

# Estimation of radiation risk in presence of classical additive and Berkson multiplicative errors in exposure doses

S. V. MASIUK\*

*State Institution “National Research Center for Radiation Medicine of the National Academy of Medical Sciences of Ukraine”, Melnykova str., 53, Kyiv, 04050, Ukraine; Ukrainian Radiation Protection Institute, Melnykova str., 53, Kyiv, 04050, Ukraine*

S. V. SHKLYAR, A. G. KUKUSH

*Taras Shevchenko National University of Kyiv, Volodymyrska Str. 64, Kyiv 01601, Ukraine*

R. J. CARROLL

*Texas A&M University, College Station, TX 77843, USA*

L. N. KOVGAN

*State Institution “National Research Center for Radiation Medicine of the National Academy of Medical Sciences of Ukraine”, Melnykova str., 53, Kyiv, 04050, Ukraine; Ukrainian Radiation Protection Institute, Melnykova str., 53, Kyiv, 04050, Ukraine*

I. A. LIKHTAROV

*State Institution “National Research Center for Radiation Medicine of the National Academy of Medical Sciences of Ukraine”, Melnykova str., 53, Kyiv, 04050, Ukraine; Ukrainian Radiation Protection Institute, Melnykova str., 53, Kyiv, 04050, Ukraine*

masja1979@gmail.com

## SUMMARY

In this paper, the influence of measurement errors in exposure doses in a regression model with binary response is studied. Recently, it has been recognized that uncertainty in exposure dose is characterized by errors of two types: classical additive errors and Berkson multiplicative errors. The combination of classical additive and Berkson multiplicative errors has not been considered in the literature previously. In a simulation study based on data from radio-epidemiological research of thyroid cancer in Ukraine caused by the Chernobyl accident, it is shown that ignoring measurement errors in doses leads to overestimation of background prevalence and underestimation of excess relative risk. In the work, several methods to reduce these biases are proposed. They are new regression calibration, an additive version of efficient SIMEX, and novel corrected score methods.

*Keywords:* Berkson measurement error; Chernobyl; Classical measurement error; Corrected scores; Dose-response; Radiation epidemiology; Regression calibration; SIMEX.

\*To whom correspondence should be addressed.

## 1. INTRODUCTION

As a result of the 1986 Chernobyl accident, significant territory of Ukraine, Russia, and Belarus were under radioactive contamination and the inhabitants of that territories suffered from radioactive exposure.

Even 5–6 years after the accident, an inflation of the incidence of thyroid cancer cases was observed for children and adolescents who lived in the territories where the estimated thyroid exposure doses were quite high, see [Likharev, Sobolev and others \(1995b\)](#), [Jacob and other \(2006\)](#), and [Buglova and others \(1996\)](#).

In fact, the growth of thyroid cancer prevalence for children and adolescents caused by internal irradiation from Chernobyl fallouts turned out to be the main (if not the unique) statistically reliable effect of the Chernobyl accident. Consequently this effect was of great interest for radiation epidemiologists all over the world, leading to a series of studies in Ukraine, Belarus and Russia, see [Likharev, Kovgan, Vavilov, Chepurny, Ron and others \(2006\)](#), [Kopecky and other \(2006\)](#), and [Zablotska and other \(2011\)](#).

However, interpretation of the results for most of the radiation epidemiological studies was based on risk estimation methods which do not take into account the presence of significant uncertainties in doses. One of the consequences of the assumption about the absence of errors in doses can be that the risk estimates are biased and the dose-response curve is distorted. The reasons for risk estimates distortions are not only systematic but also due to random errors in the dose estimates. In radiation epidemiology, various attempts have been made to construct statistical methods for analyzing not only uncertainty in the effect of the dose but also uncertainty in the dose itself, see [Mallick and other \(2002\)](#), [Carroll and other \(2006\)](#), [Lyon and other \(2006\)](#), [Kopecky and other \(2006\)](#), [Li and other \(2007\)](#), [Hofer \(2008\)](#), [Kukush and other \(2011\)](#), and [Likharev, Kovgan, Masiuk and others \(2014\)](#). The literature now recognizes that dose measurements are inevitably affected by errors of either classical or Berkson type, or a combination of the two, see [Mallick and other \(2002\)](#). Unfortunately, the most popular computer package in radiation epidemiology, EPICURE ([Preston and other, 1993](#)) does not account for dose uncertainty.

Previous attempts at dose–response estimation while accounting for uncertainties in doses have almost exclusively treated the dose uncertainties as multiplicative in structure. However, in the Chernobyl accident, recent detailed analyses of radioactivity registration mechanisms have shown that classical errors in thyroid exposure doses that were reconstructed in [Likharev, Kovgan, Vavilov, Chepurny, Bouville and others \(2005\)](#), and [Likharev, Kovgan, Masiuk and others \(2014\)](#) are of additive rather than multiplicative type, see [Likharev, Masiuk and others \(2013\)](#). In addition, [Likharev, Masiuk and others \(2013\)](#) show that thyroid radioactivity registration errors have a Poisson distribution. Because in most cases the intensity of measurements was quite high ([Likharev, Prohl and others, 1993](#); [Likharev, Goulko and others, 1995a](#)), the exposure dose measurement errors can be regarded normally distributed, although heteroscedastic, see [Likharev, Masiuk and others \(2013\)](#).

The aim of the present paper is to study radiation risk estimates and methods of risk estimation in models with *additive measurement errors* and *multiplicative Berkson errors* in exposure doses. In Section 2, we present the measurement error model and the risk model. In Section 3, we note that standard methods perform poorly in our context, and we develop three new methods: (a) a novel version of Corrected Scores, (b) a new version of Regression Calibration, and (c) a new version of efficient SIMEX (see [Cook and Stefanski, 1994](#); [Carroll and other, 2006](#); [Kukush and other, 2011](#)). Section 4 presents results of simulation studies, while Section 5 has concluding remarks. Technical details are given in Appendices of supplementary material available at *Biostatistics* online.

## 2. MODELS

2.1 *Model of dose with classical additive and Berkson multiplicative errors*

In May and June 1986, >150 000 measurements of thyroid radioactivity were made among inhabitants of the northern part of Ukraine, which suffered from the most intensive radionuclide fallouts, including

115 000 measurements among children and adolescents aged 0–18 years (Likhtarev, Prohl and others, 1993; Likhtarev, Goulko and others, 1995a). Further, the measurements will be denoted  $Q_i^{\text{mes}}$ . In what follows, a superscript “mes” refers to measured versions of the true variables, and a superscript “tr” refers to the true variables. Here  $i = 1, \dots, N$  denotes an individual.

As shown in Likhtarov, Kovgan, Vavilov, Chepurny, Bouville and others (2005) and Likhtarov, Kovgan, Masiuk and others (2014), the measured individual thyroid dose for the  $i$ th person can be written as

$$D_i^{\text{mes}} = f_i^{\text{mes}} Q_i^{\text{mes}} / M_i^{\text{mes}}, \quad (2.1)$$

where  $M_i^{\text{mes}}$  is the measured thyroid mass,  $f_i^{\text{mes}}$  is a factor that is obtained from the ecological model of radioactivity transition, and  $Q_i^{\text{mes}}$  is the measured radioactivity in the thyroid.

