_2 - Q_1$ in the situations in which (i) $\text{RDR}(Q_1) = \text{RDR}(Q_2)$ has the values 0.95, 0.80, and 0.60; (ii) ρ varies from –1 to 1; and (iii) the ratios of the between-person variances σ_1^2 and σ_2^2 (as well as the ratios of the within-person variances υ_1^2 and υ_2^2) have the values 1, 0.75, and 0.5. We first assume that there is no correlation between the within-person errors (i.e. $\tau = 0$). Under these conditions, $\text{RDR}(Q_2 - Q_1)$ (Figure 1, dotted line) decreases with a higher correlation, ρ, because of the decrease in the between-person variance and the relative increase in the within-person variance. The value of $\text{RDR}(Q_2 - Q_1)$ is smaller than that of $\text{RDR}(Q_1)$ [= $\text{RDR}(Q_2)$] for all positive correlations, ρ. Depending on the RDRs of Q_1 and Q_2, the decrease in $\text{RDR}(Q_2 - Q_1)$ can occur mainly at high correlations, ρ, or can also extend over lower correlations. For instance, for $\text{RDR}(Q_1) = \text{RDR}(Q_2) = 0.95$, $\rho > 0.8$ leads to a sudden decrease in $\text{RDR}(Q_2 - Q_1)$, whereas for lower RDRs for Q_1 and Q_2, the value of $\text{RDR}(Q_2 - Q_1)$ decreases earlier and less remarkably. Greater discrepancy in the variances of T_1 and T_2 attenuates that effect and prevents $\text{RDR}(Q_2 - Q_1)$ from decreasing beyond a certain limit. A similar but reversed situation is observed for $\text{RDR}(Q_1 + Q_2)$ (Figure 1, dashed line).

Figure 2 is a plot of the calculated RDRs for $Q_1 + Q_2$ and $Q_2 - Q_1$ under the situations in which: (i) $\text{RDR}(Q_1) = \text{RDR}(Q_2)$ has the values 0.95, 0.80, and 0.60; (ii) ρ varies from –1 to 1; and (iii) τ has the values 0, 0.3, and 0.6. We now assume equal variances, $\sigma_1^2 = \sigma_2^2$. Note that the plot lines cross at $\rho = \tau$. The value of $\text{RDR}(Q_2 - Q_1)$ declines with higher correlation, ρ. However, $\text{RDR}(Q_2 - Q_1)$ becomes more stable with increasing τ, except at very high values of ρ. Similar results are observed when the ratios of the between-person variances σ_1^2 and σ_2^2 are equal to 0.75 and 0.5, respectively (Supplementary Figure 1, available as Supplementary data at *IJE* online).

Comparison of Measures of General and Abdominal Adiposity

The Emerging Risk Factors Collaboration

The Emerging Risk Factors Collaboration (ERFC)[16] collected baseline and repeat information on height, weight, and waist and hip circumference from 12 prospective studies, in which there were 3351 fatal or first-ever non-fatal CHD events. Among 58 271 individuals in the 12 studies for whom baseline measures were available, a total of 42 300 had one or more repeat measurements, and 21 360 of these individuals, in 4 studies, had ⩾ 2 repeat measurements (Table 1). The 79 145 available repeat measurements available were derived from 18 different re-surveys spanning the interval from 2–10 years after the baseline survey. Individuals with repeat measurements of adiposity-related measures generally had somewhat higher baseline values of measures of adiposity, and were younger and more likely to be non-smokers than individuals in the same studies who did not have repeat measurements.

Statistical methods

We quantified the extent of within-person variability in WHR, WHtR, and BMI and their components by the RDR, using Rosner's regression approach.[10] Separate RDRs for each re-survey in each study were estimated by regressing the repeat measurements of the adiposity measures on their baseline values. For each study $s = 1,2,\ldots,S$, with individuals $i = 1,2,\ldots,n_s$, and repeat measurements $r = 1,2,\ldots,r_{si}$, the model can be written as

$$E_{sir} = \alpha_{sr} + \beta_{sr}E_{si} + \varepsilon_{sir},$$

where $\varepsilon_{sir} \sim N(0,\sigma_{sr}^2)$ and β_{sr} is the study and re-survey–specific RDR. E_{sir} and E_{si} represent repeat and baseline measurements of the adiposity measure E,

