

УДК: 621.039.573:616-036.8

С. В. Масюк<sup>1</sup>✉, С. В. Шкляр<sup>2</sup>, О. Г. Кукуш<sup>2</sup>

<sup>1</sup>Державна установа “Національний науковий центр радіаційної медицини Національної академії медичних наук України”, вул. Мельникова, 53, м. Київ, 04050, Україна

<sup>2</sup>Київський національний університет імені Тараса Шевченка, м. Київ, пр-т акад. Глушкова, 4-е

## ПОХИБКИ БЕРКСОНІВСЬКОГО ТИПУ В ДОЗАХ ОПРОМІНЕННЯ ТА ЇХ ВПЛИВ НА ОЦІНКУ РАДІАЦІЙНИХ РИЗИКІВ

**Мета.** Оцінити вплив похибок берксонівського типу в дозах опромінення на результати ризик-аналізу на прикладі радіоепідеміологічних досліджень онкозахворюваності щитоподібної залози.

**Матеріали та методи.** В роботі за допомогою імітаційного стохастичного моделювання досліджено вплив похибок Берксона в дозах опромінення щитоподібної залози на оцінку ризиків радіоіндукованих стохастичних ефектів.

**Результати.** Показано, що наявність похибок у дозах призводить до зсунення так званих наївних оцінок фонові захворюваності і надлишкового абсолютного ризику в лінійній логістичній моделі регресії. Зі збільшенням дозових похибок зсунення наївних оцінок практично лінійно зростає. Використання розробленого авторами методу повної максимальної вірогідності суттєво поліпшує оцінки параметрів моделі абсолютного ризику.

**Висновки.** Ігнорування значних берксонівських похибок призводить до переоцінки фонові захворюваності і недооцінки експесу абсолютного ризику. У той же час зсув значно нижчий, ніж у разі класичної похибки.

**Ключові слова:** дози опромінення, радіаційний ризик, берксонівська похибка, метод повної максимальної вірогідності, наївна оцінка.

*Проблеми радіаційної медицини та радіобіології. 2013. Вип. 18. С. 119–126.*

S. V. Masiuk<sup>1</sup>✉, S. V. Shklyar<sup>2</sup>, A. G. Kukush<sup>2</sup>

<sup>1</sup>State Institution “National Research Center for Radiation Medicine of the National Academy of Medical Sciences of Ukraine”, Melnykov str., 53, Kyiv, 04050, Ukraine

<sup>2</sup>Taras Shevchenko National University of Kyiv, Pr. Glushkov, 4-e

## Berkson errors in radiation dose assessments and their impact on radiation risk estimates

**Objective.** To estimate the influence of Berkson errors in exposure doses on the results of risk analysis within example of radiation epidemiological studies of the thyroid cancer prevalence.

**Materials and methods.** The impact of Berkson errors of the thyroid doses in a dose-response analysis is studied by the method of stochastic simulation.

**Results.** Presence of errors in doses results in bias of the naive estimations of baseline morbidity and absolute risk excess in the linear logistic regression model. With the increase of dose errors the bias of the naive estimate increases almost linearly. The use of the full maximum likelihood method developed by authors improves essentially the estimates of parameters in the excess absolute risk model.

**Conclusion.** Ignoring of the significant Berkson errors in dose assessment leads to the bias in estimates of background morbidity (overestimated) and of the excess absolute risk (underestimated). At that the bias is essentially less than the one for the case of classical error.

**Key words:** doses of exposure, radiation risk, Berkson error, full maximum likelihood method, naive estimator.

*Problems of radiation medicine and radiobiology. 2013;18:119–126.*

✉ Масюк Сергій Володимирович, e-mail: masja1979@gmail.com

© Масюк С. В., Шкляр С. В., Кукуш О. Г., 2013

Внаслідок Чорнобильської катастрофи у 1986 році значна частина територій України, Білорусії і Росії зазнала радіоактивного забруднення, а жителі цих територій зазнали радіоактивного опромінення. Найбільш істотним було опромінення щитоподібної залози (ЩЗ) внаслідок надходження в організм радіоізоотопів йоду, перш за все  $^{131}\text{I}$ .

Фактично зростання онкозахворюваності ЩЗ у дітей та підлітків, зумовлене її внутрішнім опроміненням, обумовленим чорнобильськими радіоактивними викидами, стало головним статистично достовірним віддаленим ефектом цієї аварії. Не дивно, що цей феномен викликав величезний інтерес у радіоепідеміологів усього світу і спричинив низку епідеміологічних досліджень в Україні, Білорусії й Росії.

