

*Оригінальна стаття**МЕДИЧНА ФІЗИКА*<https://doi.org/10.26565/2075-3810-2018-40-05>

УДК 577.3:616-006:575.17:519.22/25

## МЕТОДИКА ЙМОВІРНІСНОГО ВИЗНАЧЕННЯ ДІЙСНОЇ КІЛЬКОСТІ ОНКОЛОГІЧНИХ ХВОРИХ НА ОСНОВІ СТАТИСТИЧНИХ ДАНИХ

**М.А. Бондаренко***Харківський національний медичний університет, пр. Науки, 4, м. Харків, 61022, Україна**e-mail: [bondaren.koma3007@gmail.com](mailto:bondaren.koma3007@gmail.com)*

Надійшла до редакції 22 травня 2018 р.

Прийнято 10 вересня 2018 р.

**Актуальність.** Експерти в області онкології визнають, що шанси на повне вилікування від раку зростають, якщо захворювання вдається виявити на ранніх стадіях. Але щоб виявити хворобу в зародку, необхідно не нехтувати діагностичним обстеженням, яке більшість громадян нашої країни ігнорує. В даний час онкопатологія займає друге місце в світі після серцево-судинних захворювань в переліку хвороб. Україна на другому місці в Європі за темпами розповсюдження раку. Щорічно в Україні понад 160 тис. людей узнають, що вони онкохворі. В дійсності кількість таких хворих значно більша, ніж статистично виявлена. Актуальність роботи обумовлена необхідністю розробки методики визначення дійсної кількості онкохворих, що могло б поліпшити механізми раннього виявлення раку та збільшити кількість вилікованих хворих.

**Мета роботи.** Розробка методики ймовірнісного визначення дійсної онкозахворюваності населення на основі відомих статистичних даних про загальну смертність та кількість виявлених онкологічних хворих.

**Результати.** Використовуючи ймовірнісний підхід, спочатку розраховано ймовірність загибелі індивідів, які вже мали недиагностоване онкологічне захворювання у певному віці, але загинули внаслідок причин неонкологічного характеру. Далі отримано формулу розрахунку дійсної онкологічної захворюваності індивіда у певному віці. Використовуючи статистичні дані про загальну кількість померлих та про кількість виявлених онкологічних хворих у певному віці, а також використовуючи вищезазначену формулу, було одержано графік розподілу дійсної кількості онкохворих в залежності від віку. Це дало можливість розрахувати відношення сумарної дійсної кількості онкологічних онкохворих до сумарної виявленої кількості онкохворих, яке склало величину 1,95, що свідчить про значне перевищення дійсної онкологічної захворюваності над статистично виявленою.

**Висновки.** Запропоновано методику ймовірнісної оцінки дійсної онкологічної захворюваності на підставі статистично виявленої онкозахворюваності та загальної смертності. Проведені розрахунки свідчать, що показники виявленої онкологічної захворюваності майже у два рази менші за показники дійсної кількості онкохворих.

**КЛЮЧОВІ СЛОВА:** онкологія; загальна смертність індивідів; статистично виявлена та дійсна кількість онкохворих.

## METHOD OF PROBABILITY DETERMINING THE ACTUAL NUMBER OF CANCER PATIENTS USING STATISTICAL DATA

**M.A. Bondarenko***Kharkiv National Medical University, 4 Nauky Ave., Kharkiv, 61022, Ukraine*

**Background:** Experts in the field of oncology recognize that the chances of a complete cure for cancer grow if the disease can be identified at an early stage. But in order to identify the disease in the bud, it is necessary not to neglect the diagnostic examination that most citizens of our country ignore. Currently, cancer takes second place in the world after cardiovascular diseases in the list of other diseases. Ukraine is second in Europe in velocity of cancer propagation. Every year in Ukraine more than 160 thousand people learn that they are cancer patient. In fact, the number of such patients is much larger than statistically revealed. The urgency of the work is due to the need to develop a methodology for determining the actual number of cancer patients, which could improve the mechanisms of early detection of cancer and increase the number of cured patients.