Ecological coefficient  $f_i^{\text{mes}}$  includes the error of Berkson type, see Likhtarov, Kovgan, Masiuk and others (2014). Denote the factor with Berkson error  $f_i^{\text{mes}} / M_i^{\text{mes}} = F_i^{\text{mes}}$ , so that (2.1) becomes

$$D_i^{\text{mes}} = F_i^{\text{mes}} Q_i^{\text{mes}}. \quad (2.2)$$

The true dose is decomposed as

$$D_i^{\text{tr}} = F_i^{\text{tr}} Q_i^{\text{tr}}.$$

Here the relation between  $F_i^{\text{tr}}$  and  $F_i^{\text{mes}}$  includes multiplicative Berkson error of the form  $F_i^{\text{tr}} = F_i^{\text{mes}} \delta_{F,i}$ , where  $E(\delta_{F,i}) = 1$  and  $\log(\delta_{F,i}) \sim \text{Normal}(-\sigma_{F,i}^2/2, \sigma_{F,i}^2)$ , where  $\sigma_{F,i}^2$  is known. The variables  $F_i^{\text{mes}}$  and  $\delta_{F,i}$  are stochastically independent, for details, see Eq. (8) in Kukush and other (2011). The empirical distribution of  $F_i^{\text{tr}}$  and its characteristics (expectation, variance, etc.) can be obtained by the Monte-Carlo procedure described in Likhtarov, Kovgan, Masiuk and others (2014).

As shown in Likhtarov, Masiuk and others (2013), radioactivity measurements of the thyroid are now known to have additive error, so that  $Q_i^{\text{mes}}$ , the measured thyroid radioactivity, is

$$Q_i^{\text{mes}} = Q_i^{\text{tr}} + \sigma_{Q,i}^{\text{mes}} \gamma_i, \quad (2.3)$$

where the  $\gamma_i$  are independent standard normal variables, the value  $\sigma_{Q,i}^{\text{mes}}$  is known and  $Q_i^{\text{tr}}$  are independent random variables.

Plug (2.3) into (2.2) and set  $\bar{D}_i^{\text{tr}} = F_i^{\text{mes}} Q_i^{\text{tr}}$ . We get

$$D_i^{\text{mes}} = F_i^{\text{mes}} Q_i^{\text{mes}} = F_i^{\text{mes}} (Q_i^{\text{tr}} + \sigma_{Q,i}^{\text{mes}} \gamma_i) = \bar{D}_i^{\text{tr}} + F_i^{\text{mes}} \sigma_{Q,i}^{\text{mes}} \gamma_i. \quad (2.4)$$

The random variables  $\{\delta_{F,i}\}$ ,  $\{\gamma_i\}$ , and  $\{(F_i^{\text{mes}}, Q_i^{\text{tr}})\}$  are jointly independent, although we allow correlation between  $F_i^{\text{mes}}$  and  $Q_i^{\text{tr}}$ . Define  $\sigma_i = F_i^{\text{mes}} \sigma_{Q,i}^{\text{mes}}$ , then (2.4) takes a form

$$D_i^{\text{mes}} = \bar{D}_i^{\text{tr}} + \sigma_i \gamma_i, \quad (2.5)$$

$$D_i^{\text{tr}} = \bar{D}_i^{\text{tr}} \delta_{F,i}. \quad (2.6)$$

Actually, (2.5) and (2.6) are a model of dose observations with additive classical and multiplicative Berkson errors. It is straightforward to see that  $E[\bar{D}_i^{\text{tr}} | D_i^{\text{mes}}] = E[D_i^{\text{tr}} | D_i^{\text{mes}}]$ .

## 2.2 Prevalence model

In order to model cases of cancer for a fixed time interval, we use a model of rare events with binary response variable  $Y_i$ , where  $Y_i = 1$  in the case of thyroid cancer and  $Y_i = 0$  in the absence of disease.

Define  $\lambda_0$  to be background prevalence intensity, i.e., in the absence of dose, and define  $\theta = (\lambda_0, \beta)'$ . Then define

$$\lambda(D_i^{\text{tr}}, \theta) = \lambda_i = \lambda_0(1 + \beta D_i^{\text{tr}}) = \lambda_0 + \text{EAR} \cdot D_i^{\text{tr}}, \quad (2.7)$$

where EAR is excess absolute risk. Then the conditional distribution of  $Y_i$  given the exposure dose is defined by

$$P[Y_i = 1 | D_i^{\text{tr}}] = \frac{\lambda_i}{1 + \lambda_i}, \quad P[Y_i = 0 | D_i^{\text{tr}}] = \frac{1}{1 + \lambda_i}. \quad (2.8)$$

The observed sample consists of couples  $(Y_i, D_i^{\text{mes}})$ , for  $i = 1, \dots, N$ . The parameters  $\lambda_0$  and  $\beta$  (or, in other parameterization,  $\lambda_0$  and EAR), are positive and to be estimated.

### 3. METHODS

#### 3.1 Existing methods

Common methods include (a) the naïve estimator, which is maximum likelihood estimator not accounting for measurement errors in doses; (b) parametric and linear regression calibration as defined in Appendix A of supplementary material available at *Biostatistics* online; and (c) the ordinary SIMEX method (Cook and Stefanski, 1994; Carroll and other, 2006). The simulation results show that methods (a) and (b) yield estimates with significant bias, see Appendix A of supplementary material available at *Biostatistics* online. This can be explained by specific structure of the data problem, where we have a kind of mixture of lognormal and normal variables. The ordinary SIMEX has larger bias compared with the efficient SIMEX, see Kukush and other (2011). Instead, we developed three new methods described in Sections 3.2–3.5.

#### 3.2 Corrected Score estimator

Within the Corrected Score method, we adjust the unbiased estimating function to measurement errors (Carroll and other, 2006, Section 7.4). Introduce the estimating function  $\tilde{S}_C$  as a solution to the deconvolution problem

$$E[\tilde{S}_C(Y, D^{\text{mes}}; \lambda_0, \beta) | Y, D^{\text{tr}}] = \tilde{S}_{\text{ML}}(Y, D^{\text{tr}}; \lambda_0, \beta),$$

where  $S_{\text{ML}}$  is an unbiased estimating functions, see Appendix B of supplementary material available at *Biostatistics* online;  $\tilde{S}_{\text{ML}}$  is a product of a matrix and a vector

$$\tilde{S}_{\text{ML}}(Y, D^{\text{tr}}; \lambda_0, \beta) = \begin{pmatrix} 1 & \beta \\ 0 & \lambda_0 \end{pmatrix} \begin{pmatrix} (Y-1)\lambda_0(1 + D^{\text{tr}}\beta) + Y \\ (Y-1)\lambda_0(D^{\text{tr}} + (D^{\text{tr}})^2\beta) + YD^{\text{tr}} \end{pmatrix}. \quad (3.1)$$

The explicit expression for  $\tilde{S}_C$  is

$$\tilde{S}_C(Y, D^{\text{mes}}; \lambda_0, \beta) = \begin{pmatrix} 1 & \beta \\ 0 & \lambda_0 \end{pmatrix} \begin{pmatrix} (Y-1)\lambda_0(1 + D^{\text{mes}}\beta) + Y \\ (Y-1)\lambda_0(D^{\text{mes}} + ((D^{\text{mes}})^2 - \sigma^2)\beta) + YD^{\text{mes}} \end{pmatrix}. \quad (3.2)$$

A consistent estimator of  $\theta$  is a solution  $\hat{\theta}_N$  to an unbiased estimating equation, namely a solution to

$$\sum_{i=1}^N \tilde{S}_C(Y_i, D_i^{\text{mes}}; \theta) = 0. \quad (3.3)$$

Equation (3.3) is linear in  $\lambda_0$  and  $\lambda_0\beta$ , and therefore, it can be solved efficiently.