Відомо, що оцінка дози неминує супроводжується похибками класичного та/або берксонівського типу [1–3]. Однак інтерпретація результатів практично всіх радіоепідеміологічних досліджень ґрунтувалася на методах оцінки ризиків, які не враховують наявності похибок у дозах опромінення. Одним з наслідків припущення про відсутність похибок у дозах є зсунення оцінок коефіцієнтів ризику і спотворення форми кривої “доза-ефект”. Зауважимо, що до подібних спотворень оцінок ризиків призводять не лише систематичні похибки в оцінках доз, що в принципі очевидно, але й випадкові похибки. І хоча останнім часом робилися неодноразові спроби побудови математичного апарату, що враховує статистичний розподіл не лише ефекту, але й дози [4–6], нині немає остаточного висновку щодо впливу класичної або берксонівської похибки в дозових оцінках на кінцевий результат ризик-аналізу, що виражається зазвичай у величинах відносного або абсолютного ризику.

Слід зазначити, що внесок берксонівської похибки у загальну похибку дози опромінення, як правило, більший за внесок класичної похибки. Це пояснюється тим, що берксонівська похибка пов'язана з використанням методу Монте-Карло для оцінки статистичних розподілів доз опромінення [1–3]. В той же час, порівняно з класичною похибкою, вплив берксонівської похибки на оцінку радіаційного ризику є практично недослідженим. Тому дана робота присвячена саме цій темі. Дослідження проводились на імітованих даних.

## МЕТА ДОСЛІДЖЕННЯ

Оцінити вплив похибок берксонівського типу в дозах опромінення на результати ризик-аналізу на прикладі радіоепідеміологічних досліджень онкозахворюваності ЩЗ.

As a result of Chornobyl accident in 1986 a considerable part of the territory of Ukraine, Belarus and Russia appeared under contamination and the inhabitants of territory since are under radioactive exposure. The thyroid exposure was most significant as a result of intake of iodine radioisotopes and  $^{131}\text{I}$  first of all.

In fact the increasing thyroid cancer incidence of children and adolescents caused by the internal thyroid exposure due to Chornobyl radioactive fallout became the main statistically significant remote effect of the accident. No wonder this phenomenon has been of great interest by the radiation epidemiologists worldwide and several epidemiological studies were conducted in Ukraine, Belarus and Russia.

It is known that the dose estimation is inevitably followed by errors of classical and/or Berkson type [1–3]. But interpretation of results for practically all radioepidemiological studies was based on risk estimation methods not taking into account the presence of errors in the exposure dose values. The bias in risk coefficient estimates and distortion of the “dose-effect” curve are the consequences of assumption of no errors in doses. Noteworthy that not only systematic errors in dose estimates (this is in principle evident) but random errors as well yield such risk estimates distortion. And despite non-occasional attempts made last time to elaborate the mathematical tool which takes into account a statistical distribution of not only effect but dose as well [4–6] at the moment there is no ultimate conclusion about the influence of classical or Berkson error in dose estimates on the final result of risk analysis. The latter is usually expressed in the value of relative or absolute risk.

It should be noted that the contribution of Berkson error to a total exposure dose error is usually larger than the contribution of classical error. This is explained by the fact that Berkson error is related to the use of Monte Carlo method for estimation of statistical distributions of exposure doses [1–3]. Meanwhile the influence of Berkson error on the radiation risk estimate is practically unexplored in comparison with classical error. That is why the present paper deals with this topic. Research was based on simulated data.

## RESEARCH OBJECTIVE

Estimation of the influence of Berkson errors in exposure doses on the results of risk analyses on an example of radioepidemiological studies of thyroid cancer.

## МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Для моделювання виникнення випадків захворювання на рак ЩЗ на фіксованому часовому інтервалі використовувалась логістична модель рідкісних подій з бінарною залежною змінною  $Y_i$ , яка може набувати двох значень: 0 або 1 [7]. Модель логістичної регресії застосовується, коли у розпорядженні епідеміолога є індивідуальні дані спостережень. Це зазвичай типowo для когортних досліджень, суть яких полягає в тому, що спостерігається деяка група (когорти) індивідуумів, які свого часу зазнали впливу радіаційного чинника. Згодом у деяких членів цієї когорти може виявитися захворювання на рак, тобто кожному з членів когорти можна співставити бінарну змінну, що набуває значення 0 (“ $i$ -та людина *не захворіла*”) або 1 (“ $i$ -та людина *захворіла*”). Вважається, що ми володіємо повною інформацією про кожну особу, яка знаходиться в когорті (тобто знаємо її вік, стать, індивідуальну дозу, час реалізації захворювання, час що, пройшов з моменту опромінення, вік на момент опромінення тощо). Тоді умовна ймовірність  $Y_i$  при фіксованій дозі  $D_i$  опромінення визначається співвідношеннями:

$$P(Y_i = 1 | D_i) = \frac{\lambda_i(D_i, \mathbf{b})}{1 + \lambda_i(D_i, \mathbf{b})}, \quad P(Y_i = 0 | D_i) = \frac{1}{1 + \lambda_i(D_i, \mathbf{b})}, \quad (1)$$