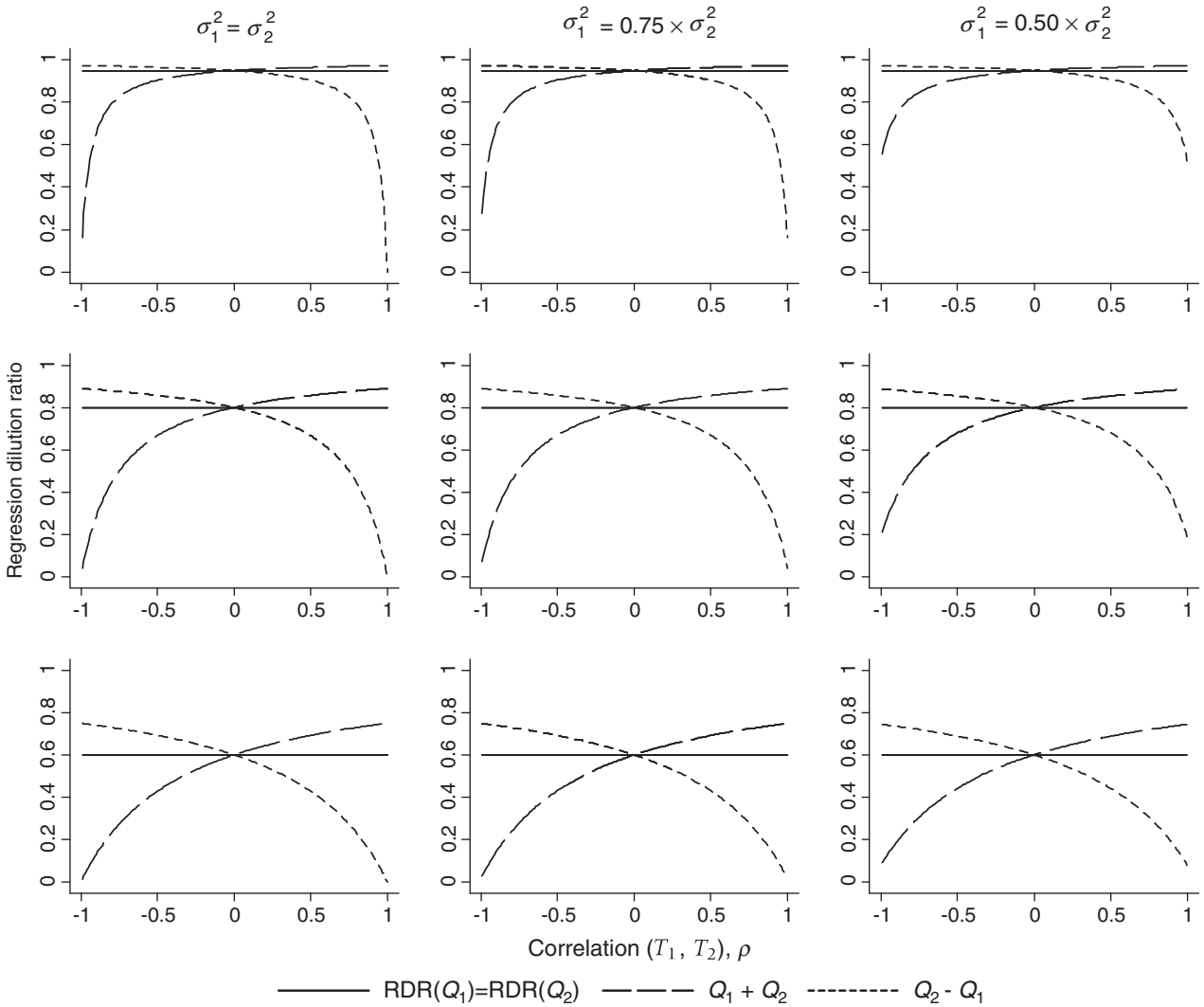

Figure 1 Changes in values of regression dilution ratios (RDRs) according to correlation, ρ, and different ratios of comparative between-person variances of components
$\mathrm{RDR}(Q_1) = \mathrm{RDR}(Q_2) = 0.95$ (top row), = 0.8 (middle row), = 0.6 (bottom row) (solid black lines). Assumption: $\tau = 0$.
Note: σ_1^2/σ_2^2 and σ_2^2/σ_1^2 produce the same results

respectively. α_{sr} represents the study and re-survey–specific intercept.

Overall RDRs were estimated from a Rosner-regression calibration model described by a single linear mixed model of the repeat measurement on the baseline measurement, adjusted for study and re-survey (to allow for any differences in mean levels between studies and at different re-surveys), and with allowance for between-study heterogeneity in the RDRs and between-person variability in mean levels (to account for multiple repeat measurements per person), represented by the model:

$$E_{sir} = \alpha_{sr} + (\beta + u_s)E_{si} + \gamma X_{si} + w_{si} + \varepsilon_{sir}$$

where, $u_s \sim N(0, \sigma_u^2)$, $w_{si} \sim N(0, \sigma_w^2)$ and $\varepsilon_{sir} \sim N(0, \sigma_e^2)$; X_{si} are other baseline covariates, such as age, sex, and smoking status. Between-study heterogeneity in the RDR value β is represented by σ_u^2. The parameters σ_w^2 and σ_e^2 represent individual-specific and residual variation, respectively. Between- and within-person variances and correlations of two log-transformed adiposity measures were estimated from a bivariate linear mixed model, using all-logarithmic baseline and re-survey measurement values as the dependent variables, regressed on dummy variables for study and for re-surveys, with allowance for between-person variability in mean levels.