In Appendix B of supplementary material available at *Biostatistics* online, we establish the asymptotic normality of  $\hat{\theta}_N$ , and construct a data-based covariance matrix estimator.

### 3.3 New regression calibration handling Berkson error

As mentioned in Section 3.1, the conventional parametric regression calibration has quite poor behavior in our simulation studies. In this section, we develop an approximation to regression calibration that has much more satisfactory behavior.

The idea is to treat the additive normal error in dose (2.5) as if it was multiplicative log-normal error, but with approximately the same conditional variance of  $D_i^{\text{mes}}$  given  $\bar{D}_i^{\text{tr}}$ . Denote the log-normal error by  $\delta_{L,i}$ ,  $\log(\delta_{L,i}) \sim \text{Normal}(0, \sigma_{L,i}^2)$ . Equating the variance of the multiplicative error  $\delta_{L,i}$  to the relative variance  $\sigma_i^2/(\bar{D}_i^{\text{tr}})^2$  and replacing the unknown  $\bar{D}_i^{\text{tr}}$  with a feasible value  $D_i^{\text{mes}}$ , we obtain

$$\text{var}(\delta_{L,i}) = \exp(2\sigma_{L,i}^2) - \exp(\sigma_{L,i}^2) = (\sigma_i/D_i^{\text{mes}})^2.$$

This yields

$$\sigma_{L,i}^2 = \log \left\{ (1/2) + \sqrt{(1/4) + (\sigma_i/D_i^{\text{mes}})^2} \right\}.$$

Then calibration is performed in the same manner as described in Kukush *and other* (2011), namely

$$E(D_i^{\text{tr}} | D_i^{\text{mes}}) \approx \exp \left\{ \frac{\sigma_{\bar{D}^{\text{tr}}}^2 \log D_i^{\text{mes}} + \sigma_{L,i}^2 \mu_{\bar{D}^{\text{tr}}} + \frac{1}{2} \sigma_{\bar{D}^{\text{tr}}}^2 \sigma_{L,i}^2}{\sigma_{\bar{D}^{\text{tr}}}^2 + \sigma_{L,i}^2} \right\}.$$

Here the estimators of  $\mu_{\bar{D}^{\text{tr}}}$  and  $\sigma_{\bar{D}^{\text{tr}}}^2$  are taken from Likhtarov, Masiuk *and others* (2013), namely

$$\hat{\mu}_{\bar{D}^{\text{tr}}} = \log\{(\hat{m}_{\bar{D}^{\text{tr}}})^2 / \sqrt{\hat{v}_{\bar{D}^{\text{tr}}} + (\hat{m}_{\bar{D}^{\text{tr}}})^2}\}, \quad (3.4)$$

$$\hat{\sigma}_{\bar{D}^{\text{tr}}}^2 = \log\{(\hat{v}_{\bar{D}^{\text{tr}}}/(\hat{m}_{\bar{D}^{\text{tr}}})^2) + 1\}, \quad (3.5)$$

where

$$\hat{m}_{\bar{D}^{\text{tr}}} = \frac{1}{N} \sum_{i=1}^N D_i^{\text{mes}},$$

$$\hat{v}_{\bar{D}^{\text{tr}}} = \frac{1}{N-1} \sum_{i=1}^N (D_i^{\text{mes}} - \hat{m}_{\bar{D}^{\text{tr}}})^2 - \frac{1}{N} \sum_{i=1}^N \sigma_i^2.$$

After preliminary calibration of doses, the maximum likelihood method described in Masiuk *and other* (2013) is used for accounting for Berkson error, see Appendix C of supplementary material available at *Biostatistics* online.

### 3.4 Efficient SIMEX

As a prerequisite to classical SIMEX method, assume that we can evaluate an estimator  $\hat{\theta} = \hat{\Theta}(D_i^{\text{tr}}, Y_i)$ ,  $i = 1, 2, \dots, N$  in the model without measurement errors (e.g., the maximum likelihood estimator).

Classical SIMEX algorithm is described in Carroll *and other* (2006, Section 5). It consists of the following steps:

- (1) Select a “large” number  $B$  and a finite set of non-negative numbers  $\Lambda$ .
- (2) For all  $b = 1, 2, \dots, B$  and all  $i = 1, 2, \dots, N$ , generate normal random variables  $U_{b,i}^* \sim \text{Normal}(0, \sigma_i^2)$ , where  $\sigma_i$  comes from (2.5).

(3) For all  $b = 1, \dots, B$  and  $\kappa \in \Lambda$ , evaluate the naive estimator for perturbed data

$$\theta_b^*(\kappa) = \hat{\Theta}(D_i^{\text{mes}} + \sqrt{\kappa}U_{b,i}^*, Y_i, i = 1, 2, \dots, N), \quad b = 1, 2, \dots, B, \kappa \in \Lambda,$$

and evaluate averaged estimate

$$\theta^*(\kappa) = \frac{1}{B} \sum_{b=1}^B \theta_b^*(\kappa), \quad \kappa \in \Lambda.$$

(4) Extrapolate  $\theta^*(\kappa)$  to point  $-1$  and assign  $\hat{\theta}_{\text{SIMEX}} = \theta^*(-1)$ .

In [Kukush and other \(2011\)](#), the “efficient SIMEX estimator” of the risk parameters of the model with multiplicative error was derived as an alternative to the ordinary SIMEX. It differed in the way that  $D_i^{\text{mes}}$  is perturbed only if  $Y_i = 1$ . Here we develop this idea in the model with additive errors.

- (i) *Setting tuning parameters.* Select a “large” number  $B$  and a finite set of non-negative numbers  $\Lambda$ . We use  $B = 1000$  and  $\Lambda = \{0, 0.2, 0.4, 0.6\}$  in our numerical work.
- (ii) *Simulation.* For all  $b = 1, 2, \dots, B$  and all  $i$  such that  $Y_i = 1$ , generate normal random variables

$$U_{b,i}^* \sim \text{Normal}(0, \sigma_i^2).$$

As an optional refinement, generate them such that  $\sum_{b=1}^B U_{b,i}^* = 0$ .