де  $\lambda_i(D_i, \mathbf{b})$  – захворюваність на рак ЩЗ, а  $\mathbf{b}$  – вектор параметрів моделі.

Далі, знаючи ймовірності індивідуумів захворіти, можна скласти функцію вірогідності, максимум якої оптимізує вектор невідомих коефіцієнтів  $\mathbf{b}$ . У випадку відсутності похибок у дозах опромінення або у випадку коли наявність похибок ігнорується (так зване наївне оцінювання), функція вірогідності для логістичної моделі описується співвідношенням

$$L(\mathbf{b}) = \prod_i P_i(D_i, \mathbf{b})^{Y_i} (1 - P_i(D_i, \mathbf{b}))^{1-Y_i} \quad (2)$$

**Застосування методу повної максимальної вірогідності (МПМВ) для оцінки ризиків за наявності мультиплікативної похибки берксонавського типу в дозах опромінення**

Нехай у дозах присутня берксонавська мультиплікативна похибка, тоді:

$$\ln D_i^{tr} = \ln D_i^{mes} + u_i, \quad i = 1, \dots, N, \quad (3)$$

$$u_i \sim N(0, \sigma_i^2),$$

де  $D_i^{tr}$  – істинна доза (невідомо),  $D_i^{mes}$  – оцінка дози (відома),  $u_i$  – гауссівська похибка.

## RESEARCH MATERIALS AND METHODS

To simulate the cases of thyroid cancer on fixed time interval we have used a logistic rare-events-model with binary response  $Y_i$  that takes values 0 and 1 [7]. The logistic regression model is used when an epidemiologist has individual observation data. Usually it is typical for the cohort studies. The main issue of them is that certain group (cohort) of persons is observed and the persons some time ago were under the influence of radiation. After a while it can be detected that some cohort members got thyroid cancer, that is each cohort member corresponds to a binary variable which takes value 0 ( $i^{\text{th}}$  person *did not* fall sick) or 1 ( $i^{\text{th}}$  person *did* fall sick). It is assumed that we possess complete information about each person from a cohort (that is we know his/her age, sex, individual radiation dose, time of the disease onset, time from the exposure moment, age at the exposure moment etc.). Then a conditional distribution of  $Y_i$  given a fixed exposure dose  $D_i$  is

where  $\lambda_i(D_i, \mathbf{b})$  is a thyroid cancer incidence, and  $\mathbf{b}$  is a vector of model parameters.

Next, because we know the probability for individuals to fall sick then one can construct the likelihood function and its maximum yields an estimator for the vector  $\mathbf{b}$  of unknown parameters. In case where errors in exposure doses are absent or in case where the presence of errors is ignored (the so-called naive estimation) the likelihood function for the logistic model is

**Full Maximum Likelihood Method (FMLM) application for risk estimation under the presence of multiplicative Berkson error in radiation exposure doses**

Let the multiplicative Berkson error be present in doses. Then

where  $D_i^{tr}$  is a true dose (unknown),  $D_i^{mes}$  is a dose estimate (known), and  $u_i$  is a Gaussian error.

Розрахуємо функцію вірогідності  $L(Y_i, D_i^{mes}, \mathbf{b})$  використовуючи наступні кроки. Нехай  $\Pr(Y_i, D_i^{mes}, D_i^{tr})$  сукупний імовірнісний розподіл величин  $Y_i, D_i^{mes}, D_i^{tr}$ . Тоді за формулою умовних ймовірностей одержимо:

$$\Pr(Y_i, D_i^{mes}, D_i^{tr}) = \Pr(Y_i | D_i^{mes}, D_i^{tr}) \cdot \Pr(D_i^{mes}, D_i^{tr}) = \Pr(Y_i | D_i^{tr}) \cdot \Pr(D_i^{tr} | D_i^{mes}) \cdot \Pr(D_i^{mes}). \quad (4)$$