**Objective:** To develop an approach to probabilistic determination of the actual oncological morbidity of the population on the basis of known statistical data on the overall mortality and the number of detected oncological patients.

**Results:** Using the probabilistic approach, the probability of death of individuals who already had undiagnosed oncological disease at a certain age was calculated first, but died as a result of non-cancer reasons. Further, a formula for calculating the actual oncological morbidity of an individual at a certain age was obtained. Using the statistical data on the total number of deaths and the number of detected cancer patients at a certain age, and using the abovementioned formula, a graph of the distribution of the actual number of cancer patients, depending on age, was obtained. This allowed us to calculate the ratio of the total actual number of oncological cancer patients to the total number of cancer patients. It was value 1.95, which indicates a significant excess of actual cancer morbidity over statistically detected.

**Conclusions:** A probabilistic approach to assessing the actual oncological morbidity based on the statistically revealed oncological morbidity and overall mortality is proposed. The performed calculations show that the indicators of the detected oncological morbidity are almost two times less than the actual number of cancer patients.

**KEY WORDS:** oncology; total mortality of individuals; diagnosed and actual number of cancer patients.

## МЕТОДИКА ВЕРОЯТНОСТНОГО ОПРЕДЕЛЕНИЯ ДЕЙСТВИТЕЛЬНОГО КОЛИЧЕСТВА ОНКОЛОГИЧЕСКИХ БОЛЬНЫХ НА ОСНОВЕ СТАТИСТИЧЕСКИХ ДАННЫХ

М.А. Бондаренко

*Харьковский национальный медицинский университет, пр. Науки, 4, г. Харьков, 61022, Украина*

**Актуальность.** Эксперты в области онкологии признают, что шансы на полное излечение от рака растут, если заболевание удастся выявить на ранних стадиях. Но чтобы выявить болезнь в зародыше, необходимо не пренебрегать диагностическим обследованием, которое большинство граждан нашей страны игнорирует. В настоящее время онкопатология занимает второе место в мире после сердечно-сосудистых заболеваний в перечне болезней. Украина на втором месте в Европе по темпам распространения рака. Ежегодно в Украине более 160 тыс. человек узнают, что они онкобольные. В действительности количество таких больных значительно больше, чем статистически выявлено. Актуальность работы обусловлена необходимостью разработки методики определения действительного количества онкобольных, что могло бы улучшить механизмы раннего выявления рака и увеличить количество излеченных больных.

**Цель работы.** Разработка методики вероятностного определения действительной онкозаболеваемости населения на основе известных статистических данных об общей смертности и количестве выявленных онкологических больных.

**Результаты.** Используя вероятностный подход, сначала рассчитана вероятность гибели индивидов, которые уже имели не диагностированное онкологическое заболевание в определенном возрасте, но погибли в результате причин неонкологического характера. Далее получена формула расчета действительной онкологической заболеваемости индивида в определенном возрасте. Используя статистические данные об общем количестве умерших и о количестве выявленных онкологических больных в определенном возрасте, а также используя вышеуказанную формулу, был получен график распределения действительного количества онкобольных в зависимости от возраста. Это позволило рассчитать отношение суммарного действительного количества онкобольных к суммарному выявленному количеству онкобольных, которое составило величину 1,95, что свидетельствует о значительном превышении действительной онкологической заболеваемости над статистически выявленной.

**Выводы.** Предложена методика вероятностной оценки действительной онкологической заболеваемости на основании статистически выявленной онкозаболеваемости и общей смертности. Проведенные расчеты свидетельствуют, что показатели выявленной онкологической заболеваемости почти в два раза меньше показателей действительного количества онкобольных.

**КЛЮЧЕВЫЕ СЛОВА:** онкология; общая смертность индивидов; выявленное и действительное количество онкологических больных.