- (iii) *Estimation.* For all  $b = 1, \dots, B$  and  $\kappa \in \Lambda$ , solve the system of equations in  $\beta$  and  $\lambda$

$$\sum_{i=1}^N (1 - Y_i)(1 + \beta D_i^{\text{mes}}) = \lambda^{-1} \sum_{i=1}^N Y_i, \quad (3.6)$$

$$\sum_{i=1}^N (1 - Y_i) = \lambda^{-1} \sum_{Y_i=1} \{1 + \beta \max(0, D_i^{\text{mes}} + \sqrt{\kappa}U_{b,i}^*)\}^{-1}. \quad (3.7)$$

The perturbed dose  $D_i^{\text{mes}} + \sqrt{\kappa}U_{b,i}^*$  can be negative, and significant negative doses break down the naive estimator. Therefore, we use the censored perturbed doses given by  $\max(0, D_i^{\text{mes}} + \sqrt{\kappa}U_{b,i}^*)$ . Denote the solution as  $\beta_b^*(\kappa) = \beta$ ,  $\lambda_{0,b}^*(\kappa) = \lambda_0$ .

For  $\kappa \in \Lambda$  average  $\lambda_{0,b}^*(\kappa)$  and  $\lambda_{0,b}^*(\kappa)\beta_b^*(\kappa)$  in  $b$ :

$$\lambda_0^*(\kappa) = \frac{1}{B} \sum_{b=1}^B \lambda_{0,b}^*(\kappa), \quad \text{EAR}^*(\kappa) = \frac{1}{B} \sum_{b=1}^B \lambda_{0,b}^*(\kappa)\beta_b^*(\kappa).$$

- (iv) *Extrapolation.* Extrapolate numerically the functions  $\lambda_0^*(\kappa)$  and  $\text{EAR}^*(\kappa)$  to  $-1$ . In extrapolation, we approximate  $\lambda_0^*(\kappa)$  and  $\text{EAR}^*(\kappa)$  with quadratic polynomial. Such a choice of extrapolant function is the simplest one, and it allows to express the estimates explicitly through  $\lambda_0^*(\kappa)$  and  $\text{EAR}^*(\kappa)$ , see [Kukush and other \(2011\)](#).

The values  $\lambda_0^*(-1)$  and  $\text{EAR}^*(-1)$  are the efficient SIMEX estimates of  $\lambda_0$  and EAR.

### 3.5 Efficient SIMEX handling Berkson error

In this section, we introduce the SIMEX estimator which uses variances of both classical and Berkson errors. We start with unbiased estimating equation in the model with Berkson error only, see Appendix C.2

of supplementary material available at *Biostatistics* online. Assume for the moment that  $\bar{D}_i$  are known. Denote the conditional probability

$$m(\bar{D}; \lambda_0, \beta, \sigma_{F,i}^2) = P[Y_i = 1 \mid \bar{D}^{\text{tr}} = \bar{D}].$$

The following equations are unbiased:

$$\begin{aligned} \sum_{i=1}^N \frac{Y_i}{m(\bar{D}_i^{\text{tr}}; \lambda_0, \beta, \sigma_{F,i}^2)} &= N, \\ \sum_{i=1}^N \frac{Y_i \bar{D}_i^{\text{tr}}}{m(\bar{D}_i^{\text{tr}}; \lambda_0, \beta, \sigma_{F,i}^2)} &= \sum_{i=1}^N \bar{D}_i^{\text{tr}}, \end{aligned}$$

that is, with true parameters substituted, expectations of the left-hand and right-hand sides of these equations coincide.

With (2.6), the expression for  $m(\bar{D}; \lambda_0, \beta, \sigma_F^2)$  is

$$\begin{aligned} m(\bar{D}; \lambda_0, \beta, \sigma_F^2) &= E \frac{\lambda_0(1 + \beta \bar{D} \delta)}{1 + \lambda_0(1 + \beta \bar{D} \delta)} \\ &= \frac{1}{\sqrt{2\pi} \sigma_F} \int \frac{\lambda_0(1 + \beta \bar{D} \exp\{u - \frac{1}{2}\sigma_F^2\})}{1 + \lambda_0(1 + \beta \bar{D} \exp\{u - \frac{1}{2}\sigma_F^2\})} \exp\left\{\frac{-u^2}{2\sigma_F^2}\right\} du, \end{aligned} \quad (3.8)$$

where expectation is taken for nonrandom  $\bar{D}$  and lognormal  $\delta$ ,  $\log \delta \sim \text{Normal}(-\frac{1}{2}\sigma_F^2, \sigma_F^2)$ , see [Likhartov, Masiuk and others \(2013\)](#). In generic case  $\bar{D} > 0$ ,  $\lambda_0 > 0$ ,  $\beta > 0$ , and the integral in (3.8) is taken from  $-\infty$  to  $+\infty$ . In other case, we integrate over the interval where the numerator  $\lambda_0(1 + \beta \bar{D} \exp\{u - \frac{1}{2}\sigma_F^2\})$  is positive.

Now, consider the model with both Berkson and classical errors. In SIMEX method, perturbed measured doses are substituted for true doses. Therefore, substitute  $D_{b,i}^*(\kappa) = D_i^{\text{mes}} + \sqrt{\kappa} U_{b,i}^*$  for  $\bar{D}_i^{\text{tr}}$ :

$$\begin{aligned} \sum_{i=1}^N \frac{Y_i}{m(D_{b,i}^*(\kappa); \lambda_0, \beta, \sigma_{F,i}^2)} &= N, \\ \sum_{i=1}^N \frac{Y_i D_{b,i}^*(\kappa)}{m(D_{b,i}^*(\kappa); \lambda_0, \beta, \sigma_{F,i}^2)} &= \sum_{i=1}^N D_{b,i}^*(\kappa). \end{aligned}$$

Change the right-hand side of the second equation to  $\sum_{i=1}^N D_i^{\text{mes}}$ . This is equivalent to formal adding to the latter equation the unbiased equation  $\sum_{i=1}^N D_{b,i}^*(\kappa) = \sum_{i=1}^N D_i^{\text{mes}}$ ; the unbiasedness holds true because

$$\begin{aligned} \sum_{i=1}^N D_{b,i}^* - \sum_{i=1}^N D_i^{\text{mes}} &= \sqrt{\kappa} \sum_{i=1}^N U_{b,i}^*, \\ E \left[ \sum_{i=1}^N D_{b,i}^* - \sum_{i=1}^N D_i^{\text{mes}} \right] &= \sqrt{\kappa} E \sum_{i=1}^N U_{b,i}^* = 0. \end{aligned}$$

This simplification is done in order to avoid perturbations of doses for non-cases  $Y_i = 0$ . We get

$$\sum_{i=1}^N \frac{Y_i}{m(D_{b,i}^*(\kappa); \lambda, \beta, \sigma_{F,i}^2)} = N, \quad (3.9)$$

$$\sum_{i=1}^N \frac{Y_i D_{b,i}^*(\kappa)}{m(D_{b,i}^*(\kappa); \lambda, \beta, \sigma_{F,i}^2)} = \sum_{i=1}^N D_i^{\text{mes}}, \quad (3.10)$$

The efficient SIMEX estimator is defined similarly to the one in Section 3.4. We just replace equations (3.6) and (3.7) with (3.9) and (3.10).