Оскільки  $\Pr(D_i^{mes})$  не несе додаткової інформації про вектор параметрів  $\mathbf{b}$ , його можна опустити. Переходячи в (4) від розподілів імовірності до функції вірогідності та зважаючи на той факт, що умовний розподіл  $D_i^{tr} | D_i^{mes}$  є логнормальним:  $D_i^{tr} | D_i^{mes} \sim \text{Lognormal}(\ln D_i^{mes}, \sigma_i^2)$ , одержимо:

$$L(Y_i, D_i^{mes}, D_i^{tr}, \mathbf{b}) = \left( \frac{\lambda_i(D_i^{tr}, \mathbf{b})}{1 + \lambda_i(D_i^{tr}, \mathbf{b})} \right)^{Y_i} \left( \frac{1}{1 + \lambda_i(D_i^{tr}, \mathbf{b})} \right)^{1-Y_i} \frac{\exp\left(-\frac{(\ln D_i^{tr} - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{D_i^{tr} \sqrt{2\pi}\sigma_i}. \quad (5)$$

Для того, щоб позбутися невідомої змінної  $D_i^{tr}$ , необхідно за цією змінною зробити згортку:

$$L(Y_i, D_i^{mes}, \mathbf{b}) = \int_0^\infty \left( \frac{\lambda_i(t, \mathbf{b})}{1 + \lambda_i(t, \mathbf{b})} \right)^{Y_i} \left( \frac{1}{1 + \lambda_i(t, \mathbf{b})} \right)^{1-Y_i} \frac{\exp\left(-\frac{(\ln t - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{t\sqrt{2\pi}\sigma_i} dt. \quad (6)$$

Таким чином, функція вірогідності, що враховує внески всіх спостережень, буде дорівнювати:

$$L(\mathbf{b}) = \prod_{i=1}^N \int_0^\infty \left( \frac{\lambda_i(t, \mathbf{b})}{1 + \lambda_i(t, \mathbf{b})} \right)^{Y_i} \left( \frac{1}{1 + \lambda_i(t, \mathbf{b})} \right)^{1-Y_i} \frac{\exp\left(-\frac{(\ln t - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{t\sqrt{2\pi}\sigma_i} dt, \quad (7)$$

а відповідна логарифмічна функція вірогідності записується так:

$$l(\mathbf{b}) = \sum_{i=1}^N \ln \int_0^\infty \left( \frac{\lambda_i(t, \mathbf{b})}{1 + \lambda_i(t, \mathbf{b})} \right)^{Y_i} \left( \frac{1}{1 + \lambda_i(t, \mathbf{b})} \right)^{1-Y_i} \frac{\exp\left(-\frac{(\ln t - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{t\sqrt{2\pi}\sigma_i} dt. \quad (8)$$

З огляду на той факт, що змінна  $Y_i$  набуває всього двох значень 0 і 1, останню рівність можна записати у вигляді:

$$l(\mathbf{b}) = \sum_{i=1}^N \left[ Y_i \ln \int_0^\infty \left( \frac{\lambda_i(t, \mathbf{b})}{1 + \lambda_i(t, \mathbf{b})} \right) \frac{\exp\left(-\frac{(\ln t - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{t\sqrt{2\pi}\sigma_i} dt + (1 - Y_i) \ln \int_0^\infty \left( \frac{1}{1 + \lambda_i(t, \mathbf{b})} \right) \frac{\exp\left(-\frac{(\ln t - \ln D_i^{mes})^2}{2\sigma_i^2}\right)}{t\sqrt{2\pi}\sigma_i} dt \right]. \quad (9)$$

Let elaborate the likelihood function  $L(Y_i, D_i^{mes}, \mathbf{b})$  using the next steps. Let  $\Pr(Y_i, D_i^{mes}, D_i^{tr})$  be a joint probability distribution of  $Y_i, D_i^{mes}, D_i^{tr}$ . Then we get by the formula of conditional probabilities that

As  $\Pr(D_i^{mes})$  does not depend of the parameter vector  $\mathbf{b}$ , it can be omitted. Passing in (4) from probability distributions to likelihood function and taking into account that the conditional distribution  $D_i^{tr} | D_i^{mes}$  is lognormal,  $D_i^{tr} | D_i^{mes} \sim \text{Lognormal}(\ln D_i^{mes}, \sigma_i^2)$  we get

To get rid of unknown covariate  $D_i^{tr}$  it is necessary to convolve by the covariate:

Thus, the likelihood function which utilizes the contribution of all observations is equal to

and the corresponding log-likelihood function is

In a view of the fact that response  $Y_i$  takes only two values of 0 and 1 one can rewrite the latter equality as

