Title
A generalisation of the method of regression calibration.

Journal
Scientific Reports, 13(1)

Authors
Little, Mark
Hamada, Nobuyuki
Zablotska, Lydia

Publication Date
2023-09-13

DOI
10.1038/s41598-023-42283-y

Peer reviewed

scientific reports

Check for updates

OPEN

A generalisation of the method of regression calibration

Mark P. Little[1]✉, Nobuyuki Hamada[2] & Lydia B. Zablotska[3]

There is direct evidence of risks at moderate and high levels of radiation dose for highly radiogenic cancers such as leukaemia and thyroid cancer. For many cancer sites, however, it is necessary to assess risks via extrapolation from groups exposed at moderate and high levels of dose, about which there are substantial uncertainties. Crucial to the resolution of this area of uncertainty is the modelling of the dose–response relationship and the importance of both systematic and random dosimetric errors for analyses in the various exposed groups. It is well recognised that measurement error can alter substantially the shape of this relationship and hence the derived population risk estimates. Particular attention has been devoted to the issue of shared errors, common in many datasets, and particularly important in occupational settings. We propose a modification of the regression calibration method which is particularly suited to studies in which there is a substantial amount of shared error, and in which there may also be curvature in the true dose response. This method can be used in settings where there is a mixture of Berkson and classical error. In fits to synthetic datasets in which there is substantial upward curvature in the true dose response, and varying (and sometimes substantial) amounts of classical and Berkson error, we show that the coverage probabilities of all methods for the linear coefficient α are near the desired level, irrespective of the magnitudes of assumed Berkson and classical error, whether shared or unshared. However, the coverage probabilities for the quadratic coefficient β are generally too low for the unadjusted and regression calibration methods, particularly for larger magnitudes of the Berkson error, whether this is shared or unshared. In contrast Monte Carlo maximum likelihood yields coverage probabilities for β that are uniformly too high. The extended regression calibration method yields coverage probabilities that are too low when shared and unshared Berkson errors are both large, although otherwise it performs well, and coverage is generally better than these other three methods. A notable feature is that for all methods apart from extended regression calibration the estimates of the quadratic coefficient β are substantially upwardly biased.

Cancer risks following exposure to moderate and high levels of radiation dose are reasonably well understood[1,2]. There are beginning to be studies yielding direct estimates of radiation risk at low dose (< 100 mGy) low-linear energy transfer (LET) radiation[3–6]. This is particularly the case for highly radiogenic sites such as thyroid[3] and leukaemia[4]. For most other cancer endpoints it is necessary to assess risks via extrapolation from groups exposed at moderate and high levels of dose. A number of recent reviews of low dose risk have been conducted, in particular those by the National Council on Radiation Protection and Measurements (NCRP)[7] and by the National Cancer Institute[8–13]. A major source of uncertainty in assessment of low dose risk concerns the extrapolation of risks at high doses and high dose-rates to those at low doses (< 0.1 gray (Gy)) and low dose-rates (< 5 mGy/hour)[14]. Crucial to the resolution of this area of uncertainty is the modelling of the dose–response relationship and the importance of both systematic and random dosimetric errors for analyses of the dose response, in particular in the Japanese atomic bomb survivors, which is central to evaluations of population risks by a number of committees assessing radiation risk[1,15]. The problem of allowing for measurement error in dose when estimating dose–response relationships has been the subject of much interest in epidemiology[16–31]. A recent review paper summarises at least some of the methods that have been used[32]. It is well recognised that measurement error can alter substantially the shape of this relationship and hence the derived population risk estimates[33]. A method that has been frequently used to correct for the effects of classical error is regression calibration, in which the terms for true dose, D_i, in regression models are replaced by the conditional expectation of true dose given the observed dose d_i, $E[D_i|d_i]$[33]. Regression calibration works well when the magnitude of errors is modest, and when

[1]Radiation Epidemiology Branch, National Cancer Institute, Room 7E546, 9609 Medical Center Drive, Bethesda, MD 20892-9778, USA. [2]Biology and Environmental Chemistry Division, Sustainable System Research Laboratory, Central Research Institute of Electric Power Industry (CRIEPI), 1646 Abiko, Chiba 270-1194, Japan. [3]Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, 550 16th Street, 2nd Floor, San Francisco, CA 94143, USA. ✉email: mark.little@nih.gov

the dose response is not substantially non-linear[33]. When errors are larger methods that take account of the full error distribution such as Monte Carlo maximum likelihood (MCML)[25–27,31] or Bayesian Markov Chain Monte Carlo (MCMC)[22–24,30] are likely to perform better.

Dose measurement errors can arise in a number of different ways. In radiotherapy (RT), for example, a machine may be used for delivering radiation doses, D_i, to a patient, and these true values are randomly distributed around the measured dial setting on the RT machine, d_i, with error U_i, so that $D_i = d_i + U_i$, implying that the d_i, U_i are independent, i.e., the Berkson error model. Alternatively, the measured "doses", d_i can be distributed at random around the true "doses", D_i, so that $d_i = D_i + U_i$ so that the D_i, U_i are independent, i.e., the "classical" error model. Although these models look very similar, they are different. In particular the crucial difference is that in the Berkson model the nominal dose and error are independent, but in the classical error model it is the true dose and the error that are independent. In the atomic bomb survivors, radiation doses are estimated by using estimates of the position of the survivors in each city, orientation with respect to the bomb and other shielding structures, e.g., buildings. In this case the estimated doses, d_i, are thought to be lognormally distributed around the true doses, D_i (i.e. classical error model)[34]. This assumption underlies many of the attempts that have been made to model dose error in the Japanese atomic bomb survivor Life Span Study (LSS) data[16–20,22–24,30]. However, some components of assessed dose to the atomic bomb survivors may be associated with Berkson error, for example that associated with estimation of the atomic bomb source term. Some attempts have been made to model this statistically[35]. Methods have been devised that allow for a combination of Berkson and classical errors in the LSS data[36,37]; although shared errors have not been explicitly modelled in the LSS they undoubtedly exist, as for example in the estimates of the bomb yield in the two cities. It is known that regression calibration can work well in cases when dose errors are not substantial and in which there is no curvature in the dose response[33]. However, it is also appreciated that there can be substantial bias in regression calibration when dose errors are substantial, also when errors are non-differential[33,38,39].

We propose a modification of the regression calibration method which is particularly suited to studies in which there is a substantial amount of shared error, and in which there may also be curvature in the true dose response. We compare the performance of this and other methods for dose error correction using synthetic data closely modelled on the Japanese atomic bomb survivor data[40].

