10

Statistical Modeling in Life Sciences and Direct Measurements

Illya Likhtarov, Sergii Masiuk, Mykola Chepurny, Alexander Kukush, Sergiy Shklyar, Andre Bouville and Lina Kovgan

10.1 Error Estimation for Direct Measurements in May–June 1986 of ^{131}I Radioactivity in Thyroid Gland of Children and Adolescents and their Registration in Risk Analysis

Abstract. A statistical model of thyroid gland radioactivity measurements is proposed. The measurement error is of classical type and heteroscedastic. Its variance can be reliably estimated. A model of thyroid exposure dose is constructed that involves both classical and Berkson errors. Two methods are proposed to deal with dose uncertainty in risk analysis: (a) parametric calibration, where the true doses are assumed log-normally distributed, and (b) nonparametric calibration, where the form of dose distribution is not specified.

Keywords. Berkson Measurement Error, Classical Measurement Error, Exposure Dose Uncertainty, Radioactivity Measurement, Regression Calibration, Thyroid Cancer

2010 Mathematics Subject Classification. 62P10

10.1.1 Introduction

In May and June 1986 millions of inhabitants of Ukraine, Belarus and Russia suffered from exposure to radiation fallout caused by the Chornobyl accident. The most substantial was the thyroid exposure as a result of iodine radioisotopes fallout, first of all ^{131}I (cf. Likhtarev et al. [9, 13, 15]).

Already within 5–6 years after the accident, sharp increasing in thyroid cancer cases were observed for children and adolescents who lived in the territories where the estimated exposure doses for this organ were quite large [1, 5, 12, 25]. In fact, increasing thyroid cancer prevalence for children and adolescents was caused by inner exposure to Chornobyl fallouts. This was the main statistically significant remote effect

* This work was supported by the Ukrainian Radiation Protection Institute of the Academy of Technological Science and by the Scientific Center for Radiation Medicine of the Academy of Medical Science of Ukraine, and by the intramural program of the U.S. National Cancer Institute, NIH, DHHS.

of the Chornobyl accident. It is not surprising that this phenomenon was of great interest for radioepidemiologists all over the world that resulted in a series of research in Ukraine, Belarus and Russia [6, 14, 24, 26]. Utmost interest to this problem is also due to the fact that there was quite extensive and reliable information about risk of radio-indicated thyroid cancer as a result of external exposure acting on this organ (cf. Ron et al. [22]). Concerning internal exposure, there is not enough data about the risk quantity [6, 14, 16, 24, 26].

Interpretation of results of most radioepidemiological investigations presented in the papers cited above was founded on a series of general approaches, first of all to estimate the acting factor, i.e., the exposure dose. Those approaches are the following:

- It was accepted that estimates of exposure dose contain uncertainty that is considerable as a rule;
- Even if it was possible to determine the size of dose, estimates errors, the applied analytic tool of risk analysis, ignored this circumstance;
- Practically in all radioepidemiological studies there were no instrumental measurements that were used in the process of their dosimetric support.

As a result of general properties of dosimetric support listed above, in analytic procedures of risk analysis only the stochastic nature of thyroid cancer cases was taken into account, whereas the exposure doses were supposed to be estimated without error. Moreover, one of the most popular instruments of such risk analysis, the computer package *EPICURE* (Preston et al. [20]) ignores the fact that the exposure doses contain significant uncertainty.

Investigations performed by the authors in [8] showed that dose uncertainties can be quite correctly taken into account within risk analysis. Some difficulty is that the main sources of dose uncertainties are related to errors in estimation of the mass of exposed organ, level of the measured activity, and ecological dose component. In Kukush et al. [8] it is shown that the thyroid mass measurements contain Berkson error and the instrumental measurements contain the classical error. While it is easy to estimate Berkson error using Monte Carlo methods, cf. Likhtarev et al. [15], for the classical error estimation a special analysis is needed. The present investigation is devoted exactly to this problem.

The paper is organized as follows. Section 10.1.2 contains the main results and includes the procedure of direct radioactivity measurements, the measurement error structure, the dose model, and the proposed methods of dealing with dose uncertainty in risk analysis. Section 10.1.3 concludes. A quadratic approximation to the conditional expectation of the latent variable given the observed variable is presented in Appendix 10.1.4. The approximation can be used within nonparametric calibration methods in risk analysis.

In the paper, $\mathbb{E}$ and var denote expectation and variance, respectively.

10.1.2 Materials and Methods

10.1.2.1 Direct Measurements of ^{131}I Radioactivity in the Thyroid

In May–June 1986 within the territory of Ukraine there were made more than 150 000 measurements of ^{131}I content in the thyroid for inhabitants of three northern Oblasts of Ukraine[1] who suffered from the most intensive radioactive nuclide fallouts, including 115 000 measurements among children and adolescents from the age of 0 to 18 years, as in Likhtarev et al. [10, 11].

Measurements were made by special accident brigades under general guidance of the Health Minister of Ukraine. Consulting help was made by a group of specialists from the Leningrad Institute of Sea Transport Hygiene. The group elaborated a general method to make measurements and provided the brigades with etalon ^{131}I sources for calibration of instruments.