В інтегралах (9) для кожного  $i=1, \dots, N$  можна зробити заміну:

$$dz = \frac{1}{t\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(\ln t - D_i^{mes})^2}{2\sigma_i^2}\right) dt,$$

$$z = \int_0^t \frac{1}{t\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(\ln t - D_i^{mes})^2}{2\sigma_i^2}\right) dt = G_i(t), \quad (10)$$

Тут  $z = G_i(t)$  – функція логнормального розподілу, тому

$$z(0) = 0, \quad z(\infty) = 1. \quad (11)$$

Підставляючи (10)–(11) у (9), одержимо:

$$l(\mathbf{b}) = \sum_{i=1}^N \left[ Y_i \ln \int_0^1 \left( \frac{\lambda_i(G_i^{-1}(z), \mathbf{b})}{1 + \lambda_i(G_i^{-1}(z), \mathbf{b})} \right) dz + (1 - Y_i) \ln \int_0^1 \left( \frac{1}{1 + \lambda_i(G_i^{-1}(z), \mathbf{b})} \right) dz \right]. \quad (12)$$

Інтеграли в (12) мають особливість в точці  $z = 1$ , оскільки функція  $G_i^{-1}(z) \rightarrow +\infty$  при  $z \rightarrow 1$ . Тим не менше ці інтеграли існують як абсолютно збіжні невластні інтеграли Рімана. Для їх обчислення можна застосувати метод Монте-Карло. При цьому можна скористатись співвідношенням:

$$G_i^{-1}(z) = \exp(D_i^{mes} + \sigma_i \Phi^{-1}(z)) \quad (13)$$

де  $\Phi(z)$  – функція розподілу стандартного нормального закону.

Таким чином, метод повної максимальної вірогідності полягає в знаходженні такого вектора параметрів  $\mathbf{b}$ , при якому функція вірогідності (12) набуває максимуму.

### Імітаційний стохастичний експеримент

При імітації даних використовувались дози ЩЗ реальної субпопуляції дітей і підлітків віком від 0 до 18 років (всього близько 13 тисяч осіб з населених пунктів Житомирської, Київської і Чернігівської областей України), які мали прямі вимірювання активності щитовидної залози в травні-червні 1986 р. Дози щитовидної залози для цієї субпопуляції були реконструйовані в рамках Міжнародного проекту по дослідженню тиреоїдозахворюваності після Чорнобильської аварії в Україні [1]. Для імітації захворюваності використовувалась лінійна модель абсолютного ризику:

$$\lambda_i = \lambda_0 + EAR \cdot D_i. \quad (14)$$

В даному випадку  $\lambda_0$  (фонова захворюваність) і  $EAR$  (ексцес абсолютного ризику) – це додатні параметри моделі, що мають бути оцінені. Спостережувана вибірка складається з пар  $(Y_i, D_i)$ ,  $i=1, \dots, N$ , де  $D_i$  – дози оп-

In integrals (9) a substitution can be made for each  $i=1, \dots, N$ :

Here  $z = G_i(t)$  is a cdf of lognormal law, therefore,

Substitute (10)–(11) into (9) and obtain

Integrals in (12) have singularity at point  $z = 1$  because  $G_i^{-1}(z) \rightarrow +\infty$  as  $z \rightarrow 1$ . Nevertheless these integrals exist as absolutely convergent improper Riemann integrals. It is possible to compute them by Monte Carlo method. At that one can use the relation

where  $\Phi(z)$  is a cdf of standard normal law.

Thus, Full Maximum Likelihood method consists in a search of parameter vector  $\mathbf{b}$  that maximizes the log-likelihood function (12).

### Simulated stochastic experiment

To simulate the data we used thyroid doses of real sub-cohort of children and adolescents aged from 0 to 18 years (totally around 13 thousand inhabitants of Zhytomyr, Kyiv and Chernihiv oblasts of Ukraine) for whom a direct measurement results of thyroid activity in May-June 1986 are available. Thyroid doses for this sub-cohort were reconstructed in the framework of the International project on study of thyroid cancer in Ukraine resulting from Chornobyl accident [1]. We used a linear absolute risk model to simulate prevalence

In this case  $\lambda_0$  (background morbidity) and  $EAR$  (excess absolute risk) are positive model parameters to be estimated. Observed sample consists of couples  $(Y_i, D_i)$ ,  $i=1, \dots, N$ , where  $D_i$  are thyroid doses (nonneg-



ромінення ЩЗ (невід'ємні числа),  $Y_i=1$  у разі наявності захворювання на рак ЩЗ на деякому часовому проміжку, а  $Y_i=0$  за відсутності цього захворювання.