Methods

Synthetic data used for assessing corrections for dose error. We used the publicly available version of the leukaemia and lymphoma data of Hsu et al.[40] to guide construction of a synthetic dataset, which we provide in outline in Table 1. Specifically we used the person year distribution by bone marrow dose groups 0–0.07, 0.08–0.19, 0.20–0.99, 1.00–2.49, ≥ 2.50 Gy. The central estimates of dose we assumed are close to the person year weighted means of these groups, and as given in Table 1, although for the uppermost dose group we assigned a central estimate of 2 Gy. The numbers of persons are close to the scaled sum of person years in these dose groups, scaling by a factor of 0.002. We assumed a composite Berkson-classical error model in which the true dose $D_{true,i,j}$ and the surrogate dose $D_{surr,i,j}$ to individual i (in dose group k_i) in simulation j are given by:

$$D_{true,i,j} = D_{cent,k_i} \exp\left[-0.5(\sigma^2_{share,Berkson} + \sigma^2_{unshare,Berkson})\right] \exp\left[\sigma_{share,Berkson}\varepsilon_j + \sigma_{unshare,Berkson}\delta_{i,j}\right] \quad (1)$$

$$D_{surr,i,j} = D_{cent,k_i} \exp\left[-0.5(\sigma^2_{share,Class} + \sigma^2_{unshare,Class})\right] \exp\left[\sigma_{share,Class}\mu_j + \sigma_{unshare,Class}\kappa_{i,j}\right] \quad (2)$$

The variables $\varepsilon_j, \delta_{i,j}, \mu_j, \kappa_{i,j}$ are independent identically distributed $N(0,1)$ random variables. The factors $D_{cent,k_i}, D_{cent,k_i}$ are the central estimates of dose, as given in Table 1. The factors $\exp\left[-0.5(\sigma^2_{share,Berkson} + \sigma^2_{unshare,Berkson})\right]$ and $\exp\left[-0.5(\sigma^2_{share,Class} + \sigma^2_{unshare,Class})\right]$ ensure that the distributions given by (1) and (2) have theoretical mean that coincides with the central estimates D_{cent,k_i}. This composite Berkson-classical error model is suggested by a similar (but purely additive) model proposed by Reeves et al.[21], whereas the errors in our model are of multiplicative form; the model of course ensures that the simulated doses are always positive. The model has the feature that when the Berkson error geometric standard deviations (GSDs) are set to 0 ($\sigma_{share,Berkson} = \sigma_{unshare,Berkson} = 0$) the model reduces to one with classical error (a mixture of shared and unshared); likewise when the classical error GSDs are set to 0 ($\sigma_{share,Class} = \sigma_{unshare,Class} = 0$) the model reduces to one with pure Berkson error (a mixture of shared and unshared).

We generated a number of different versions of the dose data, with GSD $\sigma_{share,Berkson}$, $\sigma_{unshare,Berkson}$, $\sigma_{share,Class}$, $\sigma_{unshare,Class}$ taking values of 0.2 (20%) or 0.5 (50%). We also explored 4 scenarios with pure classical error, with

Dose group	Central estimate of dose (Gy)	Scaled numbers of persons
1	0.01	2591
2	0.1	334
3	0.5	438
4	1.5	102
5	2	6

Table 1. Assumed distribution of persons by radiation dose group, based in part on distribution of person years in the Japanese atomic bomb survivor Life Span Study[40].

the Berkson error terms set to 0. This individual dose data was then used to simulate the distribution of $N = 250$ cancers for each of $m = 1000$ simulated datasets, indexed by j, using a model in which the assumed probability of being a case for individual i is given by:

$$\lambda_j[1 + \alpha D_{true,i,j} + \beta D_{true,i,j}^2] \tag{3}$$

the scaling constant λ_j being chosen for each simulation to make these sum to 1. We assumed coefficients $\alpha = 0.25/\mathrm{Gy}, \beta = 2/\mathrm{Gy}^2$, close to the values derived from fits of a similar model to the 237 leukaemias in the data of Hsu et al.[40].

A total of $m = 1000$ samples were taken of each type of dose, as given by expressions (1) and (2). A total of $n = 500$ simulations of these dose + cancer ensembles were used to fit models and evaluate fitted model means and coverage probability. Having derived synthetic individual level data, for the purposes of model fitting, for all models except MCML, the data were then collapsed (summing cases, averaging doses) into the 5 dose groups given in Table 1. Poisson linear relative risk generalised linear models[41] were fitted to this grouped data, with rates given by expression (3), using as offsets the number per group in Table 1. Models were fitted using four separate methods:

(1) unadjusted – using only the mean surrogate doses per group given by group means of the samples generated by expression (2), using a single sampled dose per individual for each of $m = 500$ dose + cancer ensembles;
(2) regression calibration adjusted – using the mean true doses per group given by group means of the samples generated by expression (1), averaged over the $n = 1000$ dose samples, for each of $m = 500$ dose + cancer ensembles;
(3) extended regression calibration adjusted – using the mean true doses per group given by group means of the samples generated by expression (1), averaged over the $n = 1000$ dose samples, for each of $m = 500$ dose + cancer ensembles, and with additional adjustments to the likelihood outlined in Appendix A;
(4) MCML, using the full set of mean true doses per group, the mean doses per group for each simulation being given by group means of the samples generated by expression (1), averaged over the $n = 1000$ dose samples.

In all cases confidence intervals were derived using the profile likelihood[41]. The Fortran 95-2003 program used to generate these datasets and perform Poisson model fitting, and the relevant steering files employed to control this program are given in online Appendix B.

Results

As shown in Table 2, the coverage probabilities of all methods for the linear coefficient α are near the desired 95% level, irrespective of the magnitudes of assumed Berkson and classical error, whether shared or unshared. However, the coverage probabilities for the quadratic coefficient β are generally too low for the unadjusted and regression calibration methods, particularly for larger magnitudes of Berkson error (with GSD = 50%), whether this is shared or unshared (Table 2). The extended regression calibration method also yields coverage probabilities that are too low when shared and unshared Berkson errors are both large (with GSD = 50%), although otherwise it performs well, and coverage is uniformly better than these other two methods (Table 2). In contrast MCML yields coverage probabilities for β that are uniformly too high (Table 2). The interindividual correlations of true dose are generally moderate to high, ranging from 0.15 to 0.84 (Table 2). The correlations between the group mean true doses are generally very high, in all cases > 0.95, for obvious reasons—as a result of the averaging the unshared errors will become relatively much less important than the shared errors (which are unaffected by averaging), and it is these that drive the correlations.