Proportional hazards Cox models were used to calculate hazard ratios (HRs) in relation to WHR, WHtR, BMI, and waist circumference for fatal or first-ever non-fatal CHD, adjusted for age, sex, and smoking status. Analyses involved a two-stage approach with estimates of association calculated separately within each study before pooling across studies by

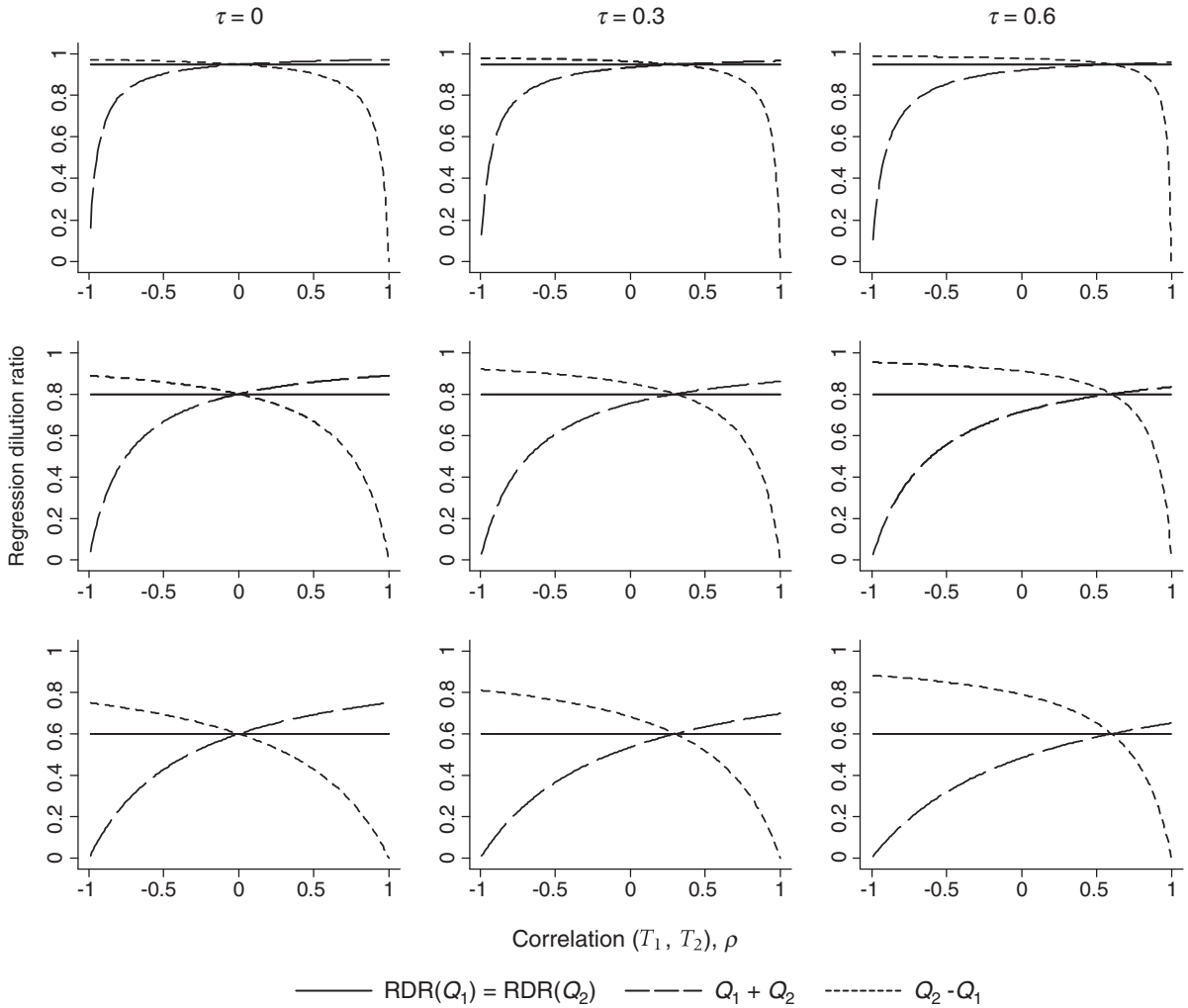

Figure 2 Changes in values of regression dilution ratios (RDRs) according to correlation, ρ, and different correlations of within-person errors of components, RDR(Q_1) = RDR(Q_2) = 0.95 (top row), = 0.8 (middle row), = 0.6 (bottom row) (solid black lines). Assumption: $\sigma_1^2 = \sigma_2^2$

random-effects meta-analysis.[18] The effect of between-study heterogeneity was quantified with the I^2 statistic.[19] To investigate the effect on these associations of different degrees of within-person variability in measures of adiposity, HRs were calculated on the error-corrected levels and presented: (i) per 1-SD change in baseline adiposity measures (SD of baseline measures for all individuals); and (ii) per 1-SD change in error-corrected levels (SD of predicted error-corrected levels for all individuals). Predicted error-corrected levels of measures of adiposity for all individuals were estimated as empirical Bayes estimates from the Rosner-regression calibration model described above.[12,20] Thus, the predicted error-corrected levels were estimated as

$$\hat{E}_{si} = \hat{\alpha}_s + (\hat{\beta} + \hat{u}_s)E_{si} + \hat{\gamma}X_{si} + \hat{w}_{si}$$

where $\hat{\alpha}_s = \frac{1}{r_s}\sum_{r=1}^{r_s}\hat{\alpha}_{sr}$ and $\hat{u}_s$ and $\hat{w}_{si}$ are the best linear unbiased predictors of the random effects. Supplementary analyses were done with averaged measures of adiposity for each individual.[21] The Supplementary Appendix (available as Supplementary data at *IJE* online) provides an example of our analyses for BMI.