At the beginning of mass measurements of ^{131}I content in the thyroid, significant numbers of adolescents from affected regions were taken to the places of traditional summer rest in southern, less affected Oblasts of the country, and the population of 30 km zone neighboring the Chornobyl Nuclear Power Plant was thoroughly evacuated. Therefore, 47 000 measurements were made within the territory of 10 Oblasts that were quite far from the Chornobyl Nuclear Power Plant, whereas 103 000 measurements were made within three northern Oblasts of Ukraine where there was considerable radionuclide pollution.

In the monitoring, about 100 instruments with scintillation detectors NaI(TI) were used. About one third of them (27 instruments) consisted of single-channel impulse radiometers of 5 different types that worked in the regime of impulse accumulation (typical accumulation time for different instruments varied from 15 to 200 seconds). The most popular instruments (about 65 ones) were integral (not energy-selective) radiometers SRP 68-01 that worked in the count rate regime. The electronics of those instruments made it possible to show permanently on the arrow indicator the number of impulses accumulated during nearly 5 second intervals.

As usual, the measurement scheme was the following. The measurements were made in well-ventilated rooms with damp cleaning every hour. In order to decrease background exposure, the instrumental detectors were defended with lead collimators, factory-made (for energy-selective instruments) or hand-made (for SRP 68-01 instruments). An instrumental detector was brought to the person's neck, and a single registration was made that was written in the list. Once an hour or once a day the background exposure was measured at the same point. Those measurements were written in the list as well. In order to calibrate an instrument, every hour or every day a registration was made from a bottle-phantom containing etalon ^{131}I solution. Alas, some of SRP 68-01 instruments were not calibrated at all.

[1] The territory of Ukraine is administratively divided into 26 Oblasts with an approximate area 20–30 thousand square km each.

10.1.2.2 Error Estimation of Direct Measurements

It is known (cf. Gol'danskiy et al. [3] as well as Ruark and Devol [23]), that for fixed intensity of radioactive source n, the probability to register k readings on measuring instrument, e.g., on Geiger–Mueller counter, during time period t is determined by Poisson law $\mathrm{Pois}(nt)$ with parameter nt,

$$p_n(k) = \frac{(nt)^k}{k!} e^{-nt}, \quad k = 0, 1, 2, \ldots \tag{10.1}$$

Based on (10.1) and methods of activity[2] measurements of ^{131}I in the thyroid that were described above, we get

$$Q^{\mathrm{mes}} = K^{\mathrm{mes}} \left(\frac{k_{\mathrm{th}}}{t_{\mathrm{th}}} - f_{\mathrm{sh}} \frac{k_{\mathrm{bg}}}{t_{\mathrm{bg}}} \right), \tag{10.2}$$

where Q^{mes} is a measured value of ^{131}I activity in the thyroid, K^{mes} is a measured value of device calibration coefficient, k_{th} is a number of impulses registered by the device in the process of ^{131}I activity measurement in the thyroid at the time interval t_{th}, k_{bg} is the number of impulses registered by the device in the process of background activity measurement at the time interval t_{bg}, and f_{sh} is a factor of background shielding (degree of weakening of background during measurements of a subject).

Because for large enough n, Poisson distribution (10.1) is close to normal (cf. Molina [19]), then

$$\mathrm{Pois}(nt) \approx \mathcal{N}(nt, nt), \tag{10.3}$$

and one can write

$$n_{\mathrm{th}}^{\mathrm{mes}} \sim \mathcal{N}(n_{\mathrm{th}}^{\mathrm{tr}}, \sigma_{\mathrm{th}}^2), \quad n_{\mathrm{bg}}^{\mathrm{mes}} \sim \mathcal{N}(n_{\mathrm{bg}}^{\mathrm{tr}}, \sigma_{\mathrm{bg}}^2), \tag{10.4}$$

where

$$n_{\mathrm{th}}^{\mathrm{mes}} = \frac{k_{\mathrm{th}}}{t_{\mathrm{th}}}, \quad n_{\mathrm{bg}}^{\mathrm{mes}} = \frac{k_{\mathrm{bg}}}{t_{\mathrm{bg}}}, \quad \sigma_{\mathrm{th}}^2 = \frac{n_{\mathrm{th}}^{\mathrm{tr}}}{t_{\mathrm{th}}}, \quad \sigma_{\mathrm{bg}}^2 = \frac{n_{\mathrm{bg}}^{\mathrm{tr}}}{t_{\mathrm{bg}}},$$

and index tr denotes the true value of a quantity.

Besides statistical registration error, variables $n_{\mathrm{th}}^{\mathrm{mes}}$ and $n_{\mathrm{bg}}^{\mathrm{mes}}$ contain additional instrumental error as well, and we denote its variance by σ_{dev}^2. Using (10.3) one can estimate total variances of ^{131}I intensities in the thyroid gland and background intensities,

$$\hat{\sigma}_{\mathrm{th}}^2 = \frac{n_{\mathrm{th}}^{\mathrm{mes}}}{t_{\mathrm{th}}} + \sigma_{\mathrm{dev}}^2, \quad \hat{\sigma}_{\mathrm{bg}}^2 = \frac{n_{\mathrm{bg}}^{\mathrm{mes}}}{t_{\mathrm{bg}}} + \sigma_{\mathrm{dev}}^2. \tag{10.5}$$

Due to the way the measuring instrument was calibrated, an approximate relation holds

$$K^{\mathrm{mes}} \approx K^{\mathrm{tr}}(1 + \sigma_K \gamma_1), \quad \gamma_1 \sim \mathcal{N}(0, 1), \tag{10.6}$$

[2] Hereafter for brevity we write *activity* instead of *radioactivity* and *source* instead of *radiosource*.

where σ_K is estimated based on the error of standard ^{131}I source and instrumental error.