Table 3 shows the coefficient mean values, averaged over all 500 simulations. A notable feature is that for all methods apart from extended regression calibration the estimates of the quadratic coefficient β are upwardly biased. There is upward bias in estimates of both α and β in the unadjusted analysis (using surrogate dose) even when there are no Berkson errors, for various magnitudes of classical errors, as shown by the first four rows of Table 3. As can be seen from Fig. 1, in this case (with shared and unshared classical errors having GSD = 50%) the mean ratio of surrogate to true dose is lognormal in the way one would expect, but as shown in Fig. 2 the fitted $\hat{\alpha}$ and $\hat{\beta}$ are markedly skew, with pronounced upper tail, particularly for $\hat{\beta}$. It is this long upper tail that accounts for the upward bias in both $\hat{\alpha}$ and $\hat{\beta}$ in the unadjusted analysis (using surrogate dose).

Discussion

We have demonstrated that the coverage probabilities of all methods for the linear coefficient α are near the desired 95% level, irrespective of the magnitudes of assumed Berkson and classical error, whether shared or unshared (Table 2). The coverage probabilities for the quadratic coefficient β are generally too low for the unadjusted and regression calibration methods, particularly for larger magnitudes of Berkson error (with GSD = 50%), whether this is shared or unshared; by contrast the coverage probabilities for β using MCML are uniformly too high (Table 2). The extended regression calibration method yields generally more satisfactory coverage probabilities, in most cases better than the other methods (Table 2). The reason for the coverage probabilities of the quadratic coefficient β being unsatisfactory may be related to the fact that for all methods apart from extended regression calibration the estimates of this parameter are upwardly biased, much more substantially so than for α (Table 3). The fact that β may not be well estimated implies that assessments of curvature may be incorrect,

Magnitude of error distribution (GSD)				Sample Pearson correlation coefficient between individual true doses	Unadjusted model		Regression calibration adjusted		Extended regression calibration adjusted		Monte Carlo maximum likelihood	
					Coverage %		Coverage %		Coverage %		Coverage %	
Unshared Berkson error (%)	Shared Berkson error (%)	Unshared classical error (%)	Shared classical error (%)		α	β	α	β	α	β	α	β
0	0	20	20	NA	95.0	80.8	95.2	94.8	95.2	94.8	95.2	94.8
0	0	20	50	NA	94.4	55.4	95.2	94.8	95.2	94.8	95.2	94.8
0	0	50	20	NA	94.4	79.6	95.2	94.8	95.2	94.8	95.2	94.8
0	0	50	50	NA	94.6	55.0	95.2	94.8	95.2	94.8	95.2	94.8
20	20	20	20	0.50	95.0	80.0	95.4	94.0	94.8	98.4	95.4	99.0
20	20	20	50	0.50	94.8	53.6	95.4	94.0	94.8	98.4	95.4	99.0
20	20	50	20	0.50	94.6	78.2	95.4	94.0	94.8	98.4	95.4	99.0
20	20	50	50	0.50	94.6	52.2	95.4	94.0	94.8	98.4	95.4	99.0
20	50	20	20	0.84	93.6	77.4	94.4	85.6	95.4	94.8	95.0	100.0
20	50	20	50	0.84	93.2	48.4	94.4	85.6	95.4	94.8	95.0	100.0
20	50	50	20	0.84	93.8	75.4	94.4	85.6	95.4	94.8	95.0	100.0
20	50	50	50	0.84	93.8	48.2	94.4	85.6	95.4	94.8	95.0	100.0
50	20	20	20	0.15	94.2	76.6	94.0	86.0	94.4	94.8	95.0	99.2
50	20	20	50	0.15	93.4	48.4	94.0	86.0	94.4	94.8	95.0	99.2
50	20	50	20	0.15	94.0	75.6	94.0	86.0	94.4	94.8	95.0	99.2
50	20	50	50	0.15	94.0	49.0	94.0	86.0	94.4	94.8	95.0	99.2
50	50	20	20	0.45	95.4	64.0	95.4	67.8	95.0	80.4	96.4	100.0
50	50	20	50	0.45	95.0	40.0	95.4	67.8	95.0	80.4	96.4	100.0
50	50	50	20	0.45	94.4	64.6	95.4	67.8	95.0	80.4	96.4	100.0
50	50	50	50	0.45	94.2	40.0	95.4	67.8	95.0	80.4	96.4	100.0

Table 2. Coverage probability of profile likelihood confidence intervals for fits of linear-quadratic model. Coverage probability evaluated using $m = 500$ dose + cancer simulations. *GSD* geometric standard deviation.

and in particular may result in overestimation of the degree of curvature in the dose response, at least for the scenarios investigated here.

An unexpected feature of our analysis is that when there is only classical error the unadjusted analysis (using surrogate dose) can result in appreciable upward bias, contrary to what is often seen when there is pure classical error (Table 3). In this case the ratio of doses (surrogate to true) is approximately lognormal (Fig. 1) and for each simulation the ratio is generally much the same in all dose groups except the topmost one, suggesting that it is the shared classical error that is dominating—the unshared error averages out in general, although it does contribute somewhat to the topmost group (data not shown). Although the distribution of fitted $\hat{\alpha}$ and $\hat{\beta}$ to some extent reflect this, as shown in Fig. 2 the distributions of both optimal $\hat{\alpha}$ and $\hat{\beta}$ are markedly skew, with pronounced upper tail, particularly for $\hat{\beta}$. This results in pronounced upward bias in the mean estimates of $\hat{\alpha}$ and $\hat{\beta}$ for the unadjusted (surrogate dose) analysis (Fig. 2). The reason for the skewness of the fitted $\hat{\alpha}$ and $\hat{\beta}$ is reasonably obvious—given the range of true doses generated (up to the level of about 2 Gy), the $\hat{\alpha}$ and $\hat{\beta}$ cannot be very substantially negative without the relative risk for the higher dose groups becoming negative, which would lead to the likelihood blowing up. It should also be noted that when there is only classical error, as implied by expression (1) all true doses used for regression calibration, extended regression calibration and MCML are precisely the central estimates given in Table 1. This implies that in this case regression calibration and MCML will yield precisely the same regression coefficients. Since the covariance term that is used to adjust the likelihood for extended regression calibration becomes trivial (i.e., 0), the second order likelihood adjustment term in Appendix A expression (A3) drops out, and extended regression calibration reduces to the standard type of calibration.