Regression dilution ratios

Overall RDRs of adiposity measures were: 0.72 (95% confidence interval 0.65–0.80) for WHR, 0.87 (0.85–0.90) for WHtR, and 0.96 (0.93–0.98) for BMI (Figure 3). Corresponding RDRs for components of these ratios were: 0.87 (0.85–0.90) for waist circumference, 0.90 (0.86–0.93) for hip circumference, 0.99 (0.98–1.00) for height, and 0.97 (0.96–0.98) for weight. There was considerable heterogeneity between the study- and resurvey–specific RDRs of WHR, with RDRs ranging from 0.48 to 0.87.

Correlations and comparative distributions of components of adiposity measures

Correlations and ratios of the between-person variances for log-transformed height, weight, waist, and

Table 1 Characteristics of studies and individuals with serial measurement of adiposity measures

		Among individuals with at least one repeat											
						Baseline mean (SD)							
Study	No of individuals with baseline values	No of individuals	No of re-surveys	No of individuals with ⩾ 2 repeats	Male (%)	Age (yrs)	WC (cm)	HC (cm)	Height (cm)	Weight (kg)	WHR	WHtR	BMI (kg/m^2)
ARIC	14 383	13 414	3	12 065	43	54 (6)	97 (14)	105 (10)	168 (9)	78 (17)	0.92 (0.08)	0.57 (0.08)	28 (5)
AUSDIAB	9204	5280	1	–	44	53 (11)	91 (14)	105 (10)	169 (9)	77 (16)	0.86 (0.09)	0.54 (0.08)	27 (5)
CHS	3881	3265	1	–	38	72 (5)	93 (13)	101 (9)	164 (9)	72 (14)	0.92 (0.09)	0.57 (0.08)	26 (4)
COPEN	8166	4332	1	–	42	54 (13)	86 (13)	99 (8)	169 (10)	73 (14)	0.86 (0.09)	0.51 (0.07)	25 (4)
EPICNOR	1417	792	1	–	67	65 (8)	92 (11)	103 (7)	168 (9)	75 (12)	0.90 (0.08)	0.55 (0.06)	27 (3)
HOORN	2226	1359	1	–	45	60 (7)	90 (10)	102 (6)	169 (9)	75 (11)	0.88 (0.09)	0.53 (0.06)	26 (3)
IKNS	1942	86	1	–	83	63 (8)	83 (8)	91 (6)	158 (7)	59 (8)	0.92 (0.06)	0.53 (0.06)	23 (3)
LASA	1806	1124	2	707	44	70 (7)	97 (11)	103 (8)	167 (9)	75 (12)	0.94 (0.08)	0.58 (0.07)	27 (4)
MESA	6768	6373	3	6002	48	62 (10)	98 (14)	106 (11)	167 (10)	79 (17)	0.93 (0.08)	0.59 (0.09)	28 (5)
RANCHO	1784	882	1	–	40	65 (10)	84 (12)	100 (8)	167 (10)	70 (14)	0.83 (0.09)	0.50 (0.06)	25 (4)
SHS	4135	3482	2	2586	37	56 (8)	106 (15)	111 (13)	165 (9)	84 (18)	0.95 (0.06)	0.64 (0.10)	31 (6)
TARFS	2559	1911	1	–	50	47 (11)	93 (12)	105 (10)	163 (9)	74 (13)	0.89 (0.09)	0.57 (0.08)	28 (5)
Overall	**58 271**	**42 300**		**21 360**	**44**	**61 (9)**	**92 (13)**	**103 (10)**	**166 (9)**	**74 (16)**	**0.90 (0.08)**	**0.56 (0.08)**	**27 (5.0)**

Study abbreviations: ARIC, Atherosclerosis Risk in Communities Study; AUSDIAB, Australian Diabetes, Obesity and Lifestyle Study; CHS, Cardiovascular Health Study; COPEN, Copenhagen City Heart Study; EPICNOR, European Prospective Investigation of Cancer Norfolk Study; HOORN, Hoorn Study; IKNS, Ikawa, Kyowa, and Noichi Study; LASA, Longitudinal Aging Study Amsterdam; MESA, Multi-Ethnic Study of Atherosclerosis; RANCHO, Rancho Bernardo Study; SHS, Strong Heart Study; TARFS, Turkish Adult Risk Factor Study (Study references are provided in ERFC, *Lancet* 2011[15]).