При імітуванні випадків раку ЩЗ параметри моделі абсолютного ризику брались близькими до оцінок, одержаних у когортних дослідженнях онкологічної захворюваності ЩЗ в Україні [8–9]:

$$\lambda_0 = 10^{-4} \frac{\text{випадків}}{\text{людино - років}},$$

$$EAR = 5 \cdot 10^{-4} \frac{\text{випадків}}{(\text{Грей}) \cdot (\text{людино - років})}$$

При моделюванні приймалось, що похибка в дозі є берксонівською мультиплікативною похибкою, розподіленою за логнормальним законом (3).

Рівень берксонівської похибки задавався таким чином, що  $GSD = \exp(\sigma_i)$  (геометричне стандартне відхилення величини) приймало значення від 2 до 10 для всіх  $i=1, \dots, N$ . У процесі моделювання для кожного рівня похибки було згенеровано 1000 наборів даних.

Для оцінки параметрів регресії  $\lambda_0$  та  $EAR$  використовувався наївний метод оцінювання (тобто той, що ігнорує наявність похибок у дозах) та МПМВ.

## РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

Результати досліджень представлені на рис. 1. З нього видно, що при великих похибках у дозах опромінення наївні оцінки параметрів фонові захворюваності  $\lambda_0$  та надлишкового абсолютного ризику  $EAR$  істотно відхиляються від відповідних модельних (дійсних) значень. Ці відхилення залежать від величини дисперсії похибки. Так, при  $GSD < 3$  наївна оцінка  $EAR$  є близькою до модельного значення, але при подальшому збільшенні  $GSD$  наївна оцінка практично лінійно спадає. При  $GSD = 10$  наївна оцінка менша за істинне значення  $EAR$  майже вдвічі (див. рис. 1б). В англійській літературі це явище називається “attenuation effect” [4], тобто ефект заниження наївних оцінок надлишкового абсолютного ризику за наявності похибок у дозах. У той же час для фонові захворюваності спостерігається зворотний ефект. При значних похибках у дозах опромінення наївні оцінки  $\lambda_0$  дещо більші за істинні (модельні) значення цього параметра (див. рис. 1а).

Таким чином, за наявності похибок берксонівського типу у дозах опромінення наївні оцінки перерозподіляються наступним чином: для  $EAR$  вони нижчі за дійсні значення, тоді як для  $\lambda_0$  навпаки – вищі. Це може бути пов'язано з нелінійністю функції вірогідності (1). Слід зазначити, що у випадку берксонівсь-

active numbers),  $Y_i=1$  in case if person  $i$  has got thyroid cancer at fixed time interval, and  $Y_i=0$  if the person is not ill.

To simulate the thyroid cancer cases the parameters of the absolute risk model were taken close to estimates obtained in a cohort study of thyroid cancer prevalence in Ukraine [8–9]:

$$\lambda_0 = 10^{-4} \frac{\text{cases}}{\text{PY}}, \quad (15)$$

$$EAR = 5 \cdot 10^{-4} \frac{\text{cases}}{\text{Gy} \times \text{PY}}.$$

It was assumed in simulations that the dose error is Berkson and multiplicative distributed by log-normal law (3).

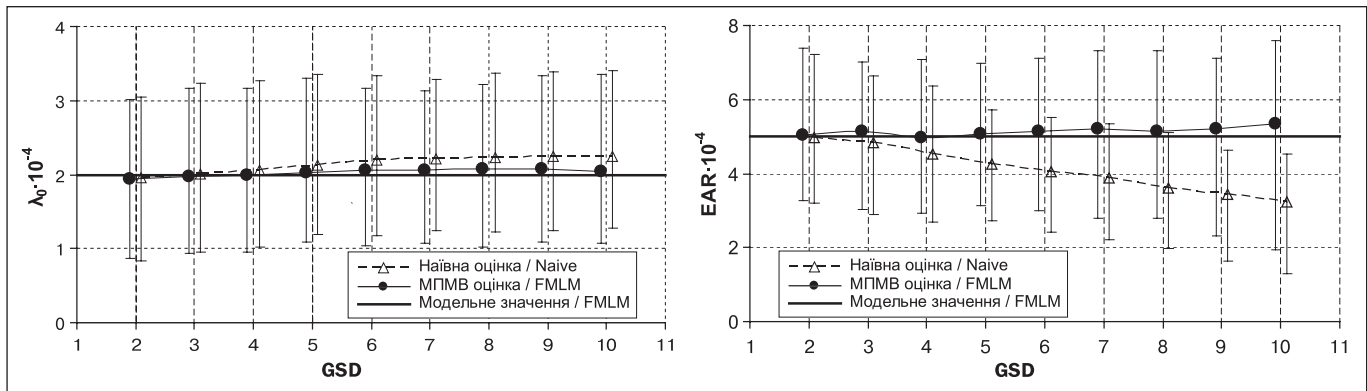
The level of Berkson error was taken such that  $GSD = \exp(\sigma_i)$  (geometric standard deviation) took values from 2 to 10, for all  $i=1, \dots, N$ . Within a modelling process we made 1,000 realizations for each level of error.