The defects in regression calibration that our modelling has revealed are not too surprising, as it is well known that this method can break down when dose error is substantial[33], as it is in many of our scenarios. The essence of regression calibration is to replace of the vector of true doses (D_i) in the expression for the theoretical likelihood $L[(y_i), \vartheta, (D_i), (Z_i)]$ by the vector of conditional expectations $(E[D_i|d_i, Z_i])$ of true dose (D_i) given the nominal or observed dose (d_i) and ancillary variables (Z_i). The method is relatively simple to apply, although it does require some method of determining the magnitude of dose error, as well as the distribution of true dose in the data. However, the distribution of true dose can be determined to some extent via deconvolution of the distribution of nominal dose. The method has the considerable advantage that once the conditional expectations have been derived conventional statistical software can be used to perform regressions. The method has been successfully applied to the LSS cohort by a number of investigators[16–20,42] and has also been used in a few other radiation exposed groups[26]. There have also been extensive applications in the non-radiation literature, reviewed by Carroll et al.[33] and more recently in a series of papers by Shaw et al.[38,39]. Calibration approaches

Magnitude of error distribution (GSD)				Unadjusted		Regression calibration		Extended regression calibration		Monte Carlo maximum likelihood	
				ERR/Gy	ERR/Gy^2	ERR/Gy	ERR/Gy^2	ERR/Gy	ERR/Gy^2	ERR/Gy	ERR/Gy^2
Unshared Berkson error (%)	Shared Berkson error (%)	Unshared classical error (%)	Shared classical error (%)	α	β	α	β	α	β	α	β
0	0	20	20	0.221	2.278	0.196	2.061	0.196	2.061	0.196	2.061
0	0	20	50	0.288	4.168	0.196	2.061	0.196	2.061	0.196	2.061
0	0	50	20	0.255	2.260	0.196	2.061	0.196	2.061	0.196	2.061
0	0	50	50	0.328	4.136	0.196	2.061	0.196	2.061	0.196	2.061
20	20	20	20	0.220	2.469	0.195	2.233	0.125	2.132	0.288	2.207
20	20	20	50	0.287	4.523	0.195	2.233	0.125	2.132	0.288	2.207
20	20	50	20	0.255	2.451	0.195	2.233	0.125	2.132	0.288	2.207
20	20	50	50	0.328	4.492	0.195	2.233	0.125	2.132	0.288	2.207
20	50	20	20	0.262	2.983	0.227	2.707	0.109	2.393	0.370	3.007
20	50	20	50	0.354	5.426	0.227	2.707	0.109	2.393	0.370	3.007
20	50	50	20	0.303	2.962	0.227	2.707	0.109	2.393	0.370	3.007
20	50	50	50	0.401	5.390	0.227	2.707	0.109	2.393	0.370	3.007
50	20	20	20	0.259	2.986	0.224	2.709	0.121	2.354	0.337	2.678
50	20	20	50	0.347	5.441	0.224	2.709	0.121	2.354	0.337	2.678
50	20	50	20	0.299	2.964	0.224	2.709	0.121	2.354	0.337	2.678
50	20	50	50	0.395	5.401	0.224	2.709	0.121	2.354	0.337	2.678
50	50	20	20	0.243	3.703	0.209	3.349	0.038	2.795	0.362	3.401
50	50	20	50	0.332	6.744	0.209	3.349	0.038	2.795	0.362	3.401
50	50	50	20	0.286	3.682	0.209	3.349	0.038	2.795	0.362	3.401
50	50	50	50	0.383	6.703	0.209	3.349	0.038	2.795	0.362	3.401
True value				0.25	2.0	0.25	2.0	0.25	2.0	0.25	2.0

Table 3. Mean over $m = 500$ dose + cancer simulations of regression coefficients in fits of linear-quadratic model. *GSD* geometric standard deviation, *ERR* excess relative risk.

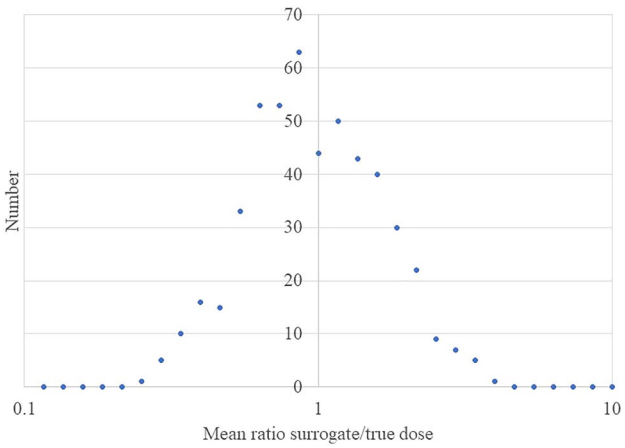

Figure 1. Distribution of weighted mean ratio of surrogate to true dose when there is 50% shared classical error, 50% shared classical error, no Berkson error (as in 4th row of Table 3). A logarithmic X-axis is used, with step size = $10^{(1/15)}$.

that take account of mixtures of Berkson and classical error have also been developed and used to fit domestic radon case–control data[21].

The relatively poor performance of MCML is perhaps more surprising. MCML relies on replacing the likelihood, as a function of the true dose vectors (D_i), by its expectation with respect to the nominal dose array (d_i), $E\left[L[(y_i), \vartheta, (D_i), (d_i)]|(d_i)\right] = \int L[(y_i), \vartheta, (D_i), (d_i)]\, dP(D_i|d_i)$. The marginal likelihood thus derived can then be used for likelihood-based inference in the usual way[43]. The integration is often achieved via Monte Carlo samples, produced from a Monte Carlo dosimetry system (MCDS) that can simulate true doses based on often quite complex dosimetric models, which can incorporate uncertainties in many dosimetric and other parameters.