For significant perturbations, the modified dose  $D_i^{\text{mes}} + \sqrt{\kappa} U_{b,i}^*$  may be negative, which may break down the estimation procedure. Therefore, the negative doses are changed to zeros, i.e.,  $D_{b,i}^*(\kappa) = \max\{0, D_i^{\text{mes}} + \sqrt{\kappa} U_{b,i}^*\}$  is used instead of  $D_{b,i}^*(\kappa) = D_i^{\text{mes}} + \sqrt{\kappa} U_{b,i}^*$ .

## 4. SIMULATION STUDY

### 4.1 Simulation setup

In order to simulate exposure doses, we used a real subpopulation of children and adolescents under 18, consisting of  $\sim 13\,000$  persons from the settlements of Zhytomyr, Kyiv, and Chernihiv, which had direct measurements of thyroid activity in May–June 1986. Exposure doses for this subpopulation were constructed via the framework of the Ukrainian-American project on thyroid cancer prevalence in Ukraine after the Chernobyl accident; see [Likhтарov, Kovgan, Masiuk and others \(2014\)](#).

Parameters of the absolute risk model (2.7) for the observation period from 1991 to 2001 were given by values close to ones obtained in epidemiological studies of thyroid cancer in Ukraine, see [Jacob and other \(2006\)](#) and [Likhтарov, Kovgan, Vavilov, Chepurny, Ron and others \(2006\)](#), namely

$$\lambda_0 = 2 \times 10^{-3} \frac{\text{cases}}{\text{persons}}, \quad \text{EAR} = 5 \times 10^{-3} \frac{\text{cases}}{\text{Gray} \times \text{persons}}. \quad (4.1)$$

In our simulation study, 1000 different data sets were simulated for different levels of classical ( $\delta_Q$ ) and Berkson ( $\sigma_F$ ) uncertainty. The classical error level was defined as the constant value  $\delta_Q = \sigma_{Q,i}^{\text{mes}} / Q_i^{\text{tr}}$  varied from 0.2 to 1. The Berkson error level was set in such a way that geometric standard deviation of  $F^{\text{tr}}$  given  $F^{\text{mes}}$ ,  $\text{GSD}_F = \exp(\sigma_F)$ , took on the values 1 (no error), 1.5, 2, 3, 5, and 8. All the listed values are realistic.

Simulation study is performed in four steps:

- (1) Initial doses  $\bar{D}_i^{\text{tr}}$  are taken from the real thyroid doses of children and adolescents internally exposed to  $^{131}\text{I}$  in 1986, see Figure 1.
- (2) True dose values are generated for the cohort by using  $\bar{D}_i^{\text{tr}}$  and taking into account the uncertainty levels  $\text{GSD}_F$  given in the first column of Tables 1 and 2, see (2.6).
- (3) Using the data from Step (2), as well as the model in equations (2.7) and (2.8), with the parameter values  $\lambda_0$  and EAR in (4.1), a disease vector is generated.
- (4) Initial doses  $\bar{D}_i^{\text{tr}}$  were perturbed, and thus, the measured doses  $D_i^{\text{mes}}$  were generated according to equation (2.5), with the error standard deviation  $\sigma_i = \delta_Q \bar{D}_i^{\text{tr}}$ , where  $\delta_Q$  enters the second column of Tables 1 and 2. As a result, we obtain an observation model with classical additive and Berkson multiplicative errors in doses.



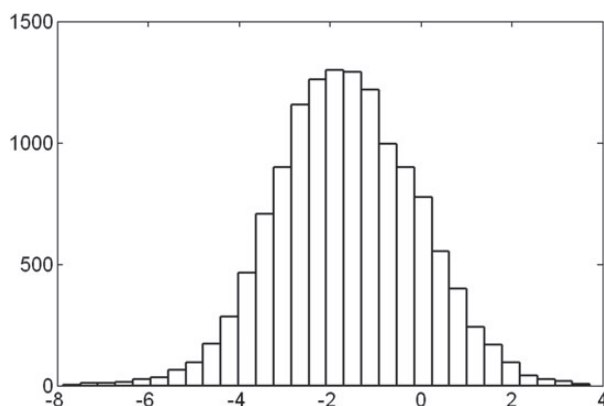


Fig. 1. Histogram of  $\log(\bar{D}^{\text{tr}})$ .

- (5) Based on the measured doses  $D_i^{\text{mes}}$ , the information of measurement errors  $\text{GSD}_F$  and  $\sigma_i$ , as well as the disease vector generated in Step (3), the parameter values  $\lambda_0$  and EAR are estimated by three methods.

Steps (1) to (5) are repeated 1000 times and the median values of the estimated risk coefficients as well as standard deviations are presented in Tables 1 and 2.

Sometimes measured doses  $D_i^{\text{mes}}$  can take negative values as a result of large errors in the additive error model (2.5). In such cases, negative doses were replaced by a small positive number, except for the Corrected Score estimator, because the Corrected Score method can handle negative doses.

For each of the various values of  $\text{GSD}_F$ , the averaged number of cases over 1000 realizations was 68, with corresponding frequency of thyroid cancer disease 0.51%.

## 4.2 Results and discussion

Estimation of absolute risk parameters was performed by the naïve method, the Corrected Score method presented in Section 3.2 that takes into account only classical error, and also by the new regression calibration method and the efficient SIMEX method described in Sections 3.3 and 3.5, respectively. The latter two methods take into account both classical and Berkson errors. Because in our case the distribution of data set  $\bar{D}^{\text{tr}}$  is strictly positive and its logarithm is approximately symmetric (see Figure 1), in our simulation any parametric method assumes a log-normal distribution of  $\bar{D}^{\text{tr}}$ .

The medians of the estimates of the baseline incidence rates and the standard deviations (SD) of the estimates are given in Table 1, while the medians of the estimates of the excess absolute risk and the standard deviations of the estimates are given in Table 2. In Appendix D of supplementary material available at *Biostatistics* online, we display 95% deviance intervals computed based on the obtained empirical distribution for risk parameters estimators with truncation of 2.5% quantiles from both sides, and hence an interval estimate for risk parameters.