Зростання онкологічної захворюваності в світі обумовлює підвищення інтересу до наукових досліджень, пов'язаних з онкологічною проблематикою, зокрема, до використання математичного моделювання для чисельної оцінки значень статистичних характеристик онкологічної захворюваності. До однієї з таких найпростіших (і в той же час – найважливіших) характеристик відноситься ймовірність виникнення

онкологічного захворювання у довільно обраного з популяції індивіда на певному році його життя.

Статистичні дані, що відносяться до онкологічної захворюваності, широко відомі [1-10]. Однак слід мати на увазі те, що частина індивідів, що вже мають онкологічне захворювання, можуть загинути до його виявлення з причин, не пов'язаних з онкологією (наприклад, внаслідок серцевого нападу або в результаті нещасного випадку).

У зв'язку з цим можна говорити про те, що існують характеристики онкологічної захворюваності, які статистично виявлені, та дійсні. Виявлені характеристики - це офіційна статистика, яка не враховує тих хворих, що загинули до того, як у них могло бути діагностовано онкологічне захворювання. Що стосується дійсної захворюваності, то вона включає і тих індивідів, у яких онкологічне захворювання було виявлено, і вони потрапили в статистичні дані, і тих, у яких не було виявлено онкологічного захворювання внаслідок того, що вони загинули раніше, ніж воно у них було діагностовано. У зв'язку зі сказаним зрозуміло, що величина дійсної захворюваності більша за статистично виявлену.

Оцінити величину дійсної онкологічної захворюваності можна, використовуючи ймовірнісний підхід.

Питання про дійсну та виявлену онкологічну захворюваність населення в науковій літературі не розглядалося, а, отже, не було запропоновано будь-якої методики визначення дійсної кількості онкохворих.

Метою дослідження стала розробка методики визначення дійсної онкозахворюваності населення на основі відомих статистичних показників онкозахворюваності та загальної смертності.

### **МЕТОДИКА ВИЗНАЧЕННЯ ДІЙСНОЇ КІЛЬКОСТІ ОНКОЛОГІЧНИХ ХВОРИХ**

Введемо деякі позначення. Нехай  $k$  – вік (в роках) індивіда, що довільно обраний з популяції людей;  $s_k$  – загальна (за будь-якими можливими причинами) емпірична ймовірність загибелі індивіда у віці  $k$ ;  $n_k$  – величина кількості виявлених онкологічних захворювань у популяції у віці  $k$ , виходячи з відомих статистичних даних;  $i_k$  – величина дійсної кількості онкологічних захворювань, що утворилися в віці  $k$ , розрахована з використанням математичного моделювання.

Спочатку знайдемо величину кількості загиблих індивідів, які вже мали ще не діагностоване онкологічне захворювання у віці  $k$ , але загинули внаслідок причин неонкологічного характеру. З імовірнісної точки зору ця величина дорівнює добутку величин  $s_k$  та  $i_k$ . Тоді

$$i_k = n_k + s_k \cdot i_k. \quad (1)$$

Перетворюючи вираз (1), отримуємо формулу розрахунку дійсної онкологічної захворюваності для індивідів у віці  $k$ :

$$i_k = \frac{n_k}{1 - s_k}. \quad (2)$$

Виходячи з цієї формули і маючи дані про розподіл величин  $n_k$  і  $s_k$  в залежності від віку, можна отримати розподіл дійсної кількості ( $i_k$ ) онкологічних захворювань, що утворилися в віці  $k$ .

На рис. 1 представлений графік розподілу загальної емпіричної ймовірності  $s_k$  загибелі індивіда в залежності від віку.