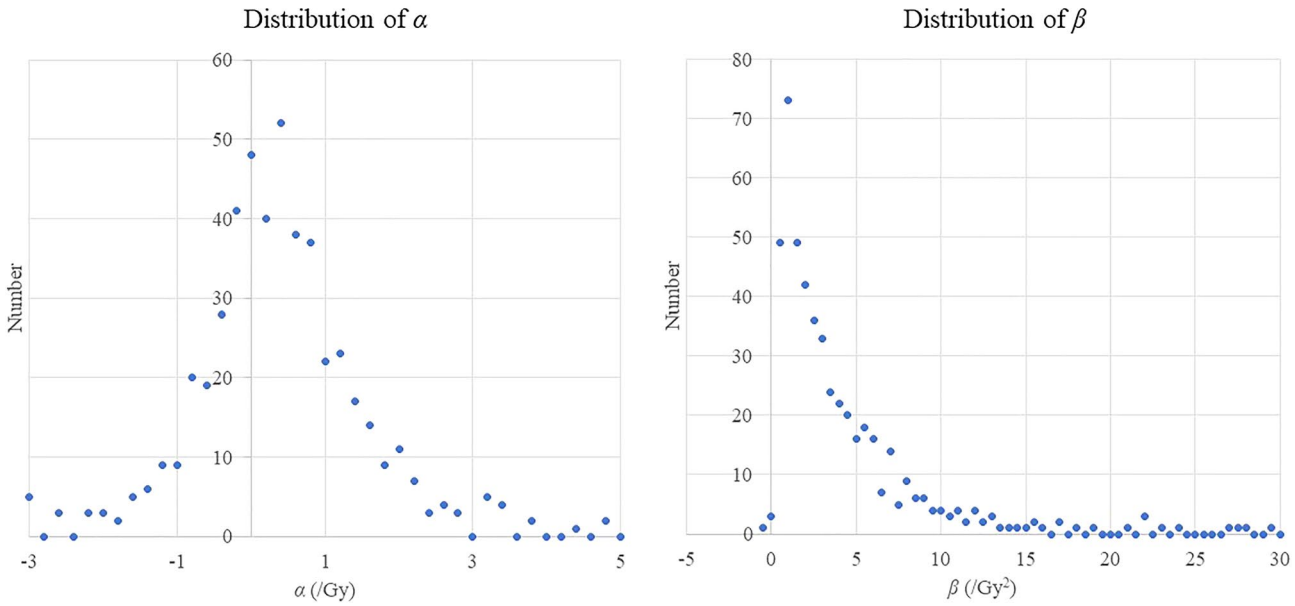

Figure 2. Distribution of fitted linear and quadratic coefficients when there is 50% shared classical error, 50% shared classical error, no Berkson error (as in 4th row of Table 3). The step size used for α is 0.2, the step size used for β is 0.5.

Implementation of MCML relies on specialist software, often written in high level languages such as Fortran or C/C++, and is generally highly computationally burdensome. It may suffer from the additional problem occasioned by attempting to sample from very high dimensional distributions, the so-called curse of dimensionality, which implies that a large part of the overall distribution of true dose will not have been sampled. However, whether this is a problem in practice is not always altogether clear—for example the underlying set of parameters being sampled may be in some cases quite low dimensional. In particular, the Monte Carlo simulations inspired by the Mayak worker data exhibit little evidence of upward bias, at most 15% or so, arguably of little material significance given the uncertainties[44]. Even where such problems may arise there may be ways round this, for example by using importance Monte Carlo sampling, as outlined by Dai et al.[45]. MCML has been used for analysis of nuclear workers[46], indoor radon data[47] and in a number of studies of Chernobyl-exposed groups[25–27,31], and in a few other datasets[48]. The poor performance of MCML in our study may reflect the fact that there is hidden correlation within each group, which MCML cannot take into account, given the collapsed nature of the data that we use.

A Bayesian approach to the measurement error problem has been developed over the last 30 years which rests on the formulation of conditional independence relationships between different model components[49,50], following the general structure outlined by Clayton[51]. In this approach three basic sub-models are distinguished and linked: the disease model, the measurement model and the exposure model. The power of this Bayesian approach, as with MCML, is that the dosimetric uncertainty is (in principle) reflected in the variability of the model parameters relating dose to health effects. An adapted Bayesian method of correction for measurement error has been fitted to various versions of the LSS mortality data[22–24,30], also to an older version of the LSS incidence data[23]. Derivation of the posterior distribution is generally analytically infeasible, and relies on the MCMC algorithm, specifically the Metropolis sampler, which converges to the posterior distribution of the risk parameters. However, the speed of convergence is not known, and in practice one relies on a number of ad hoc tests of convergence such as the Brooks-Gelman-Rubin statistic[52,53] and other less formal methods, e.g., inspection of caterpillar plots. As such all one can know is when convergence has not taken place. Flexible and efficient software exists to run this on a number of platforms e.g., OpenBUGS[54] or rjags[55]. The method is exceptionally computationally burdensome. As with all Bayesian methods the choice of prior is critical.

Some other methods of more limited utility have been developed for dealing with dosimetric error, which we briefly review. The simulation-extrapolation (SIMEX) method was developed by Cook and Stefanski[56]. It was originally proposed for datasets where the error is of pure classical form, and where the precise magnitude of the dose error is known. The method proceeds by adding classical random error with progressively larger GSD to the nominal dose estimates, performing regression analyses, this Monte Carlo procedure being repeated a large number of times to reduce sampling uncertainties. A curve is then fitted to the regression estimates as a function of magnitude of dose error, and the curve used to extrapolate back to 0 error. It is computationally highly intensive. R packages exist (e.g. **simex**[57]) to fit at least certain types of generalised linear model[41] although not the linear relative risk models in common use in epidemiological analysis of radio-epidemiological data. Quite apart from the computational difficulties, the method relies on a substantial extrapolation (from the given level of dose error to 0 error), a jump that may be difficult to justify. An attempt has been made to expand SIMEX to allow for a mixture of classical and Berkson errors utilising the LSS data[37]. Perhaps due to the computational cost with the cross-tabulation and because of the limited types of error structure that can be handled it has been used only twice to our knowledge, in analysis of the LSS data[28,37].

The so-called two dimensional Monte Carlo using Bayesian model averaging (2DMC-BMA) method relies on Monte Carlo simulations from an MCDS. The key aspect is that ensembles of doses $(D_{ijk})_{j=1\ k=1}^{N\ n_j}$ are produced for all individuals for a large number of scenarios $i, 1 \leq i \leq M$. However, unlike other uses of MCDS it is assumed that only one of the dose scenarios i, and therefore one of the sets of dose realisations $(D_{ijk})_{j=1\ k=1}^{N\ n_j}$ is the correct one. Essentially this method therefore assumes something like a combination of functional and structural approaches—there are assumed to be random errors in the data, but certain parameters are assumed fixed (but unknown). The BMA approach is used to reweight the scenarios depending on the goodness of fit[29]. So realisations where the risk-dose relationship was linear would be much more highly weighted than realisations where this was not the case. The contrast with MCML is quite pronounced—MCML works by averaging the likelihood in one go and then maximising the averaged likelihood with respect to the parameters of interest. The 2DMC-BMA method appears designed for applications where there is a substantial amount of shared error. This method has been applied to analysis of thyroid nodules in a dataset of persons exposed to atmospheric weapons tests in the former Soviet Union[58]. The method has been much discussed[44]. Stram et al.[44] suggested that the method will produce substantially upwardly biased estimates of risk, also that the coverage may be poor. The implementation of the methodology presently relies on proprietary software, and has only been used by the group that developed it. Another substantial problem with the method is the use of BMA, reflecting general criticism made of this class of models in the literature[59,60]. An implicit assumption of BMA is that one of the underlying models is the "true" one with convergence guaranteed to the "true" model[61]. As with all Bayesian methods the choice of prior is critical.