Using (10.6) together with (10.4) and (10.5), formula (10.2) is rewritten as

$$Q^{\text{mes}} \approx K^{\text{tr}}(1 + \sigma_K \gamma_1)(n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}} + \sigma_n \gamma_2), \quad (10.7)$$

with $\sigma_n = \sqrt{\hat{\sigma}_{\text{th}}^2 + f_{\text{sh}}^2\, \hat{\sigma}_{\text{bg}}^2}$.

From (10.7) we get

$$Q^{\text{mes}} \approx K^{\text{tr}}(n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}} + (n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}})\sigma_K \gamma_1 + \sigma_n \gamma_2 + \sigma_K \sigma_n \gamma_1 \gamma_2). \quad (10.8)$$

Because

$$Q^{\text{tr}} = K^{\text{tr}}(n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}}), \quad (10.9)$$

we obtain after substitution (10.9) in (10.8) that

$$Q^{\text{mes}} \approx Q^{\text{tr}} + K^{\text{tr}}(\sigma_n \gamma_2 + (n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}})\sigma_K \gamma_1 + \sigma_K \sigma_n \gamma_1 \gamma_2) \approx Q^{\text{tr}} + \sigma_Q \gamma,$$

with $\sigma_Q = K^{\text{tr}} \sqrt{\sigma_n^2 + \sigma_n^2 \sigma_K^2 + (n_{\text{th}}^{\text{tr}} - f_{\text{sh}} n_{\text{bg}}^{\text{tr}})^2 \sigma_K^2}, \quad \gamma \sim \mathcal{N}(0, 1)$.

Because $n_{\text{th}}^{\text{tr}}$ and $n_{\text{bg}}^{\text{tr}}$ are unknown, we estimate σ_Q by

$$\sigma_Q^{\text{mes}} = K^{\text{mes}} \sqrt{\sigma_n^2 + \sigma_n^2 \sigma_K^2 + (n_{\text{th}}^{\text{mes}} - f_{\text{sh}} n_{\text{bg}}^{\text{mes}})^2 \sigma_K^2}.$$

Finally, we get the observation model of activity with additive error

$$Q^{\text{mes}} = Q^{\text{tr}} + \sigma_Q^{\text{mes}} \gamma. \quad (10.10)$$

10.1.2.3 Dose Model

Likhtarev et al. [13, 15] show that the true individual thyroid dose of i th person from a cohort consisting of N persons is equal to

$$D_i^{\text{tr}} = f_i^{\text{tr}} Q_i^{\text{tr}} / M_i^{\text{tr}}, \quad (10.11)$$

where M_i^{tr} is the true thyroid mass, f_i^{tr} is a multiplier obtained using an ecological model of radiation transition along trophic chains, and Q_i^{tr} is the true ^{131}I activity in thyroid.

Denote $f_i^{\text{tr}} / M_i^{\text{tr}} = F_i^{\text{tr}}$. Then relation (10.11) takes the form

$$D_i^{\text{tr}} = F_i^{\text{tr}} Q_i^{\text{tr}}. \quad (10.12)$$

But the true dose D_i^{tr} is unknown, because parameters F_i^{tr} and Q_i^{tr} are unknown. Only the measured dose is given,

$$D_i^{\text{mes}} = F_i^{\text{mes}} Q_i^{\text{mes}}. \quad (10.13)$$

Here, the relation between F_i^{tr} and F_i^{mes} is described by additive[3] Berkson error $F_i^{\text{tr}} = F_i^{\text{mes}} + \delta_i, \mathbb{E}\delta_i = 0$, F_i^{mes} and δ_i are stochastically independent; Q_i^{mes} is the measured thyroid activity that can be written in the form (see formula (10.10))

$$Q_i^{\text{mes}} = Q_i^{\text{tr}} + \sigma_{Q,i}^{\text{mes}} \gamma_i, \quad i=1,\ldots,N, \tag{10.14}$$

where $\gamma_1, \ldots, \gamma_n$ are standard normal variables, the value $\sigma_{Q,i}^{\text{mes}}$ is given and nonrandom, and variables γ_i, Q_i^{tr}, $i=1,\ldots,N$, are jointly independent.

The empirical distribution of a multiplier F_i^{tr} and its characteristics (expectation, variance, etc.) can be obtained by Monte Carlo procedures described in Likhtarev et al. [15].

Consider several methods of measurement error registration for radiation risk estimation.

10.1.2.4 Methods of Dose Uncertainty Registration in Risk Analysis

There exist several methods for how to include errors in independent variables in the procedure of regression parameters estimation, see [2, 8, 17, 18, 21]. But the simplest and the most popular method is regression calibration. Its popularity is related to the fact that after calibration (that is, substitution of an independent variable with its conditional expectation) it is possible to utilize standard procedures of regression parameters estimation and to use for that purpose the computer package *EPICURE*. There exist two kinds of calibration, parametric and nonparametric. Both are considered in the present paper.