To estimate the regression parameters  $\lambda_0$  and  $EAR$  we used the naive estimation method (that ignores the errors in doses) and FMLM.

## RESULTS AND DISCUSSION

The results of research are presented in Figure 1. One can see that for large dose errors the naive estimates of the background morbidity  $\lambda_0$  and excess absolute risk  $EAR$  deviate significantly from the corresponding model (true) values. The deviations depend on the error variance. Thus for  $GSD < 3$  the naive estimate of  $EAR$  is close to the model value, but with further increasing of  $GSD$  the naive estimator decreases practically linearly. For  $GSD=10$  the naive estimate is almost twice less than the true value of  $EAR$  (see Fig. 1b). In the literature this phenomenon is called an “attenuation effect”, that is the effect of underestimation of excess absolute risk in the presence of dose errors. At that for background morbidity the opposite effect is observed. For considerable errors in exposure doses the naive estimates of are somewhat larger than the true (model) values of this parameter (see Fig. 1a).

Thus, under Berkson errors in exposure doses the naive estimates are redistributed as follows: for  $EAR$  they are lower than the true values, at that for  $\lambda_0$  vice versa they are higher. This can be related to nonlinearity of likelihood function (1). It should be noted that in case of a Berkson error



**Рисунок 1.** Залежності оцінок фонові захворюваності (а) та надлишкового абсолютного ризику (б) від величини берксонівської мультиплікативної похибки в дозах опромінення ЩЗ

**Figure 1.** Dependence of estimates for background morbidity (a) and excess absolute risk (b) from the level of Berkson multiplicative error in thyroid doses

кої похибки “attenuation effect” проявляється тільки при великих похибках і є не таким істотним, як у випадку класичної похибки [10].

Навіть при значних похибках у дозах оцінки обох параметрів регресії  $\lambda_0$  та  $EAR$  значно покращуються при використанні методу повної максимальної вірогідності, який враховує наявність похибки за допомогою інтегралу-згортки (8).

## ВИСНОВКИ

За допомогою імітаційного стохастичного моделювання у роботі досліджено вплив похибок берксонівського типу в дозах опромінення ЩЗ на оцінку ризиків радіоіндукованих стохастичних ефектів. Показано, що ігнорування значних берксонівських похибок у дозах призводить до зсунення оцінок фонові захворюваності  $\lambda_0$  та надлишкового абсолютного ризику  $EAR$ . У той же час ці зсунення є значно меншими, ніж у випадку класичної похибки. Використання МПМВ суттєво покращує оцінку як  $\lambda_0$ , так і  $EAR$ .

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Questionnaire- and measurement-based individual thyroid doses in Ukraine resulting from the Chernobyl nuclear reactor accident / I. Likhtarev, A. Bouville, L. Kovgan [et al.] // Radiat. Res. – 2006. – Vol. 166, No. 1, Pt 2. – P. 271–286.
2. Post-Chernobyl thyroid cancers in Ukraine. Report 1: Estimation of thyroid doses / I. Likhtarev, L. Kovgan, S. Vavilov [et al.] // Radiat. Res. – 2005. – Vol. 163, No 2. – P. 125–136.
3. Ron E. Uncertainties in radiation dosimetry and their impact on dose-response analysis : proc. of a workshop held September 3–5, 1997 in Bethesda, Maryland. NIH Publication No. 99-4541 / eds. E. Ron, F. O. Hoffman. – Bethesda, MD: National Cancer Institute, 1999.
4. Measurement error in nonlinear models. A modern perspective / R. J. Carroll, D. Ruppert, L. A. Stefanski, C. A. Crainiceanu – Boca Raton : Chapman and Hall / CRC, 2006. – 438 p.

the “attenuation effect” acts only for large errors and is not as significant as in case of a classical error [10].

Even for the significant dose errors the estimates of regression parameters  $\lambda_0$  and  $EAR$  are considerably improved by Full Maximum Likelihood method that takes into account the presence of an error via integral convolution (8).