Zhang et al.[62] developed their corrected information matrix (CIM) method for analysis of datasets where there is pure Berkson error in radiation dose, a substantial part of it shared. This entails an extensive calculation, which requires specially written software, which the authors have developed in Python[63] specifically applied to the Mayak worker lung cancer data. R code has also been developed for fitting this model to US radiologic technologists (USRT) cataract data for relative risk and absolute risk Poisson models[64]. The calculations result in inflation of the confidence intervals (CI) on the regression estimate—the central estimate is largely unchanged. Arguably the assumptions underlying the CIM method, that all dose simulations are samples from the true dose, may be unlikely, but this assumption is arguably less implausible than that made for 2DMC-BMA, which assumes that one realisation is true. The method appears to be well adapted to analysis of the Mayak data[63], where there is a substantial amount of shared error. In the USRT cataract data, the amount of shared error is small, and the method yields largely trivial adjustments to CI[64].

A relatively novel method of measurement error correction has been recently introduced, moment reconstruction (MR)[65]. The basic idea is that one substitutes for the nominal dose estimate d_i a new quantity M_{d_i,Y_i} which is chosen to have the same first two moments (with the outcome variable Y_i) of the joint distribution as (D_i, Y_i). It can be shown[65] that the solution is given by $M_{d_i,Y_i} = E[d_i|Y_i](1-G) + d_iG$ where $G = G(Y) = \mathrm{cov}[D_i|Y_i]^{0.5}(\mathrm{cov}[d_i|Y_i])^{-0.5}$. Under linear regression it is easily shown that MR is entirely equivalent to regression calibration[65]. It has the advantage over regression calibration that it yields consistent estimates even when the model is non-linear, or when the errors in dose are non-differential[65]. Moment-adjusted imputation (MAI) is a generalisation of MR, in which the moments of (D_i, Y_i) are matched by M_{d_i,Y_i}, usually up to at least the 4th order[66,67]. However, both MR and MAI require knowledge of second and higher order moments of the true dose distribution in conjunction with the disease endpoint, information that would generally have to come from a gold standard sample, which is not often available in radiation studies. Although MR and MAI can be more efficient than regression calibration there are circumstances when efficiency is reduced compared with regression calibration[39]. Perhaps for all these reasons, to the best of our knowledge neither method has been used in radiation applications.

Conclusions

We have outlined a modification of the regression calibration method[33] which is particularly suited to studies where there is a substantial amount of shared error, and where there may also be curvature in the true dose response. We have shown in fits to a number of synthetic datasets in which there is substantial upward curvature in the true dose response, and varying (and sometimes substantial) amounts of classical and Berkson error, that the coverage probabilities of all methods for the linear coefficient are near the desired level, irrespective of the magnitudes of assumed Berkson and classical error, whether shared or unshared. However, the coverage probabilities for the quadratic coefficient are generally too low for the unadjusted and regression calibration methods, particularly for larger magnitudes of the Berkson error, whether this is shared or unshared, while MCML yields coverage probabilities for the quadratic coefficient that are uniformly too high. The extended regression calibration method yields coverage probabilities that are too low when shared and unshared Berkson errors are both large, although otherwise it performs well, and coverage is generally better than these other methods. A notable feature is that for all methods apart from extended regression calibration the estimates of the quadratic coefficient are substantially upwardly biased.

Data availability

The datasets generated and analysed in the current study are available by running the Fortran 95/2003 program fitter_shared_error_simulation_reg_cal.for, given in the online web repository, with any of the five steering input files given there. All are described in Appendix B. The datasets are temporarily stored in computer memory, and the program uses them for fitting the Poisson models described in the "Methods" section.

Received: 9 August 2023; Accepted: 7 September 2023
Published online: 13 September 2023