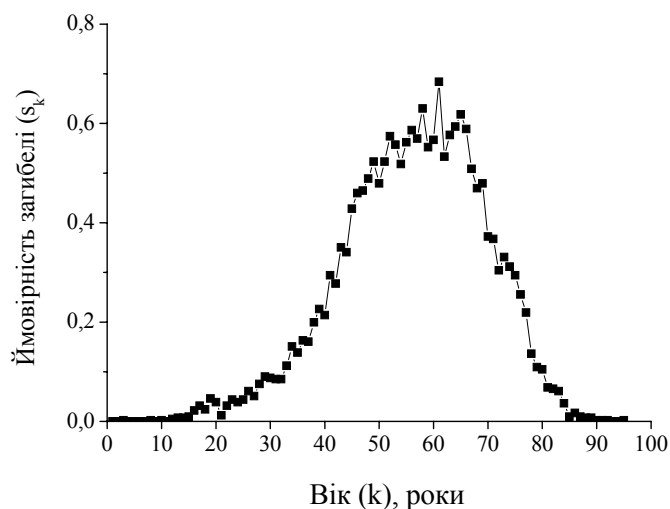


Рис. 1. Залежність емпіричної ймовірності  $S_k$  загибелі індивіда від віку індивіда.

Значний розкид даних, показаних на рис. 1, пов'язаний з невисоким обсягом вибірки наявних даних (411 осіб). Дані одержані у Головному управлінні статистики у Харківській області.

На рис. 2 представлені графіки розподілу величин  $i_k$  (дійсна захворюваність) і  $n_k$  (виявлена захворюваність) в залежності від віку  $k$ . Дані по кількості виявлених онкохворих (8288 осіб) були отримані в архіві Харківського інституту медичної радіології ім. С. П. Григор'єва, м. Харків, 2004-2006 р.р. Розрахунки дійсної кількості онкохворих (значення величини  $i_k$ ) були проведені за допомогою табличного процесора з використанням останньої формули для обчислення величини  $i_k$  та на основі відомих статистичних даних про розподіл виявленої кількості онкохворих ( $n_k$ ) та відомих даних про загальну смертність в залежності від віку (рис. 1).

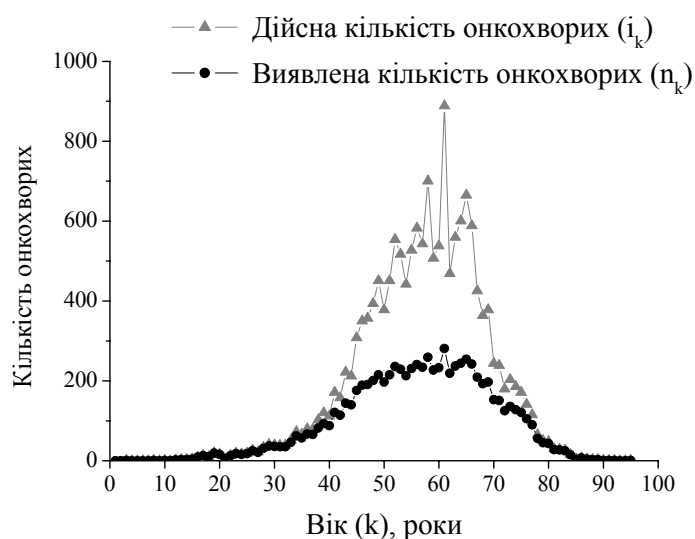


Рис. 2. Графіки розподілу виявленої кількості  $n_k$  онкологічних захворювань у популяції в віці  $k$  та дійсної кількості  $i_k$  онкологічних захворювань у віці  $k$ .

Для оцінки того, наскільки різняться значення показників виявленої та дійсної онкозахворюваності, розраховуємо суму ( $N$ ) значень величини  $n_k$  і суму ( $I$ ) значень  $i_k$  (2) за весь період життя індивіда до віку  $K$  – максимального віку індивіда в вибірці:

$$N = \sum_{k=1}^K n_k ; I = \sum_{k=1}^K i_k . \quad (3)$$

В наших розрахунках для величини  $K$  приймалося значення  $K = 95$  років, що визначалося наявними у нас експериментальними даними. Для величин  $N$  та  $I$  були одержані значення  $I = 16151$ ,  $N = 8288$ . Тоді відношення дійсної кількості онкохворих до виявленої кількості онкохворих дорівнює

$$\frac{I}{N} = 1,95 .$$

Як бачимо, величина дійсної онкологічної захворюваності з урахуванням наших регіональних особливостей (Харківщина, Україна) майже в два рази перевищує величину статистично виявленої захворюваності.