## CONCLUSION

In the present paper we studied using a stochastic simulation the impact of Berkson error in thyroid doses on the of risk estimates for the radio-induced stochastic effects. It is shown that ignoring the significant Berkson errors in doses leads to the bias in estimates of the background morbidity  $\lambda_0$  and excess absolute risk  $EAR$ . At that the bias is essentially less than one for the case of classical error. The FMLM improves essentially the estimate of both  $\lambda_0$  and  $EAR$ .

## REFERENCES

1. Likhtarev I, Bouville A, Kovgan L, Luckyanov N, Voilleque P, Chepurny M. Questionnaire- and measurement-based individual thyroid doses in Ukraine resulting from the Chernobyl nuclear reactor accident. Radiat Res. 2006;166(1 Pt 2):271–86.
2. Likhtarov I, Kovgan L, Vavilov S, Chepurny M, Bouville A, Luckyanov N, Jacob P, Voilleque P, Voigt G. Post-Chernobyl thyroid cancers in Ukraine. Report 1: Estimation of thyroid doses. Radiat Res. 2005 Feb;163(2):125–36.
3. Ron E, Hoffman FO, editors. Uncertainties in Radiation Dosimetry and Their Impact on Dose-Response Analysis. Proceedings of a Workshop at the National Cancer Institute; 1997 Sep 3–5; Bethesda, Maryland. NIH Publication No. 99-4541. Bethesda, MD: National Cancer Institute; 1999.

5. Hofer, E. How to account for uncertainty due to measurement errors in an uncertainty analysis using Monte Carlo simulation / E. Hofer // Health Physics. – 2008. – Vol. 95, No 3. – P. 277–290.
6. Methods for estimation of radiation risk in epidemiological studies accounting for classical and Berkson errors in doses / A. Kukush, S. Shklyar, S. Masiuk [et al.] // Int. J. Biostat. – 2011. – Vol. 7, Iss. 1. – P. 1–30.
7. Hosmer D. Applied Logistic Regression / D. Hosmer, S. Lemeshow – New-York : John Wiley and Sons, 2000. – 375 p.
8. A cohort study of thyroid cancer and other thyroid diseases after the Chernobyl accident: thyroid cancer in Ukraine detected during first screening / M. D. Tronko. G. R. Howe, T. I. Bogdanova [et al.] // J. Natl. Cancer Inst. – 2006. – Vol. 98, Iss. 13. – P. 897–903.
9. Post-Chernobyl thyroid cancers in Ukraine. Report 2: Risk analysis / I. Likhtarev, L. Kovgan, S. Vavilov [et al.] // Radiat. Res. – 2006. – Vol. 166, No 2. – P. 375–386.
10. Масюк, С.В. Вплив похибок вимірювання в дозах опромінення щитоподібної залози на оцінку радіаційних ризиків / С.В. Масюк, С.В. Шкляр, О.Г. Кукуш // Проблеми радіаційної медицини та радіобіології. – 2011. – Вип. 16. – С. 25–29.
4. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CA. Measurement Error in Nonlinear Models. A Modern Perspective. Boca Raton: Chapman and Hall; 2006. 438 p.
5. Hofer E. How to account for uncertainty due to measurement errors in an uncertainty analysis using Monte Carlo simulation. Health Physics. 2008;95(3):277–90.
6. Likhtarov I, Kovgan L, Carroll RJ, Bouville A. Methods for estimation of radiation risk in epidemiological studies accounting for classical and Berkson errors in doses. Int J Biostat. 2011 Feb; 7(1):1–30.
7. Hosmer D, Lemeshow S. Applied Logistic Regression. New-York: John Wiley and Sons; 2000. 375 p.
8. Tronko MD, Howe GR, Bogdanova TI, Bouville AC, Epstein OV, Brill AB, et al. A cohort study of thyroid cancer and other thyroid diseases after the Chernobyl accident: thyroid cancer in Ukraine detected during first screening. J Natl Cancer Inst. 2006 Jul 5; 98(13):897–903.
9. Likhtarov I, Kovgan L, Vavilov S, Chepurny M, Ron E, Lubin J, et al. Post-Chernobyl thyroid cancers in Ukraine. Report 2: risk analysis. Radiat Res. 2006 Aug;166(2):375–86.
10. Masiuk S, Shklyar S, Kukush A. [The influence of measurement errors in irradiation doses of thyroid gland to assess radiation risks]. Problems of Radiation Medicine and Radiobiology. 2011; (16):25–9. Ukrainian.

*Стаття надійшла до редакції 05.08.2013*

*Received: 05.08.2013*