**4.2.1 Naïve estimator:** The simulation results showed that the naïve method underestimates EAR and overestimates background prevalence intensity. The risk estimates have larger bias for larger measurement errors in doses. For  $\delta_Q = 1$ , EAR is underestimated twice. The level of uncertainty  $\delta_Q = 1$  for additive measurement errors in doses corresponds to the geometric standard deviation equal 2.3 for multiplicative errors. Comparison with results from [Kukush and other \(2011\)](#) shows reasonable consistency. It is worth

Table 1. *Estimates of baseline incidence rate (medians over 1000 simulations and standard deviations)*

Estimates of $\lambda_0 \times 10^3$ by different methods									
Error		Naïve		New calibrate handling Berkson error		Corrected Score		Efficient SIMEX handling Berkson error	
GSD <sub>F</sub>	$\delta_Q$	Median	SD	Median	SD	Median	SD	Median	SD
1 no error	0	1.95	(0.53)	1.95	(0.53)	1.93	(0.97)	1.94	(0.53)
	0.2	1.99	(0.54)	2.00	(0.55)	1.95	(1.01)	1.94	(0.55)
	0.4	2.20	(0.56)	2.11	(0.57)	1.98	(1.11)	1.93	(0.75)
	0.6	2.57	(0.59)	2.46	(0.58)	1.93	(1.28)	2.52	(1.11)
	0.8	2.91	(0.61)	2.79	(0.59)	1.90	(1.49)	3.52	(1.39)
	1	3.15	(0.62)	3.05	(0.60)	1.98	(1.80)	4.49	(1.59)
1.5	0	1.96	(0.55)	1.95	(0.55)	1.97	(0.97)	1.95	(0.55)
	0.2	2.01	(0.56)	2.01	(0.58)	1.97	(1.01)	1.95	(0.56)
	0.4	2.20	(0.58)	2.14	(0.60)	1.97	(1.12)	1.93	(0.73)
	0.6	2.57	(0.58)	2.46	(0.59)	2.00	(1.25)	2.40	(1.10)
	0.8	2.89	(0.59)	2.79	(0.59)	2.04	(1.48)	3.50	(1.44)
	1	3.11	(0.60)	3.06	(0.58)	2.06	(1.79)	4.43	(1.59)
2	0	1.95	(0.54)	1.94	(0.55)	2.05	(0.94)	1.93	(0.54)
	0.2	2.00	(0.54)	1.99	(0.56)	2.05	(1.02)	1.94	(0.54)
	0.4	2.21	(0.54)	2.14	(0.59)	2.05	(1.10)	1.95	(0.71)
	0.6	2.57	(0.58)	2.47	(0.59)	2.07	(1.24)	2.40	(1.13)
	0.8	2.90	(0.58)	2.80	(0.60)	2.06	(1.42)	3.43	(1.38)
	1	3.13	(0.59)	3.04	(0.58)	2.11	(1.67)	4.39	(1.56)
3	0	2.00	(0.56)	1.97	(0.57)	2.18	(0.94)	1.95	(0.56)
	0.2	2.04	(0.56)	2.02	(0.57)	2.20	(0.94)	1.95	(0.56)
	0.4	2.23	(0.56)	2.15	(0.60)	2.19	(1.03)	1.94	(0.74)
	0.6	2.60	(0.57)	2.42	(0.58)	2.20	(1.18)	2.43	(1.14)
	0.8	2.89	(0.59)	2.82	(0.61)	2.24	(1.34)	3.46	(1.34)
	1	3.12	(0.59)	3.06	(0.60)	2.23	(1.62)	4.38	(1.52)
5	0	2.12	(0.55)	2.03	(0.57)	2.37	(0.83)	1.94	(0.55)
	0.2	2.17	(0.56)	2.02	(0.59)	2.38	(0.88)	1.95	(0.56)
	0.4	2.34	(0.57)	2.12	(0.59)	2.39	(0.96)	1.95	(0.73)
	0.6	2.65	(0.57)	2.44	(0.58)	2.38	(1.08)	2.40	(1.07)
	0.8	2.94	(0.57)	2.75	(0.57)	2.39	(1.21)	3.40	(1.37)
	1	3.14	(0.57)	2.99	(0.57)	2.44	(1.46)	4.24	(1.46)
8	0	2.23	(0.56)	2.07	(0.56)	2.59	(0.76)	1.92	(0.57)
	0.2	2.27	(0.56)	2.04	(0.58)	2.58	(0.78)	1.93	(0.56)
	0.4	2.43	(0.57)	2.13	(0.58)	2.59	(0.85)	1.92	(0.76)
	0.6	2.71	(0.57)	2.45	(0.57)	2.59	(0.91)	2.39	(1.08)
	0.8	2.96	(0.57)	2.71	(0.56)	2.61	(1.03)	3.25	(1.38)
	1	3.13	(0.57)	2.93	(0.55)	2.64	(1.26)	4.04	(1.40)

True value  $\lambda_0 = 2 \times 10^{-3}$ .

Table 2. Estimates of absolute excess risk (medians over 1000 simulations and standard deviations)

Estimates of $\text{EAR} \times 10^3$ by different methods									
Error		Naïve		New calibrate handling Berkson error		Corrected Score		Efficient SIMEX handling Berkson error	
$\text{GSD}_F$	$\delta_Q$	Median	SD	Median	SD	Median	SD	Median	SD
1 no error	0	4.98	(1.03)	5.01	(1.03)	5.03	(1.58)	4.99	(1.04)
	0.2	4.93	(1.02)	5.05	(1.02)	4.97	(1.64)	5.00	(1.03)
	0.4	4.63	(1.01)	4.88	(1.03)	4.93	(1.76)	5.04	(1.27)
	0.6	4.01	(0.92)	4.07	(0.87)	4.98	(2.03)	4.07	(1.65)
	0.8	3.39	(0.84)	3.31	(0.80)	5.02	(2.42)	2.36	(1.90)
	1	2.91	(0.73)	2.80	(0.72)	5.02	(2.91)	0.84	(2.09)
1.5	0	4.98	(1.04)	4.99	(1.04)	5.00	(1.59)	4.99	(1.04)
	0.2	4.90	(1.02)	5.07	(0.99)	4.94	(1.65)	4.99	(1.04)
	0.4	4.57	(0.98)	4.90	(0.99)	4.93	(1.77)	5.04	(1.31)
	0.6	3.98	(0.88)	4.05	(0.92)	4.94	(2.03)	4.17	(1.73)
	0.8	3.38	(0.83)	3.34	(0.82)	4.94	(2.42)	2.46	(1.96)
	1	2.91	(0.76)	2.83	(0.71)	4.90	(2.86)	0.88	(2.13)
2	0	4.90	(1.06)	5.04	(1.06)	4.84	(1.59)	5.00	(1.07)
	0.2	4.84	(1.02)	5.05	(1.04)	4.84	(1.68)	5.01	(1.06)
	0.4	4.52	(0.99)	4.89	(1.05)	4.78	(1.85)	5.04	(1.34)
	0.6	3.92	(0.89)	4.03	(0.92)	4.82	(2.07)	4.13	(1.85)
	0.8	3.33	(0.85)	3.31	(0.79)	4.80	(2.35)	2.57	(2.02)
	1	2.87	(0.77)	2.80	(0.70)	4.79	(2.73)	0.98	(2.06)
3	0	4.75	(1.02)	5.15	(1.01)	4.54	(1.57)	5.03	(1.08)
	0.2	4.68	(1.01)	5.08	(1.07)	4.51	(1.64)	5.02	(1.09)
	0.4	4.36	(0.97)	4.89	(1.07)	4.48	(1.79)	5.03	(1.37)
	0.6	3.75	(0.90)	4.02	(1.34)	4.44	(1.96)	4.13	(1.90)
	0.8	3.19	(0.84)	3.30	(0.82)	4.40	(2.26)	2.37	(2.15)
	1	2.74	(0.76)	2.78	(0.73)	4.42	(2.71)	0.90	(2.18)
5	0	4.30	(0.99)	5.07	(0.99)	3.79	(1.42)	5.00	(1.22)
	0.2	4.24	(0.98)	5.13	(1.25)	3.80	(1.49)	4.98	(1.23)
	0.4	3.94	(0.96)	4.99	(1.26)	3.77	(1.57)	4.96	(1.49)
	0.6	3.37	(0.90)	4.04	(1.06)	3.73	(1.71)	4.07	(2.09)
	0.8	2.85	(0.81)	3.32	(0.94)	3.72	(1.95)	2.28	(2.38)
	1	2.50	(0.74)	2.80	(0.83)	3.71	(2.38)	0.79	(2.26)
8	0	3.64	(0.90)	5.15	(1.15)	2.99	(1.25)	4.98	(1.43)
	0.2	3.59	(0.89)	5.19	(1.48)	2.96	(1.29)	5.00	(1.44)
	0.4	3.30	(0.86)	4.98	(1.41)	2.89	(1.37)	4.94	(1.79)
	0.6	2.88	(0.79)	4.02	(1.22)	2.87	(1.55)	3.96	(2.21)
	0.8	2.41	(0.75)	3.30	(1.08)	2.89	(1.72)	2.23	(2.51)
	1	2.08	(0.67)	2.78	(1.95)	2.84	(2.04)	0.65	(2.51)