Abbreviations: BMI, body-mass index; HC, hip circumference; SD, standard deviation; WC, waist circumference; WHR, waist-to-hip ratio; WHtR, waist-to-height ratio.

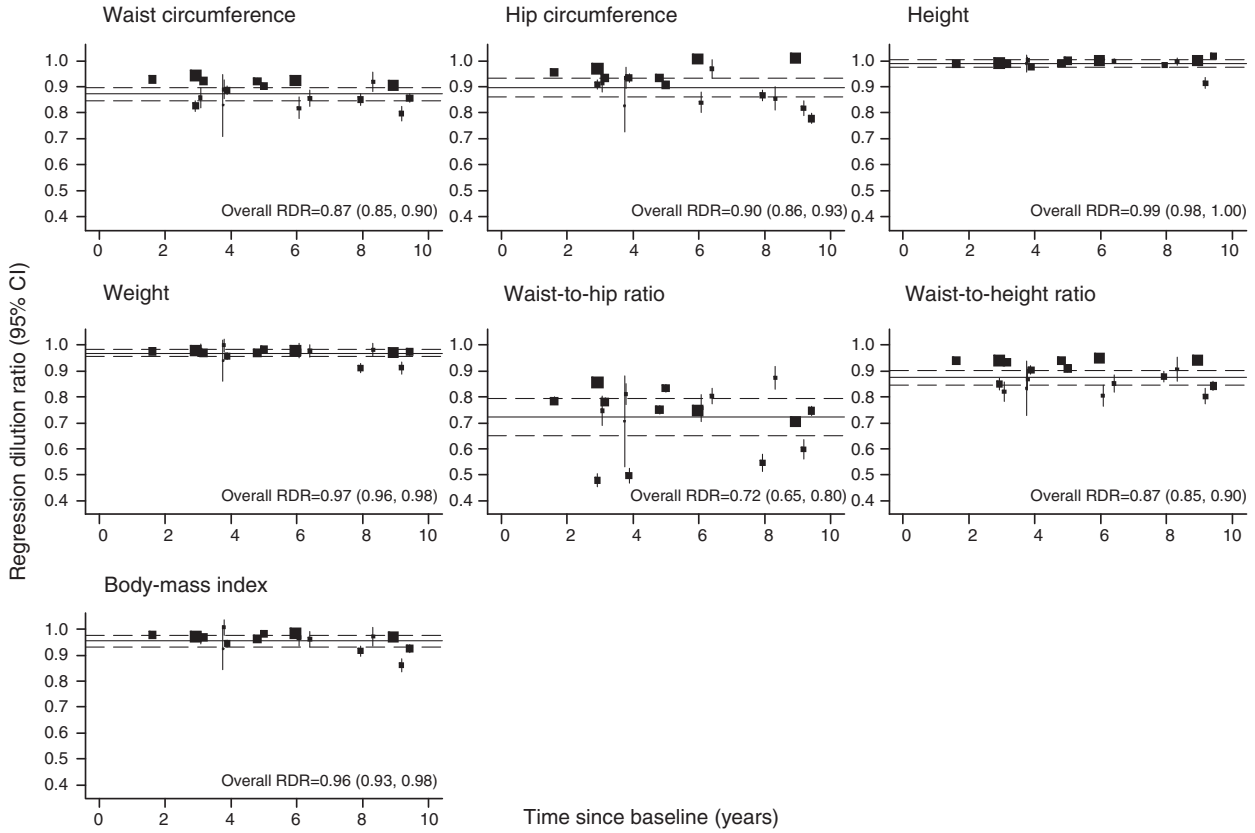

Figure 3 Regression dilution ratios of adiposity measures and anthropometric indicators plotted against time since baseline measurement by study
Regression dilution ratios (RDRs) were unadjusted for covariates. The sizes of data markers are proportional to the inverse of the variance of the RDR. CI indicates confidence interval

hip circumference measurements are shown in Table 2. Underlying waist and hip circumference were more strongly correlated with one another than were either waist circumference and height or weight and height. Additionally, the errors for waist and hip circumference were strongly correlated, which is likely to be due to similar variation in these two measures over time.

Associations corrected for within-person variability

Without correction for regression dilution bias, the HRs of CHD per 1-SD change in baseline levels were similar for BMI, waist circumference, WHR and WHtR (Table 3). However, in analyses that corrected for regression dilution bias and calculated HRs per 1-SD change in baseline adiposity measure, CHD was distinctly more strongly associated with WHR than with WHtR, BMI, or waist circumference. Broadly similar magnitudes of associations were observed in analyses that corrected for regression dilution bias and calculated HRs per 1-SD change in error-corrected levels of adiposity levels. The between-study heterogeneity of HRs decreased somewhat after correction for between-person variability (Table 3). Error-corrected HRs were consistently smaller in analyses that used averaged levels of adiposity measures (Supplementary Table 1, available as Supplementary data at *IJE* online), primarily because this approach fails to correct for regression dilution bias in individuals without repeat measurements.