The main idea of regression calibration (see Carroll et al. [2] and Kukush et al. [8]) is as follows. In the radiation risk model (see *Health Risks from Exposure to Low Levels of Ionizing Radiation* [4]) instead of true doses D_i^{tr}, their conditional expectations are used,

$$\mathbb{E}(D_i^{\text{tr}} \mid D_i^{\text{mes}}). \tag{10.15}$$

Substitute (10.12) and (10.13) in (10.15) and obtain

$$\mathbb{E}(D_i^{\text{tr}} \mid D_i^{\text{mes}}) = \mathbb{E}(F_i^{\text{tr}} Q_i^{\text{tr}} \mid D_i^{\text{mes}}) = \mathbb{E}(F_i^{\text{mes}} Q_i^{\text{tr}} \mid D_i^{\text{mes}}).$$

Denote $\overline{D}_i^{\text{tr}} = F_i^{\text{mes}} Q_i^{\text{tr}}$ and get

$$\mathbb{E}(D_i^{\text{tr}} \mid D_i^{\text{mes}}) = \mathbb{E}(\overline{D}_i^{\text{tr}} \mid D_i^{\text{mes}}).$$

Using (10.14) we obtain

$$D_i^{\text{mes}} = F_i^{\text{mes}} Q_i^{\text{mes}} = F_i^{\text{mes}} (Q_i^{\text{tr}} + \sigma_{Q,i}^{\text{mes}} \gamma_i) = \overline{D}_i^{\text{tr}} + F_i^{\text{mes}} \sigma_{Q,i}^{\text{mes}} \gamma_i. \tag{10.16}$$

[3] The error can be either additive or multiplicative. In this case it is not so important, and the main requirement is that the equality holds $\mathbb{E}(F_i^{\text{mes}}) = \mathbb{E}(F_i^{\text{tr}})$.

Random variables $\{\delta_i,\ i\geq 1\}$, $\{\gamma_i,\ i\geq 1\}$ and vectors $\{(F_i^{\text{mes}}, Q_i^{\text{tr}}),\ i\geq 1\}$ are jointly independent, at that F_i^{mes} and Q_i^{tr} could be correlated. Denote $\sigma_i = F_i^{\text{mes}}\sigma_{Q,i}^{\text{mes}}$, then (10.16) takes the form

$$D_i^{\text{mes}} = \overline{D}_i^{\text{tr}} + \sigma_i\gamma_i. \tag{10.17}$$

In fact (10.17) is a dose observation model with classical error.

Parametric calibration method assumes that the sample $\overline{D}_i^{\text{tr}}$, $i=1,\dots,N$, comes from a known distribution (or the latter can be estimated reliably). Because the radiation doses of thyroids are essentially positive, and their distribution is left-asymmetric, see Likhtarev et al. [15], a log-normal distribution provides a good approximation. Therefore, we suppose that

$$\log \overline{D}^{\text{tr}} \sim \mathcal{N}(\mu_{\overline{D}^{\text{tr}}}, \sigma^2_{\overline{D}^{\text{tr}}}). \tag{10.18}$$

If the measured doses D_i^{mes}, $i=1,\dots,N$, are given, then the parameters of the distribution (10.18) can be easily estimated. First, we estimate $m_{\overline{D}^{\text{tr}}} = \mathbb{E}\overline{D}^{\text{tr}}$ and $v_{\overline{D}^{\text{tr}}} = \text{var}(\overline{D}^{\text{tr}})$,

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

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

Using relations for the moments of a log-normal distribution, see Korolyuk et al. [7], it is easy to construct the estimators of $\mu_{\overline{D}^{\text{tr}}}$ and $\sigma^2_{\overline{D}^{\text{tr}}}$:

$$\hat{\mu}_{\overline{D}^{\text{tr}}} = \log\left(\frac{(\widehat{m}_{\overline{D}^{\text{tr}}})^2}{\sqrt{\hat{v}_{\overline{D}^{\text{tr}}} + (\widehat{m}_{\overline{D}^{\text{tr}}})^2}}\right), \quad \hat{\sigma}^2_{\overline{D}^{\text{tr}}} = \log\left(\frac{\hat{v}_{\overline{D}^{\text{tr}}}}{(\widehat{m}_{\overline{D}^{\text{tr}}})^2} + 1\right).$$

If the sequence $\{\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2,\ i\geq 1\}$ is bounded, then those estimators are strongly consistent, that is, converge to $\mu_{\overline{D}^{\text{tr}}}$ and $\sigma^2_{\overline{D}^{\text{tr}}}$ a.s., as $n\to\infty$. In what follows the parameters $\mu_{\overline{D}^{\text{tr}}}$ and $\sigma^2_{\overline{D}^{\text{tr}}}$ will be assumed known.

Let ρ_{tr} be a pdf of $\overline{D}_i^{\text{tr}}$, ρ_{mes} be a pdf of D_i^{mes}, $\rho_{\text{tr,mes}}$ be a joint pdf of $\overline{D}_i^{\text{tr}}$ and D_i^{mes}, and ρ_γ be a pdf of $\sigma_i\gamma_i$. Then

$$\rho_{\text{tr,mes}}(\overline{D}_i^{\text{tr}}, D_i^{\text{mes}}) = \rho_{\text{tr}}(\overline{D}_i^{\text{tr}})\cdot\rho_\gamma(D_i^{\text{mes}} - \overline{D}_i^{\text{tr}}),$$

and the conditional pdf is

$$\rho_{\text{tr}|\text{mes}}(\overline{D}_i^{\text{tr}}) = \frac{\rho_{\text{tr,mes}}(\overline{D}_i^{\text{tr}}, D_i^{\text{mes}})}{\int_0^\infty \rho_{\text{tr,mes}}(t, D_i^{\text{mes}})\,dt} = \frac{\rho_{\text{tr}}(\overline{D}_i^{\text{tr}})\cdot\rho_\gamma(D_i^{\text{mes}} - \overline{D}_i^{\text{tr}})}{\rho_{\text{mes}}(D_i^{\text{mes}})}.$$