References

1. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). *UNSCEAR 2006 Report. Annex A. Epidemiological Studies of Radiation and Cancer.* 13–322 (United Nations, New York, 2008).
2. Armstrong, B. *et al. Radiation. A Review of Human Carcinogens.* Vol. 100D. 1–341 (International Agency for Research on Cancer, Lyon, 2012).
3. Lubin, J. H. *et al.* Thyroid cancer following childhood low-dose radiation exposure: A pooled analysis of nine cohorts. *J. Clin. Endocrinol. Metab.* **102**, 2575–2583. https://doi.org/10.1210/jc.2016-3529 (2017).
4. Little, M. P. *et al.* Leukaemia and myeloid malignancy among people exposed to low doses (<100 mSv) of ionising radiation during childhood: A pooled analysis of nine historical cohort studies. *Lancet Haematol.* **5**, e346–e358. https://doi.org/10.1016/S2352-3026(18)30092-9 (2018).
5. Little, M. P. *et al.* Review of the risk of cancer following low and moderate doses of sparsely ionising radiation received in early life in groups with individually estimated doses. *Environ. Int.* **159**, 106983. https://doi.org/10.1016/j.envint.2021.106983 (2022).
6. Little, M. P. *et al.* Cancer risks among studies of medical diagnostic radiation exposure in early life without quantitative estimates of dose. *Sci. Total Environ.* **832**, 154723. https://doi.org/10.1016/j.scitotenv.2022.154723 (2022).
7. National Council on Radiation Protection and Measurements (NCRP). *Implications of Recent Epidemiologic Studies for the Linear-Nonthreshold Model and Radiation Protection. NCRP Commentary No 27.* i–ix, 1–199 (National Council on Radiation Protection and Measurements (NCRP), Bethesda, 2018).
8. Berrington de Gonzalez, A. *et al.* Epidemiological studies of low-dose ionizing radiation and cancer: Rationale and framework for the monograph and overview of eligible studies. *J. Natl. Cancer Inst. Monogr.* **2020**, 97–113. https://doi.org/10.1093/jncimonographs/lgaa009 (2020).
9. Hauptmann, M. *et al.* Epidemiological studies of low-dose ionizing radiation and cancer: Summary bias assessment and meta-analysis. *J. Natl. Cancer Inst. Monogr.* **2020**, 188–200. https://doi.org/10.1093/jncimonographs/lgaa010 (2020).
10. Linet, M. S., Schubauer-Berigan, M. K. & Berrington de Gonzalez, A. Outcome assessment in epidemiological studies of low-dose radiation exposure and cancer risks: Sources, level of ascertainment, and misclassification. *J. Natl. Cancer Inst. Monogr.* **2020**, 154–175. https://doi.org/10.1093/jncimonographs/lgaa007 (2020).
11. Schubauer-Berigan, M. K. *et al.* Evaluation of confounding and selection bias in epidemiological studies of populations exposed to low-dose, high-energy photon radiation. *J. Natl. Cancer Inst. Monogr.* **2020**, 133–153. https://doi.org/10.1093/jncimonographs/lgaa008 (2020).
12. Gilbert, E. S., Little, M. P., Preston, D. L. & Stram, D. O. Issues in interpreting epidemiologic studies of populations exposed to low-dose, high-energy photon radiation. *J. Natl. Cancer Inst. Monogr.* **2020**, 176–187. https://doi.org/10.1093/jncimonographs/lgaa004 (2020).
13. Daniels, R. D., Kendall, G. M., Thierry-Chef, I., Linet, M. S. & Cullings, H. M. Strengths and weaknesses of dosimetry used in studies of low-dose radiation exposure and cancer. *J. Natl. Cancer Inst. Monogr.* **2020**, 114–132. https://doi.org/10.1093/jncimonographs/lgaa001 (2020).
14. Wakeford, R. & Tawn, E. J. The meaning of low dose and low dose-rate. *J. Radiol. Prot.* **30**, 1–3. https://doi.org/10.1088/0952-4746/30/1/E02 (2010).
15. International Commission on Radiological Protection (ICRP). The 2007 Recommendations of the International Commission on Radiological Protection ICRP publication 103. *Ann. ICRP* **37**, 1–332. https://doi.org/10.1016/j.icrp.2007.10.003 (2007).
16. Pierce, D. A., Stram, D. O. & Vaeth, M. Allowing for random errors in radiation dose estimates for the atomic bomb survivor data. *Radiat. Res.* **123**, 275–284 (1990).
17. Pierce, D. A., Stram, D. O., Vaeth, M. & Schafer, D. W. The errors-in-variables problem: Considerations provided by radiation dose-response analyses of the A-bomb survivor data. *J. Am. Stat. Assoc.* **87**, 351–359. https://doi.org/10.1080/01621459.1992.10475214 (1992).
18. Little, M. P. & Muirhead, C. R. Evidence for curvilinearity in the cancer incidence dose-response in the Japanese atomic bomb survivors. *Int. J. Radiat. Biol.* **70**, 83–94 (1996).
19. Little, M. P. & Muirhead, C. R. Curvilinearity in the dose-response curve for cancer in Japanese atomic bomb survivors. *Environ. Health Perspect.* **105**(Suppl 6), 1505–1509 (1997).
20. Little, M. P. & Muirhead, C. R. Curvature in the cancer mortality dose response in Japanese atomic bomb survivors: Absence of evidence of threshold. *Int. J. Radiat. Biol.* **74**, 471–480 (1998).
21. Reeves, G. K., Cox, D. R., Darby, S. C. & Whitley, E. Some aspects of measurement error in explanatory variables for continuous and binary regression models. *Stat. Med.* **17**, 2157–2177 (1998).
22. Little, M. P., Deltour, I. & Richardson, S. Projection of cancer risks from the Japanese atomic bomb survivors to the England and Wales population taking into account uncertainty in risk parameters. *Radiat. Environ. Biophys.* **39**, 241–252 (2000).
23. Bennett, J., Little, M. P. & Richardson, S. Flexible dose-response models for Japanese atomic bomb survivor data: Bayesian estimation and prediction of cancer risk. *Radiat. Environ. Biophys.* **43**, 233–245. https://doi.org/10.1007/s00411-004-0258-3 (2004).
24. Little, M. P. *et al.* New models for evaluation of radiation-induced lifetime cancer risk and its uncertainty employed in the UNSCEAR 2006 report. *Radiat. Res.* **169**, 660–676. https://doi.org/10.1667/RR1091.1 (2008).
25. Kesminiene, A. *et al.* Risk of thyroid cancer among Chernobyl liquidators. *Radiat. Res.* **178**, 425–436. https://doi.org/10.1667/RR2975.1 (2012).
26. Little, M. P. *et al.* Impact of uncertainties in exposure assessment on estimates of thyroid cancer risk among Ukrainian children and adolescents exposed from the Chernobyl accident. *PLoS ONE* **9**, e85723. https://doi.org/10.1371/journal.pone.0085723 (2014).
27. Little, M. P. *et al.* Impact of uncertainties in exposure assessment on thyroid cancer risk among persons in Belarus exposed as children or adolescents due to the Chernobyl accident. *PLoS ONE* **10**, e0139826. https://doi.org/10.1371/journal.pone.0139826 (2015).
28. Allodji, R. S. *et al.* Simulation-extrapolation method to address errors in atomic bomb survivor dosimetry on solid cancer and leukaemia mortality risk estimates, 1950–2003. *Radiat. Environ. Biophys.* **54**, 273–283. https://doi.org/10.1007/s00411-015-0594-5 (2015).
29. Kwon, D., Hoffman, F. O., Moroz, B. E. & Simon, S. L. Bayesian dose-response analysis for epidemiological studies with complex uncertainty in dose estimation. *Stat. Med.* **35**, 399–423. https://doi.org/10.1002/sim.6635 (2016).
30. Little, M. P. *et al.* Lifetime mortality risk from cancer and circulatory disease predicted from the Japanese atomic bomb survivor Life Span Study data taking account of dose measurement error. *Radiat. Res.* **194**, 259–276. https://doi.org/10.1667/RR15571.1 (2020).
31. Little, M. P. *et al.* Impact of uncertainties in exposure assessment on thyroid cancer risk among cleanup workers in Ukraine exposed due to the Chornobyl accident. *Eur. J. Epidemiol.* **37**, 837–847. https://doi.org/10.1007/s10654-022-00850-z (2022).
32. Wu, Y. *et al.* Methods to account for uncertainties in exposure assessment in studies of environmental exposures. *Environ. Health* **18**, 31. https://doi.org/10.1186/s12940-019-0468-4 (2019).