Discussion

This paper explores the extent of within-person variability in calculated risk factors and its implication for epidemiological studies. Our findings show that RDRs in calculated variables can be considerably different from those in the directly measured risk factors. This difference depends most strongly on the strength of correlations and similarity in the distributions of the directly measured risk factors. Our data on repeat measurements of adiposity showed that the overall RDR of WHR is considerably lower than that of its components or of BMI and WHtR.

The main explanation for the lower RDR of WHR is that overall waist and hip circumference are more strongly correlated with one another and have more similar between-person coefficients of variation than do height and weight or waist circumference and

Table 2 Within- and between-person correlations (95% CI) and variances of components of body-mass index, waist-to-hip ratio, and waist-to-height ratio

T_2/T_1	Between-person correlation of log T_1 and log T_2, ρ	Correlation of within-person errors, τ	Between-person variance(log T_2), σ_2^2	Between-person variance(log T_1), σ_1^2	Ratio of between-person variances, σ_1^2/σ_2^2
Body-mass index (weight/height^2)	0.524 (0.516, 0.531)	0.050 (0.043, 0.057)	0.0394	0.0126	0.32
Waist-to-hip ratio	0.810 (0.806, 0.813)	0.574 (0.569, 0.578)	0.0172	0.0081	0.47
Waist-to-height ratio	0.227 (0.198, 0.217)	-0.084 (-0.091, -0.077)	0.0172	0.0032	0.18

CI indicates confidence interval.

Table 3 Associations of body-mass index, waist circumference, waist-to-hip ratio, and waist-to-height ratio with coronary heart disease risk, with and without correction for within-person variability, per 1-standard deviation (SD) change in baseline or error-corrected levels

	Uncorrected for within-person variability			Corrected for within-person variability					Overall RDR
Coronary heart disease	Per 1-SD in baseline levels	HR (95% CI)	I^2 (95% CI)	Per 1-SD in baseline levels	HR (95% CI)	Per 1-SD in error-corrected levels	HR (95% CI)	I^2 (95% CI)	
Body-mass index	4.95	1.26 (1.17, 1.37)	69 (44, 83)	4.95	1.24 (1.15, 1.34)	4.91	1.24 (1.15, 1.33)	63 (31, 80)	0.96 (0.94, 0.98)
Waist circumference	13.29	1.30 (1.20, 1.40)	64 (33, 81)	13.29	1.30 (1.20, 1.40)	12.41	1.27 (1.18, 1.39)	59 (22, 78)	0.88 (0.86, 0.91)
Waist-to-hip ratio	0.084	1.30 (1.18, 1.44)	79 (65, 88)	0.084	1.44 (1.29, 1.62)	0.067	1.35 (1.23, 1.47)	65 (35, 81)	0.66 (0.59, 0.72)
Waist-to-height ratio	0.080	1.31 (1.21, 1.42)	66 (38, 82)	0.080	1.32 (1.22, 1.43)	0.076	1.30 (1.20, 1.41)	64 (33, 81)	0.88 (0.85, 0.91)

Hazard ratios (HRs) and regression dilution ratios (RDRs) were adjusted for age, sex, and smoking status. Analyses were restricted to participants with BMI $\geqslant 20\,kg/m^2$. Analyses were based on 3351 cases from 12 studies. I^2 is a measure of consistency across studies: the percentage of variance in estimated log HRs that is attributable to between-study variation as opposed to sampling variation. Values of I^2 close to 0 correspond to lack of heterogeneity. CI indicates confidence interval.

height. Hip- and waist-circumference measurements can be made with the same measuring technique (e.g. with a tape measure, and usually by same observer) and have potentially similar variations over time, explaining the strong positive correlations in the within-person errors of hip and waist circumference. In contrast, weight and height are measured with independent standardised techniques (e.g. weighing scales and wall charts, respectively) and height has little variation over time. Our investigation across 12 studies allowed us to observe considerable between-study heterogeneity in the RDRs of WHR, which was largely explained by variation in the correlations and between-person coefficients of variation of waist and hip circumference.

We have shown that HRs for CHD per 1-SD change in error-corrected level of BMI, WHR, WHtR and waist circumference are broadly similar, but quite different conclusions would be drawn had the error-corrected associations been presented per 1-SD of baseline levels of these variables. The objectives of many aetiological studies are to estimate associations between underlying (i.e. error-corrected) levels of risk factors and the likelihood of disease, expressed as risk ratios for some appropriate change in the risk factors. For continuous variables, the unit of change is often chosen as an SD in the observed baseline risk factor, which appropriately allows direct comparisons of: (i) risk associations for several baseline risk factors measured on different scales, uncorrected for within-person variability; and (ii) risk associations for a single risk factor before and after correction for within-person variability. However, we argue that use of an SD in baseline risk factors may be inappropriate for the comparison of different risk factors after correction for within-person variability. Correcting for within-person variability in a single risk factor can be viewed as shrinking the observed distribution of the risk factor to its true error-corrected distribution, and the degree of shrinkage will depend on the extent of within-person variability. Thus, for risk factors with substantial within-person variability, the SD for the error–corrected levels will be much smaller than the SD of the observed baseline levels of these factors. Given the aetiological objectives, it is more appropriate to present the risk ratios per SD change in the error-corrected levels to allow a direct comparison of risk associations between error-corrected levels of several risk factors with different degrees of within-person variability, such as we present for the different adiposity markers in relation to the risk of CHD. These results may resemble the risk associations uncorrected for within-person variability, because the use of smaller unit changes counteracts the effect of correcting for regression dilution bias. However, this similarity is not guaranteed, especially in the case of multivariate regression dilution corrections.