True value  $\text{EAR} = 5.0 \times 10^{-3}$ .

mentioning that for  $\delta_Q \leq 0.2$  and for  $\text{GSD}_F \leq 2$ , the bias of the background prevalence and the bias of EAR do not exceed 5%. Thus, for small level of uncertainty, the naïve method gives quite satisfactory results, as expected.

Nevertheless the effect of Berkson error on the results of risk analysis is much smaller. If  $\text{GSD}_F \leq 2$ , then the effect is negligible. When  $\text{GSD}_F$  is increasing up to 3 and more, then the bias of the estimate is more essential and should be taken into account.

**4.2.2 Regression calibration and efficient SIMEX.** Though parametric regression calibration defined in [Likhtarov, Masiuk and others \(2013\)](#) takes into account the shape of the distribution of  $\bar{D}^{\text{tr}}$ , the estimates computed by this method are considerably biased, with underestimated background prevalence intensity and overestimated of EAR (the results are shown in Appendix A of supplementary material available at *Biostatistics* online). This is unexpected effect compared with simulation results from [Kukush and other \(2011\)](#), where for multiplicative measurement errors in doses, the parametric estimates were quite acceptable. It looks like the reason for this is the structure of the normal measurement errors  $\sigma_i \gamma_i$  and the log-normal distribution of  $\bar{D}^{\text{tr}}$ , but we have no definite explanation.

Estimates obtained by the new regression calibration are much more stable and less biased compared with the ones obtained by other methods of regression calibration, and are quite satisfactory when the classical error in dose is not too large, in particular for  $\delta_Q \leq 0.4$ . However, when  $\delta_Q \geq 0.6$ , there is considerable bias.

Estimates of absolute risk parameters obtained by efficient SIMEX method fit the model values only for small classical errors. The estimates are satisfactory (that is bias does not exceed 10%) if  $\delta_Q \leq 0.4$ . However, when  $\delta_Q \geq 0.6$ , there is considerable bias.

These methods can handle quite large Berkson errors.

**4.2.3 Corrected Score method.** The Corrected Score estimator is the least biased of all ones presented in this paper. For the error-level  $\text{GSD}_F \leq 2$ , the maximal absolute bias for EAR and for  $\lambda_0$  does not exceed 5%. Of course, the Corrected Score estimator has the widest deviance intervals, reflecting the well-known phenomenon that bias correction typically leads to increased variability of estimates.

Using this estimator, only classical error in the factor  $Q^{\text{mes}}$  (see (2.3)) was taken into account. This leads to bias for large Berkson errors.

**4.2.4 Influence of Berkson error.** For moderate levels  $\text{GSD}_F \leq 2$ , the effect of Berkson error on ultimate estimates is insignificant. But if  $\text{GSD}_F$  increases to 3 and more, then the influence of Berkson error is indeed significant and should be taken into account. Simulation showed that in the naïve estimates the Berkson error, as well as the classical error but to a smaller extent, leads to underestimation of EAR and overestimation of  $\lambda_0$ .

## 5. CONCLUSIONS

There are classical additive errors and Berkson multiplicative errors in exposure doses in the linear model for rare events. That is a fact that requires a new statistical methodology. To solve this problem, we have developed new methods of regression calibration, corrected scores, and efficient SIMEX that are appropriate for the actual dose uncertainties. We performed simulations based on real data from epidemiological studies. The thyroid absorbed doses were taken from the results of Ukrainian–American project involving the Chernobyl accident, and cases were modeled based on the underlying risk model. The true absolute

risk parameters were chosen to be typical for the epidemiological studies in this important context. Estimators of the parameters were constructed by the naïve method (that is without taking into account dose measurement errors) with the package EPICURE and also by the methods mentioned above.

We showed that the naïve estimator has significant bias. The bias increases as the classical or Berkson error variance increases. The efficient SIMEX and new regression calibration approaches improve the estimators, but mainly for moderate classical uncertainty levels such as  $\delta_Q \leq 0.4$ . They give quite good result for significant Berkson error. The new Corrected Score estimator has little bias for small Berkson errors. However, this estimator has the largest deviance intervals, and it does not take Berkson error into account.

In general, methods of radiation risk estimation in cases of the classical additive dose error work more poorly than in case of the classical multiplicative error (Kukush and other, 2011). At first glance the reason is as follows: the size  $N$  of underlying cohort in the latter paper is larger, namely  $N$  around 70 000 persons vs.  $N$  around 13 000 persons in the present paper. However, additional simulations showed that in this case artificial enlargement of the sample size does not significantly improve the risk estimates. Therefore, we believe that this phenomenon has to do with the combination of normal dose errors  $\sigma_i \gamma_i$  and lognormally distributed random variables  $\bar{D}_i^R$ . This assertion is confirmed by other investigations we have done but that are not reported in the present paper.

Choosing among the methods, other than the naïve estimate which is clearly unacceptable, is difficult. However, for a concrete radiation risk estimation problem, it is reasonable to perform a preliminary simulation study. Such a simulation will make it possible, for a given dose distribution and prevalence level, to analyze the behavior of estimates obtained by various methods and also the influence of nuisance parameters on the model, such as effect modifiers and confounders, see [Health Risks from Exposure to Low Levels of Ionizing Radiation \(2006\)](#).

#### ACKNOWLEDGMENTS

*Conflict of Interest:* None declared.