33. Carroll, R. J., Ruppert, D., Stefanski, L. A. & Crainiceanu, C. M. *Measurement error in nonlinear models. A modern perspective.* 1–488 (Chapman and Hall/CRC, Boca Raton, 2006).
34. Jablon, S. in *ABCC Technical Report TR/23–71* (Atomic Bomb Casualty Commission, Hiroshima, 1971).
35. Pierce, D. A. & Kellerer, A. M. Adjusting for covariate errors with nonparametric assessment of the true covariate distribution. *Biometrika* **91**, 863–876. https://doi.org/10.1093/biomet/91.4.863 (2004).
36. Pierce, D. A., Vaeth, M. & Cologne, J. B. Allowance for random dose estimation errors in atomic bomb survivor studies: A revision. *Radiat. Res.* **170**, 118–126. https://doi.org/10.1667/RR1059.1 (2008).
37. Misumi, M., Furukawa, K., Cologne, J. B. & Cullings, H. M. Simulation-extrapolation for bias correction with exposure uncertainty in radiation risk analysis utilizing grouped data. *J. R. Stat. Soc. Ser. C-Appl. Stat.* **67**, 275–289. https://doi.org/10.1111/rssc.12225 (2018).
38. Keogh, R. H. *et al.* STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1-Basic theory and simple methods of adjustment. *Stat Med* **39**, 2197–2231. https://doi.org/10.1002/sim.8532 (2020).
39. Shaw, P. A. *et al.* STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2-More complex methods of adjustment and advanced topics. *Stat Med* **39**, 2232–2263. https://doi.org/10.1002/sim.8531 (2020).
40. Hsu, W.-L. *et al.* The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950–2001. *Radiat. Res.* **179**, 361–382. https://doi.org/10.1667/RR2892.1 (2013).
41. McCullagh, P. & Nelder, J. A. *Generalized linear models*. 2nd edition. 1–526 (Chapman and Hall/CRC, Boca Raton, 1989).
42. Little, M. P. & Muirhead, C. R. Derivation of low-dose extrapolation factors from analysis of curvature in the cancer incidence dose response in Japanese atomic bomb survivors. *Int. J. Radiat. Biol.* **76**, 939–953 (2000).
43. Schervish, M. J. *Theory of statistics*. 1–724 (Springer Verlag, Berlin, 1995)
44. Simon, S. L., Hoffman, F. O. & Hofer, E. Letter to the Editor Concerning Stram et al.: "Lung Cancer in the Mayak Workers Cohort: Risk Estimation and Uncertainty Analysis". *Radiat. Res.* **196**, 449–451. https://doi.org/10.1667/rade-21-00106.1 (2021).
45. Dai, C., Heng, J., Jacob, P. E. & Whiteley, N. An invitation to sequential Monte Carlo samplers. *J. Am. Stat. Assoc.* **117**, 1587–1600. https://doi.org/10.1080/01621459.2022.2087659 (2022).
46. Stayner, L. *et al.* A Monte Carlo maximum likelihood method for estimating uncertainty arising from shared errors in exposures in epidemiological studies of nuclear workers. *Radiat. Res.* **168**, 757–763. https://doi.org/10.1667/RR0677.1 (2007).
47. Fearn, T., Hill, D. C. & Darby, S. C. Measurement error in the explanatory variable of a binary regression: Regression calibration and integrated conditional likelihood in studies of residential radon and lung cancer. *Stat. Med.* **27**, 2159–2176. https://doi.org/10.1002/sim.3163 (2008).
48. Little, M. P. *et al.* Association of chromosome translocation rate with low dose occupational radiation exposures in U.S. radiologic technologists. *Radiat. Res.* **182**, 1–17. https://doi.org/10.1667/RR13413.1 (2014).
49. Richardson, S. & Gilks, W. R. A Bayesian approach to measurement error problems in epidemiology using conditional independence models. *Am. J. Epidemiol.* **138**, 430–442 (1993).
50. Richardson, S. & Gilks, W. R. Conditional independence models for epidemiological studies with covariate measurement error. *Stat. Med.* **12**, 1703–1722 (1993).
51. Clayton, D. The analysis of event history data: A review of progress and outstanding problems. *Stat. Med.* **7**, 819–841 (1988).
52. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 (1992).
53. Brooks, S. P. & Gelman, A. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* **7**, 434–455. https://doi.org/10.2307/1390675 (1998).
54. Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. *OpenBUGS version 3.2.3*, http://www.openbugs.net/w/FrontPage (2016).
55. rjags. Bayesian graphical models using MCMC. Version 4-13 (CRAN - The Comprehensive R Archive Network, 2022).
56. Cook, J. R. & Stefanski, L. A. Simulation-extrapolation estimation in parametric measurement error models. *J. Am. Stat. Assoc.* **89**, 1314–1328. https://doi.org/10.2307/2290994 (1994).
57. simex. Version 1.8 (CRAN - The Comprehensive R Archive Network, 2019).
58. Land, C. E. *et al.* Accounting for shared and unshared dosimetric uncertainties in the dose response for ultrasound-detected thyroid nodules after exposure to radioactive fallout. *Radiat. Res.* **183**, 159–173. https://doi.org/10.1667/RR13794.1 (2015).
59. Claeskens, G. & Hjort, N. L. *Cambridge Series in Statistical and Probabilistic Mathematics*. 1–312 (Cambridge University Press, 2008).
60. Dormann, C. F. *et al.* Model averaging in ecology: A review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecol. Monogr.* **88**, 485–504. https://doi.org/10.1002/ecm.1309 (2018).
61. Gelfand, A. E. & Dey, D. K. Bayesian model choice: Asymptotics and exact calculations. *J. R. Stat. Soc. Ser. B* **56**, 501–514 (1994).
62. Zhang, Z. *et al.* Correction of confidence intervals in excess relative risk models using Monte Carlo dosimetry systems with shared errors. *PLoS ONE* **12**, e0174641. https://doi.org/10.1371/journal.pone.0174641 (2017).
63. Stram, D. O. *et al.* Lung cancer in the Mayak workers cohort: Risk estimation and uncertainty analysis. *Radiat. Res.* **195**, 334–346. https://doi.org/10.1667/RADE-20-00094.1 (2021).
64. Little, M. P., Patel, A., Hamada, N. & Albert, P. Analysis of cataract in relationship to occupational radiation dose accounting for dosimetric uncertainties in a cohort of U.S. radiologic technologists. *Radiat. Res.* **194**, 153–161. https://doi.org/10.1667/RR15529.1 (2020).
65. Freedman, L. S., Fainberg, V., Kipnis, V., Midthune, D. & Carroll, R. J. A new method for dealing with measurement error in explanatory variables of regression models. *Biometrics* **60**, 172–181. https://doi.org/10.1111/j.0006-341X.2004.00164.x (2004).
66. Thomas, L., Stefanski, L. A. & Davidian, M. Moment adjusted imputation for multivariate measurement error data with applications to logistic regression. *Comput. Stat. Data Anal.* **67**, 15–24. https://doi.org/10.1016/j.csda.2013.04.017 (2013).
67. Thomas, L., Stefanski, L. & Davidian, M. A moment-adjusted imputation method for measurement error models. *Biometrics* **67**, 1461–1470. https://doi.org/10.1111/j.1541-0420.2011.01569.x (2011).

Acknowledgements

The Intramural Research Program of the National Institutes of Health, the National Cancer Institute, Division of Cancer Epidemiology and Genetics supported the work of MPL. The work of LBZ was supported by National Cancer Institute and National Institutes of Health (Grant No. R01CA197422). The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication. The authors are grateful for the detailed and helpful comments of the two referees.

Author contributions

M.P.L. devised the analytic strategy, wrote the software, performed the analysis and produced the first draft of the paper. L.B.Z. and N.H. contributed to subsequent drafting of the paper.

Funding
Open Access funding provided by the National Institutes of Health (NIH).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-42283-y.

Correspondence and requests for materials should be addressed to M.P.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.