This implies that

$$\mathbb{E}(\overline{D}_i^{\mathrm{tr}} \mid D_i^{\mathrm{mes}}) = \int_0^\infty t\rho_{\mathrm{tr|mes}}(t)\,dt = \frac{1}{\rho_{\mathrm{mes}}(D_i^{\mathrm{mes}})}\int_0^\infty t\rho_{\mathrm{tr}}(t)\rho_\gamma(D_i^{\mathrm{mes}}-t)\,dt,$$

with

$$\rho_{\mathrm{mes}}(D_i^{\mathrm{mes}}) = \int_0^\infty \rho_{\mathrm{tr}}(t)\cdot\rho_\gamma(D_i^{\mathrm{mes}}-t)\,dt.$$

Because of (10.18),

$$\rho_{\mathrm{tr}}(t) = \frac{1}{t\sqrt{2\pi}\sigma_{\overline{D}^{\mathrm{tr}}}}\exp\left(-\frac{(\log t-\mu_{\overline{D}^{\mathrm{tr}}})^2}{2\sigma^2_{\overline{D}^{\mathrm{tr}}}}\right).$$

Then

$$\rho_{\mathrm{mes}}(D_i^{\mathrm{mes}}) = \\ = \int_0^\infty \frac{1}{t\sqrt{2\pi}\sigma_{\overline{D}^{\mathrm{tr}}}}\exp\left(-\frac{(\log t-\mu_{\overline{D}^{\mathrm{tr}}})^2}{2\sigma^2_{\overline{D}^{\mathrm{tr}}}}\right)\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(D_i^{\mathrm{mes}}-t)^2}{2\sigma_i^2}\right)dt. \tag{10.19}$$

Similarly,

$$\int_0^\infty t\cdot\rho_{\mathrm{tr}}(t)\cdot\rho_\gamma(D_i^{\mathrm{mes}}-t)\,dt = \\ = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_{\overline{D}^{\mathrm{tr}}}}\exp\left(-\frac{(\log t-\mu_{\overline{D}^{\mathrm{tr}}})^2}{2\sigma^2_{\overline{D}^{\mathrm{tr}}}}\right)\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{(D_i^{\mathrm{mes}}-t)^2}{2\sigma_i^2}\right)dt. \tag{10.20}$$

We change variables in integrals (10.19) and (10.20):

$$\begin{aligned} dz &= \frac{1}{t\sqrt{2\pi}\sigma_{\overline{D}^{\mathrm{tr}}}}\exp\left(-\frac{(\log t-\mu_{\overline{D}^{\mathrm{tr}}})^2}{2\sigma^2_{\overline{D}^{\mathrm{tr}}}}\right)dt,\\ z &= \int_0^t \frac{1}{t\sqrt{2\pi}\sigma_{\overline{D}^{\mathrm{tr}}}}\exp\left(-\frac{(\log t-\mu_{\overline{D}^{\mathrm{tr}}})^2}{2\sigma^2_{\overline{D}^{\mathrm{tr}}}}\right)dt = G(t),\\ t &= G^{-1}(z). \end{aligned} \tag{10.21}$$

Because $z = G(t)$ is a cdf of a log-normal law, then

$$z(0) = 0, \quad z(\infty) = 1. \tag{10.22}$$

We plug-in (10.21) and (10.22) into (10.19) and (10.20) and get

$$\rho_{\text{mes}}(D_i^{\text{mes}}) = \int_0^1 \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(D_i^{\text{mes}} - G^{-1}(z))^2}{2\sigma_i^2} \right) dz, \tag{10.23}$$

$$\int_0^\infty t\rho_{\text{tr}}(t) \cdot \rho_\gamma(D_i^{\text{mes}} - t)\, dt = \int_0^1 \frac{G^{-1}(z)}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(D_i^{\text{mes}} - G^{-1}(z))^2}{2\sigma_i^2} \right) dz. \tag{10.24}$$

In spite of $G^{-1}(z) \to +\infty$ as $z \to 1$, the integrals (10.23) and (10.24) are proper, because the integrand can be defined 0 by continuity at point $z = 1$. To compute the integrals we use an identity

$$G^{-1}(z) = \exp(\mu_{\overline{D}^{\text{tr}}} + \sigma_{\overline{D}^{\text{tr}}} \Phi^{-1}(z)),$$

where $\Phi(z)$ stands for the cdf of standard normal law. The function $\Phi^{-1}(z)$ is tabulated, and therefore one can use the trapezoid formula to compute the integrals (10.23) and (10.24).

Within *nonparametric* calibration we do not parametrize the distribution of doses $\overline{D}_i^{\text{tr}}$, $i=1,\dots,N$. Therefore, in order to compute the conditional expectation $\mathbb{E}(\overline{D}_i^{\text{tr}} \mid D_i^{\text{mes}})$, one can use a polynomial approximation in powers of D_i^{mes} as described in Appendix, or estimation methods of the distribution of $\overline{D}^{\text{tr}}$ based on the maximum likelihood function presented in Kukush et al. [8].