#### FUNDING

The research was supported by the Ukrainian Radiation Protection Institute. Carroll's research was supported by a grant from the National Cancer Institute (R37-CA057030). This publication is based in part on work supported by Award Number KUS-CI-016-04, made by King Abdullah University of Science and Technology (KAUST).

#### SUPPLEMENTARY MATERIAL

Supplementary material is available at <http://biostatistics.oxfordjournals.org>.

#### REFERENCES

- BUGLOVA, E. E., KENIGSBERG, J. E. AND SERGEEVA, N. V. (1996). Cancer risk estimation in Belarussian children due to thyroid irradiation as a consequence of the Chernobyl nuclear accident. *Health Physics* **71**, 45–49.
- CARROLL, R. J., RUPPERT, D., STEFANSKI, L. A. AND CRAINICEANU, C. A. (2006) *Measurement Error in Nonlinear Models. A Modern Perspective*, 2nd edition. Boca Raton: Chapman and Hall/CRC.
- COOK, J. R. AND STEFANSKI, L. A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association* **89**, 1314–1328.

- Health Risks from Exposure to Low Levels of Ionizing Radiation (BEIR VII Phase 2, 2006). Washington: National Academy Press.
- HOFER, E. (2008). How to account for uncertainty due to measurement errors in the uncertainty analysis using Monte-Carlo simulation. *Health Physics* **95**, 277–290.
- JACOB, P., BOGDANOVA, T. I., BUGLOVA, E., CHEPURNY, M., DEMIDCHIK, Y., GAVRILIN, Y., KENIGSBURG, J., MECKBACH, R., SCHOTOLA, C., SHINKAREV, S. *and others* (2006). Thyroid cancer risk in areas of Ukraine and Belarus affected by the Chernobyl accident. *Radiation Research* **165**, 1–8.
- KOPECKY, K. J., STEPANENKO, V., RIVKIND, N., VOILLEQUE, P., ONSTAD, L., TROSHIN, V., ROMANOVA, G., DOROSHENKO, V., PROSHIN, A., TSYB, A. AND DAVIS, S. (2006). Childhood thyroid cancer, radiation dose from Chernobyl and dose uncertainties in Bryansk Oblast, Russia: A population-based case-control study. *Radiation Research* **166**, 367–374.
- KUKUSH, A., SHKLYAR, S., MASIUK, S., LIKHTAROV, I., KOVGAN, L., CARROLL, R. J. AND BOUVILLE, A. (2011). Methods for estimation of radiation risk in epidemiological studies accounting for classical and Berkson errors in doses. *The International Journal of Biostatistics* **7**(1, Article 15). doi:10.2202/1557-4679.1281
- LIKHTAREV, I. A., GOULKO, G. M., SOBOLEV, B. G., KAIRO, I. A., PROHL, G., RATH, P. AND HENRICH, K. (1995a). Evaluation of the  $^{131}\text{I}$  thyroid-monitoring measurements performed in Ukraine during May and June of 1986. *Health Physics* **69**, 6–15.
- LIKHTAROV, I., KOVGAN, L., MASIUK, S., TALERKO, M., CHEPURNY, M., IVANOVA, O., GERASYMENKO, V., BOYKO, Z., VOILLEQUÉ, P., DROZDOVITCH, V. AND BOUVILLE, A. (2014). Thyroid cancer study among Ukrainian children exposed to radiation after the Chornobyl accident: improved estimates of the thyroid doses to the cohort members. *Health Physics* **106**, 370–396.
- LIKHTAROV, I., KOVGAN, L., VAVILOV, S., CHEPURNY, M., BOUVILLE, A., LUCKYANOV, N., JACOB, P., VOILLEQUÉ, P. AND VOIGT, G. (2005). Post-Chornobyl thyroid cancers in Ukraine. Report 1: estimation of thyroid doses. *Radiation Research* **163**, 125–136.
- LIKHTAROV, I., KOVGAN, L., VAVILOV, S., CHEPURNY, M., RON, E., LUBIN, J., BOUVILLE, A., TRONKO, N., BOGDANOVA, T., GULAK, L., ZABLITSKA, L. AND HOWE, G. (2006). Post-Chornobyl thyroid cancers in Ukraine. Report 2: Risk analysis. *Radiation Research* **166**, 375–386.
- LIKHTAROV, I., MASIUK, S., CHEPURNY, M., KUKUSH, A., SHKLYAR, S., BOUVILLE, A. AND KOVGAN, L. (2013). Error estimation for direct measurements in May–June 1986 of  $^{131}\text{I}$  radioactivity in thyroid gland of children and adolescents and their registration in risk analysis. In Antoniouk, A. and Melnik, R. (editors), *Mathematics and Life Sciences*. Berlin/Boston: Walter de Gruyter GmbH, pp. 231–244.
- LIKHTAREV, I. A., PROHL, G. AND HENRICH, K. (1993). Reliability and accuracy of the  $^{131}\text{I}$  thyroid activity measurements performed in the Ukraine after the Chernobyl accident in 1986. *GSF-Bericht 19/93*, Institut für Strahlenschutz, Munich.
- LIKHTAREV, I. A., SOBOLEV, B. G., KAIRO, I. A., TRONKO, N. D., BOGDANOVA, T. I., OLEINIC, V. A., EPSHTEIN, E. V. AND BERAL, V. (1995b). Thyroid cancer in the Ukraine. *Nature* **375**, 365–378.
- LI, Y., GUOLO, A., HOFFMAN, F. O. AND CARROLL, R. J. (2007). Shared uncertainty in measurement error problems, with application to Nevada Test Site fallout data. *Biometrics* **63**, 1226–1236.
- LYON, J. L., ALDER, S. C., STONE, M. B., SCHOLL, A., READING, J. C., HOLUBKOV, R., SHENG, X., WHITE, G. L., HEGMANN, K. T., ANSPAUGH, L. *and others* (2006). Thyroid disease associated with exposure to the Nevada Test Site radiation: a reevaluation based on corrected dosimetry and examination data. *Epidemiology* **17**, 604–614.
- MALLICK, B., HOFFMAN, F. O. AND CARROLL, R. J. (2002). Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada Test Site. *Biometrics* **58**, 13–20.
- MASIUK, S. V., SHKLYAR, S. V. AND KUKUSH, A. G. (2013). Berkson errors in radiation dose assessments and their impact on radiation risk estimates. *Problems of Radiation Medicine and Radiobiology* **18**, 119–126.

PRESTON, D. L., LUBIN, J. H., PIERCE, D. A. AND MCCONNEY, M. E. (1993) *EPICURE User's Guide*. Seattle, Washington: Hirosoft Corporation.

ZABLOTSKA, L. B., RON, E., ROZHKO, A., HATCH, M., POLYANSKAYA, O. N., BRENNER, A. V., LUBIN, J., ROMANOV, G. N., MCCONNELL, R. J., O'KANE, P. *and others* (2011). Thyroid cancer risk in Belarus among children and adolescents exposed to radioiodine after the Chernobyl accident. *British Journal of Cancer* **104**, 181–187.

[Received April 23, 2015; revised December 2, 2015; accepted for publication December 2, 2015]