This paper has focused on within-person variability in calculated ratios for measures of adiposity markers, but our numerical examples have implications for other commonly used ratios, such as those for lipids (e.g. the ratio of total to high-density lipoprotein cholesterol), apolipoproteins (e.g. the ratio of apolipoprotein AI to apolipoprotein-B), fatty acids (e.g. the ratio of omega-6 to omega-3 fatty acids), and for simple sums and differences of risk factors (e.g. change in body weight). We have not investigated beyond combining two variables, as is required for determining the calculated serum level of low-density lipoprotein cholesterol,[22] for example, but we expect similar determinants to affect the combined within-person variability of three or more variables.

Among various statistical considerations[13] is that the ratio of two normally distributed variables cannot strictly be normally distributed, violating the assumptions of the additive measurement error model and the parametric Rosner-regression calibration model. However, we observed approximately normal distributions for height, weight, and hip and waist circumference and their corresponding ratios. Depending on the original distribution, a log-transformation can be a useful tool with which to gain a better approximation of normality,[23] although using an RDR or Rosner-regression calibration model for a log-transformed ratio would be appropriate only if the ratio is also log-transformed in the risk-regression model, which may create interpretive difficulties.

We have assumed that disease risk depends on a single, underlying, long-term error-corrected exposure level, and have used repeat measurements made over a long time span. The methodology presented here is also suitable for assessing alternative research hypotheses, such as relating disease risk to the error-corrected exposure level at a point in time (e.g. BMI at age 20 years): in this case, RDRs should be estimated only with the use of repeated measures made over a shorter time span (e.g. BMI repeat measures taken during the 20th year of age).

The limitations of methods to correct for regression dilution bias are well-known.[17] In a more realistic situation with true, time-varying underlying exposure, regression dilution corrections are valid if disease risk depends only on the current true underlying exposure, or if RDRs are constant over time; otherwise, these corrections typically overcorrect.[24] We further assumed a classical, additive non-differential measurement error model, but multiplicative measurement error models that allow within-person variability to increase with the level of exposure may be preferred. However, we observed no important time trend in RDRs over a 10-year time span (Figure 3), nor an increasing within-person variability with level, suggesting that our corrections are likely to be appropriate for adiposity measures. Nevertheless, our observed associations of measures of adiposity with CHD risk may reflect residual bias caused by unmeasured confounders (e.g. dietary intake or physical activity), rather than being causal associations. Corrections for

the extent of within-person variability amplify the effect of such non-causal associations.[25] If these confounding factors were available, they would be measured with error, and correction methods would need to be extended for such analyses, such as by using the *multivariate* Rosner-regression model.[1] As is appropriate, the current study did not adjust for factors (e.g. systolic blood pressure or lipids) on the biological pathways between adiposity and CHD.

Our findings indicate that using calculated variables as aetiologic risk factors can be problematic, but we acknowledge that there may be some practical advantages in using such variables. These include the simplicity in the clinical interpretation of disease-risk associations with a single calculated summary variable, and the applicability of simple correction methods to a single risk factor, whereas multiple error-prone risk factors require more complex multivariate correction methods. However, using any calculated variable in place of its separate components forces constraints on the estimated risk associations, which may not be appropriate or optimal, especially for risk prediction.

Supplementary Data

Supplementary data are availabe at *IJE* online.

Funding

The ERFC Coordinating Centre is underpinned by a programme grant from the British Heart Foundation (RG/08/014) and grants from the UK Medical Research Council and the National Institute of Health, Cambridge Biomedical Research Centre. A variety of sources have supported recruitment, follow-up, and laboratory measurements in the cohorts contributing to the ERFC. Investigators in several of these studies have contributed to a list naming some of these funding sources, which can be found at http://www.phpc.cam.ac.uk/ceu/research/erfc/studies/.

Acknowledgements

We thank Stephen Kaptoge, George Davey Smith, Sarah Lewington, Gary Whitlock, and two reviewers for helpful comments on an earlier draft of this paper.

Conflict of interest: None declared.