10.1.3 Conclusion and Discussion

In present paper we constructed a statistical model of radioactivity measurements. It was shown that the measurements always contain random additive error of the classical type. It was shown as well that the measurement error can be regarded as normally distributed with sufficient accuracy. Estimation methods of the measurement error variance were elaborated.

In the paper, a model of thyroid dose was constructed that takes into account the presence of both classical error, that is present in thyroid radioactivity measurements, and Berkson error, that is inevitably present in the ecological dosimetry model. Summarizing, we get the observation model of the dose with classical additive normal error. This model differs from the one proposed by the authors in [8]. In that model methods were elaborated to estimate radiation risks under uncertainty in doses. In this relation, the error was assumed multiplicative without particular investigation. In the present paper, the dose model is more realistic.

Also the authors elaborated two methods to deal with dose uncertainty in risk analysis. One of them, parametric calibration of dose, is based on the log-normality assumption for the sample of true doses. This assumption is based on the paper by Likhtarev et al. [15]. The second method, nonparametric calibration of dose, uses polynomial approximation for conditional expectation of the true dose in powers of the measured dose.

A simulation study is not presented here and will be reported in a forthcoming paper. Information about the efficiency of parametric and nonparametric calibration and other methods dealing with uncertainty in risk analysis can be found in [8, 17, 18, 22].

10.1.4 Appendix. Approximation of Conditional Expectations

Let $w_n = x_n + u_n,\ n{=}1,2,\ldots,N$ be realizations of a random variable w, $u_n \sim \mathcal{N}(0,\sigma_n^2),\ \sigma_n > 0$, and variables $\{x_n,\ u_n,\ n{\ge}1\}$ are jointly independent. Here, x_n is the unknown true value of the observable variable (dose), and w_n is a measurement of x_n. Denote $\mathbb{E}x = \mu_x,\ \ \mathrm{var}x = \sigma_x^2$. Suppose also that $\mathbb{E}x^2 < \infty$ and the sequence $\{\frac{1}{N}\sum_{n=1}^N \sigma_n^2,\ N{\ge}1\}$ is bounded.

10.1.4.1 Linear Approximation of $\mathbb{E}(x_n \mid w_n)$

We search for the conditional expectation in a form

$$\mathbb{E}(x_n \mid w_n) = a + bw_n. \tag{10.25}$$

From (10.25) we have

$$\begin{aligned}
\mathbb{E}x &= \mu_x = \mathbb{E}(a + bw_n),\\
\mathbb{E}x_n w_n &= \mathbb{E}x^2 = \sigma^2 + \mu_x^2 = \mathbb{E}(w_n \mathbb{E}(x_n \mid w_n))\\
&= \mathbb{E}(aw_n + bw_n^2) = a\mu_x + b(\sigma_{w_n}^2 + \mu_x^2).
\end{aligned}$$

Solve the system of two algebraic equations in a and b and obtain

$$A = (1-K)\mu_x, \quad b = K,$$

where $K = \frac{\sigma_x^2}{\sigma_{w_n}^2} = \frac{\sigma_x^2}{\sigma_x^2+\sigma_n^2}$ is the reliability ratio.

Parameters μ_x and σ_x^2 can be easily estimated from observations,

$$\begin{aligned}
\hat{\mu}_x &= \overline{w} = \frac{1}{N}\sum_{n=1}^N w_n,\\
(\hat{\sigma}_x)^2 &= \frac{1}{N-1}\sum_{n=1}^N (w_n - \overline{w})^2 - \overline{\sigma^2},
\end{aligned}$$

with $\overline{\sigma^2} = \frac{1}{N}\sum_{n=1}^N \sigma_n^2$.

Thus,

$$\mathbb{E}(x_n \mid w_n) \approx (1-K)\mu_x + K w_n. \tag{10.26}$$

In case $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$, then in (10.26) the equality is attained, that is,

$$\mathbb{E}(x_n \mid w_n) = (1-K)\mu_x + K w_n.$$

10.1.4.2 Quadratic Approximation of $\mathbb{E}(x_n \mid w_n)$

Now, suppose that $\mathbb{E}x^4 < \infty$ and the sequence $\{\frac{1}{N}\sum_{n=1}^{N}\sigma_n^4,\ N \geq 1\}$ is bounded. We search for the conditional expectation in a form

$$\mathbb{E}(x_n \mid w_n) = \sum_{i=0}^{2} a_i w_n^i. \tag{10.27}$$

From (10.27) we have for $j = 0, 1, 2$:

$$\mathbb{E}(w_n^j\, \mathbb{E}(x_n \mid w_n)) = \mathbb{E}\,\mathbb{E}(w_n^j x_n \mid w_n) = \mathbb{E}(w_n^j x_n) = \mathbb{E}((x_n + u_n)^j x_n) =: b_j,$$

$$b_j = \mathbb{E}\left(w_n^j \sum_{i=0}^{2} a_i w_n^i\right) = \sum_{i=0}^{2} a_i \mathbb{E} w_n^{i+j}. \tag{10.28}$$

Equalities (10.28) form a system of three linear equations with three unknowns a_i, $i=0,1,2$. The matrix of the system

$$\mathbf{G} = (G_{ij}) = (\mathbb{E} w_n^{i+j}), \quad i, j = 0, 1, 2,$$

is a Gram matrix. It is positive definite, because random variables 1, w_n, w_n^2 are linearly independent in the space $L_2(\Omega, \mathcal{F}, \mathbb{P})$ of random variables with finite second moment (the linear independence follows from the fact that w_n has continuous cdf and, therefore, is not concentrated at two points).