З огляду на досить невеликий обсяг даних, використаних у роботі, отримані результати можуть розглядатися тільки в якості оціночних даних.

В даній статті запропоновано саме методика дослідження величини дійсної онкологічної захворюваності. Зрозуміло, що для жителів різних країн результати можуть істотно відрізнятися. Так, в країнах з кращими діагностичними можливостями у порівнянні з Україною це співвідношення повинно бути меншим, ніж 1,95.

## ВИСНОВКИ

1. Аналізуючи статистику онкозахворюваності необхідно розрізняти поняття виявленої та дійсної захворюваності, оскільки дійсна захворюваність враховує не тільки наявні статистичні дані про кількість онкохворих, але й тих онкологічних хворих, які померли до того, як це захворювання було у них діагностовано.


2. В роботі запропонована методика ймовірнісної оцінки величини дійсної онкологічної захворюваності на підставі статистично виявленої онкозахворюваності та загальної смертності.

3. Проведені розрахунки показують, що різниця між показниками виявленої онкологічної захворюваності та дійсної є значною.

## КОНФЛІКТ ІНТЕРЕСІВ

Автори повідомляють про відсутність конфлікту інтересів.

### Author's ORCID ID

М.А. Bondarenko  <https://orcid.org/0000-0002-0203-0161>

## REFERENCES

1. Zavarykina, T.M. (2015). Issledovanie polimorfnykh markerov genov reparatsii DNK hOGG1, XRCC1 i ERCC2 i ikh assotsiatcii s riskom razvitiia nemelkokletochnogo raka legkogo i raka verkhnikh dykhatelnykh putei. *Estestvennye i matematicheskie nauki v sovremennom mire*, 2(26), 1-11. (in Russian).
2. Klark, M., & Beker, M. (2006). Rakovye stvolovye kletki. *V mire nauki*, 10, 29-35. (in Russian).
3. Leshchenko Ia.A. (2014). Onkologicheskaiia zaboлеваemost naseleniia promyshlennogo goroda Irkutskoi oblasti. *Mezhdunarodnyi zhurnal prikladnykh i fundamentalnykh issledovaniy*, 11(4), 666-670. Retrieved from <https://www.applied-research.ru/ru/article/view?id=6205> (in Russian).
4. International Agency for Research on Cancer. (2012). *All Cancers (excluding non-melanoma skin cancer) Estimated Incidence, Mortality and Prevalence Worldwide in 2012*. Retrieved from [http://globocan.iarc.fr/Pages/fact\\_sheets\\_cancer.aspx](http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx).
5. Hanahan, D., & Weinberg, R.A. (2000). The hallmarks of cancer. *Cell*, 100, 57-70.

6. Bravo, R., & Axelrod, D. E. (2013). A calibrated agent-based computer model of stochastic cell dynamics in normal human colon crypts useful for in silico experiments. *Theoretical Biology and Medical Modelling*, 10(66), 1-24.
7. Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA Cancer J. Clin.*, 65, 5-29.
8. Siegel, R. L., Miller, K. D., & Jemal, A. (2017). Cancer statistics, 2017. *CA Cancer J. Clin.*, 67(1), 7-30.
9. Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., & Matrisian, L. M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.*, 74(11), 2913-2921.
10. Chen, W., Zheng, R., & Baade, P. D. (2016). Cancer statistics in China, 2015. *CA Cancer J. Clin.*, 66(2), 115-132.