Emerging Risk Factors Collaboration Investigators: G Davey Smith; **ARIC** AR Folsom, LE Chambless, J Stevens; **AUSDIAB:** R Atkins, E Barr; **CHS**: KJ Mukamal, JR Kizer (see http://www.chs-nhlbi.org for acknowledgements); **COPEN**: A Tybjærg-Hansen, BG Nordestgaard, R Frikke-Schmidt; **EPICNOR**: K-T Khaw, NJ Wareham; **HOORN**: JM Dekker, G Nijpels, CDA Stehouwer; **IKNS:** H Iso, A Kitamura, K Yamagishi, H Noda; **LASA**: D Deeg JL Poppelaars; **MESA:** M Cushman, AR Folsom, BM Psaty, S Shea (see http://www.mesa-nhlbi.org for acknowledgements); **RANCHO:** E Barrett-Connor, DL Wingard, R Bettencourt; **SHS:** BV Howard, Y Zhang, J Umans; **TARFS:** A Onat.

Data Management Team: M Walker, S Watson.

Coordinating Centre M Alexander, AS Butterworth, E Di Angelantonio, P Gao, P Haycock, S Kaptoge, S Lewington, L Pennells, SG Thompson, M Walker, S Watson, IR White, AM Wood, D Wormser, J Danesh (principal investigator).

KEY MESSAGES

- The extent of within-person variability relative to between-person variability in calculated risk factors can be considerably larger (or smaller) than in its components.
- Regression dilution ratios in calculated risk factors depend strongly on the regression dilution ratios, correlation, and comparative distributions of the components of these risk factors.
- Aetiological associations of risk factors should be appropriately compared through the use of error-corrected hazard ratios per standard deviation change in error-corrected levels.

References

[1] Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol* 1990;**132:**734–45.

[2] Spearman C. The proof and measurement of association between two things. *Int J Epidemiol* 2010;**39:** 1137–50.

[3] Lovie S, Lovie P. Commentary: Charles Spearman and correlation: a commentary on 'The proof and measurement of association between two things'. *Int J Epidemiol* 2010;**39:**1151–53.

[4] Spiegelman D. Commentary: some remarks on the seminal 1904 paper of Charles Spearman 'The proof and measurement of association between two things'. *Int J Epidemiol* 2010;**39:**1156–59.

[5] de Klerk N. Commentary: Spearman's 'The proof and measurement of association between two things'. *Int J Epidemiol* 2010;**39:**1159–61.
[6] Perks J. Commentary: 'The next trick is impossible'. *Int J Epidemiol* 2010;**39:**1153–55.
[7] MacMahon S, Peto R, Cutler J *et al*. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;**335:**765–74.
[8] Law MR, Wald NJ, Wu T, Hackshaw A, Bailey A. Systematic underestimation of association between serum cholesterol concentration and ischaemic heart disease in observational studies: data from the BUPA study. *BMJ* 1994;**308:**363–66.
[9] Phillips AN, Davey Smith G. How independent are 'independent' effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clin Epidemiol* 1991;**44:**1223–31.
[10] Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med* 1989;**8:**1051–69.
[11] Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;**360:**1903–13.
[12] Fibrinogen Studies Collaboration. Regression dilution methods for meta-analysis: assessing long-term variability in plasma fibrinogen among 27 247 adults in 15 prospective studies. *Int J Epidemiol* 2006;**35:**1570–78.
[13] Allison DB, Paultre F, Goran MI, Poehlman ET, Heymsfield SB. Statistical considerations regarding the use of ratios to adjust data. *Int J Obes Relat Metab Disord* 1995;**19:**644–52.
[14] The Emerging Risk Factors Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009;**302:**1993–2000.
[15] The Emerging Risk Factors Collaboration. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet* 2011;**377:** 1085–95.
[16] The Emerging Risk Factors Collaboration. The Emerging Risk Factors Collaboration: analysis of individual data on lipid, inflammatory and other markers in over 1.1 million participants in 104 prospective studies of cardiovascular diseases. *Eur J Epidemiol.* 2007;**22:**839–69.
[17] Frost C, Thompson S. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *J R Statist Soc A.* 2000;**163:**173–89.
[18] Thompson SG, Kaptoge S, White IR, Wood AM, Perry P, Danesh J. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *Int J Epidemiol.* 2010;**39:**1345–59.
[19] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003; **327:**557–60.
[20] Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. *Stat Med* 2009; **28:**1067–92.
[21] Emberson JR, Shaper AG, Wannamethee SG, Morris RW, Whincup PH. Alcohol intake in middle age and risk of cardiovascular disease and mortality: accounting for intake variation over time. *Am J Epidemiol.* 2005;**161:**856–63.
[22] Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;**18:**499–502.
[23] Feng C, Wang H, Lu N, Tu XM. Log transformation: application and interpretation in biomedical research. *Stat Med* 2012.
[24] Frost C, White IR. The effect of measurement error in risk factors that change over time in cohort studies: do simple methods overcorrect for 'regression dilution'? *Int J Epidemiol* 2005;**34:**1359–68.
[25] Davey Smith G, Phillips AN. Inflation in epidemiology: ''the proof and measurement of association between two things'' revisited. *BMJ* 1996;**312:**1659–61.