Let $a = (a_0, a_1, a_2)^T \in \mathbb{R}^{3\times 1}$, $b = (b_0, b_1, b_2)^T \in \mathbb{R}^{3\times 1}$. Then (10.28) can be written in a vector form $\mathbf{G}a = b$. Hence,

$$a = \mathbf{G}^{-1} b.$$

To estimate entries G_{ij} of the matrix $\mathbf{G}$ we construct estimators of $\mathbb{E} w_n^k$, $0 \leq n \leq 4$. We have

$$\begin{aligned}
\mathbb{E} w_n^0 &= 1, \\
\mathbb{E} w_n^1 &= \mathbb{E}x, \\
\mathbb{E} w_n^2 &= \mathbb{E}x^2 + \sigma_n^2, \\
\mathbb{E} w_n^3 &= \mathbb{E}x^3 + 3\sigma_x^2 \mathbb{E}x, \\
\mathbb{E} w_n^4 &= \mathbb{E}x^4 + 6\sigma_n^2 \mathbb{E}x^2 + 3\sigma_n^4.
\end{aligned}$$

Next,

$$b_0 = \mathbb{E}x,$$
$$b_1 = \mathbb{E}(x_n + u_n)x_n = \mathbb{E}x^2,$$
$$b_2 = \mathbb{E}(x_n + u_n)^2 x_n = \mathbb{E}x^3 + \sigma_n^2 \mathbb{E}x.$$

Estimators of $\mathbb{E}x$, $\mathbb{E}x^2$, $\mathbb{E}x^3$, $\mathbb{E}x^4$ can be constructed based on observations w_n:

$$\widehat{\mathbb{E}}x = \overline{w},$$
$$\widehat{\mathbb{E}}x^2 = \overline{w^2} - \overline{\sigma^2},$$
$$\widehat{\mathbb{E}}x^3 = \overline{w^3} - 3\overline{w\sigma^2},$$
$$\widehat{\mathbb{E}}x^4 = \overline{w^4} - 6\overline{\sigma^2}\,\widehat{\mathbb{E}}x^2 - 3\overline{\sigma^4} = \overline{w^4} - 6\,\overline{\sigma^2}\ \overline{w^2} + 6\left(\overline{\sigma^2}\right)^2 - 3\overline{\sigma^4}.$$

Here hat denotes the estimator, and bar stands for the average in $n = 1, 2, \dots, N$.

Thus, $\hat{a} = (\widehat{\mathbf{G}})^{-1}\hat{b}$, and the final approximation is

$$\mathbb{E}(x_n \mid w_n) \approx \sum_{i=1}^{2} \hat{a}_i w_n^i.$$

Bibliography

[1] E. E. Buglova, J. E. Kenigsberg, N. V. Sergeeva, Cancer risk estimation in Belarussian children due to thyroid irradiation as a consequence of the Chernobyl nuclear accident, *Health Phys* **71** (1996), 45–49.

[2] R. J. Carroll, D. Ruppert, L. A. Stefanski, C. A. Crainiceanu, *Measurement Error in Nonlinear Models. A Modern Perspective*, Chapman and Hall/CRC, Boca Raton, 2006.

[3] V. I. Gol'danskiy, A. V. Kutsenko, M. I. Podgoretskiy, Statistics of Counting-out During Registration of Nuclear Particles. Moscow, 1959 (in Russian).

Note 53: AU: [3] Publisher is missing.

[4] *Health Risks from Exposure to Low Levels of Ionizing Radiation*, BEIR VII Phase 2, National Academy Press, Washington D.C., 2006.

[5] P. Jacob, T. I. Bogdanova, E. Buglova et al., Thyroid cancer risk in areas of Ukraine and Belarus affected by the Chernobyl accident, *Radiation Research* **165** (2006), 1–8.

[6] K. J. Kopecky, V. Stepanenko, N. Rivkind et al., Childhood thyroid cancer, radiation dose from Chernobyl and dose uncertainties in Bryansk Oblast, Russia: A population-based case-control study, *Radiation Research* **166** (2006), 367–374.

[7] V. S. Korolyuk, N. I. Portenko, A. V. Skorokhod, A. F. Turbin, *Reference Book on Probability Theory and Mathematical Statistics*, Moscow, 1985 (In Russian).

Note 54: AU: [6] Publisher is missing.

[8] A. Kukush, S. Shklyar, S. Masiuk, I. Likhtarov, L. Kovgan, R. Carroll, A. Bouville, 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, (2011), DOI: 10.2202/1557-4679.1281.

[9] I. A. Likhtarev, N. K. Shandala, G. M. Goulko, I. A. Kairo, Exposure doses to thyroid of the Ukrainian population after the Chernobyl accident *Health Phys.* **64** (1993), 594–599.

[10] I. A. Likhtarev, G. Prohl, K. Henrichs, Reliability and accuracy of the ^{131}I thyroid activity measurements performed in the Ukraine after the Chernobyl accident in 1986, Munich, GSF-Bericht 19/93, Institut für Strahlenschutz (1993).

[11] I. A. Likhtarev, G. M. Goulko, B. G. Sobolev, I. A. Kairo, G. Prohl, P. Rath, K. Henrichs, Evaluation of the ^{131}I thyroid-monitoring measurements performed in Ukraine during May and June of 1986. *Health Phys.* **69** (1995), 6–15.

[12] I. A. Likhtarev, B. G. Sobolev, I. A. Kairo et al., Thyroid cancer in Ukraine. *Nature* **375** (1995), 365–378.

[13] I. Likhtarev, L. Kovgan, S. Vavilov, M. Chepurny, A. Bouville, N. Luckyanov, P. Jacob, P. Voilleque, G. Voigt, Post-Chornobyl thyroid cancers in Ukraine. Report 1: Estimation of thyroid doses. *Radiation Research* **163** (2005), 125–136.

[14] I. Likhtarev, L. Kovgan, S. Vavilov, M. Chepurny, A. Bouville, N. Luckyanov, P. Jacob, P. Voilleque, G. Voigt, Post-Chornobyl thyroid cancers in Ukraine. Report 2: Risk analysis. *Radiation Research* **166** (2006), 375–386.

[15] I. Likhtarev, A. Bouville, L. Kovgan, N. Luckyanov, P. Voilleque, M. Chepurny, Questionnaire- and measurement-based individual thyroid doses in Ukraine resulting from the Chornobyl Nuclear Reactor accident. *Radiation Research* **166** (2006), 271–286.

[16] J. L. Lyon, S. C. Alder, M. B. Stone, A. Scholl, J. C. Reading, R. Holubkov, X. Sheng, G. L. White, K. T. Hegmann, L. Anspaugh, F. O. Hoffman, S. L. Simon, B. Thomas, R. J. Carroll, A. W. Meikle, Thyroid disease associated with exposure to the Nevada Test Site radiation: a reevaluation based on corrected dosimetry and examination data, *Epidemiology* **17** (2006), 604–614.

[17] B. Mallick, F. O. Hoffman, R. J. Carroll, Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada Test Site. *Biometrics* **58** (2002), 13–20.

[18] S. V. Masiuk, S. V. Shklyar, A. G. Kukush, S. E. Vavilov, Impact of doses uncertainties on the radiation risks estimation. *Radiation and Risk – Bulletin of the National Radiation and Epidemiological Registry* **17(3)** (2008), 64–75 (in Russian). http://mechmat.univ.kiev.ua/eng/ppages/kukush/Public/files/Unsert_risk_influence.doc.

[19] E. C. Molina, *Poisson's Exponential Binomial Limit*, D. Van Nostrand Co., New York, 1945.

[20] D. L. Preston, J. H. Lubin, D. A. Pierce, M. E. McConney, *EPICURE User's Guide*, Hirosoft Corporation, Seattle (WA), 1993.

[21] E. Ron, F. O. Hoffman, Uncertainties in radiation dosimetry and their impact on dose-response analysis. Proceedings of a workshop held September 3–5, 1997 in Bethesda, Maryland, NIH Publication No. 99-4541, 1999.

[22] E. Ron, J. Lubin, R. Shore et al., Thyroid cancer after exposure to external radiation: a pooled analysis of seven studies, *Radiation Research* **141** (1995), 259–277.

[23] A. Ruark, L. Devol, The General Theory of Fluctuations in Radioactive Disintegration, *Phys. Rev.* **49** (1936), 355–367.

[24] M. D. Tronko, G. R. Howe, T. I. Bogdanova et al., A cohort study of thyroid cancer and other thyroid diseases after the Chornobyl accident: thyroid cancer in Ukraine detected during first screening. *Journal of the National Cancer Institute* **98** (2006), 897–903.

[25] A. F. Tsyb, E. M. Parshkov, V. V. Shakhtarin et al., Thyroid cancer in children and adolescents of Bryansk and Kaluga regions. Proceedings of the first international conference, in: *The Radiological Consequences of the Chernobyl Accident*, 691–698, Minsk, Belarus, 1996.

Note 55: AU: [25] Publisher is missing.

[26] L. Zablotska, E. Ron, A. Rozhko, Thyroid cancer risk in Belarus among children and adolescents exposed to radioiodine after the Chornobyl accident, *British Journal of Cancer* **104** (2011), 181–187.

Author Information

Illya Likhtarov, Scientific Center for Radiation Medicine of the Academy of Medical Sciences of Ukraine, Ukraine
E-mail: Likh@rpi.kiev.ua

Sergii Masiuk, Scientific Center for Radiation Medicine of the Academy of Medical Sciences of Ukraine, Ukraine
E-mail: masja1979@gmail.com

Mykola Chepurny, Scientific Center for Radiation Medicine of the Academy of Medical Sciences of Ukraine, Ukraine
E-mail: ch@rpi.kiev.ua

Alexander Kukush, National Taras Shevchenko University of Kyiv, Ukraine
E-mail: alexander_kukush@mail.univ.kiev.ua

Sergiy Shklyar, National Taras Shevchenko University of Kyiv, Ukraine
E-mail: shklyar@mail.univ.kiev.ua

Andre Bouville, National Cancer Institute, NIH, DHHS, Bethesda, Maryland, USA
E-mail: bouvilla@exchange.nih.gov

Lina Kovgan, Scientific Center for Radiation Medicine of the Academy of Medical Sciences of Ukraine, Ukraine
E-mail: lina@rpi.kiev.ua