

УДК 616.44-006-089.17-036

ЄВГЕНІЯ БОРИСІВНА РАДЗІШЕВСЬКА<sup>1,2</sup>

<sup>1</sup> ДУ «Інститут медичної радіології ім. С. П. Григор'єва НАМН України», Харків

<sup>2</sup> Харківський національний медичний університет

## РЕЗУЛЬТАТИ РОЗВІДУВАЛЬНОГО АНАЛІЗУ ВІДДАЛЕНИХ НАСЛІДКІВ ЛІКУВАННЯ ХВОРИХ НА РАК ЩИТОПОДІБНОЇ ЗАЛОЗИ

**Мета роботи.** Проведення розвідувального аналізу бази даних перебігу та післялікувального скринінгу хворих на рак щитоподібної залози (РЩЗ) для виявлення статистично значущих залежностей між катамнестичними даними анамнезу хворих і перебігом захворювання та фактом появи відстрочених онкологічних наслідків лікування захворювання.

**Матеріали і методи.** Проаналізовано 215 паперових історій хвороби та внесено до електронної бази даних інформацію про перебіг захворювання 68 пацієнтів за критеріями або появи будь-якого злоякісного захворювання через 3 та більше років після проведеного лікування, або появи віддалених терапевтичних ускладнень. Онкологічні наслідки мали 33 пацієнти, у 35 осіб були лише терапевтичні наслідки. Статистичну обробку проводили методами непараметричної статистики за допомогою пакета програм Statistica Basic Academic 13 for Windows.

**Результати.** Хворі на РЩЗ, які на віддалених термінах після лікування мають онкологічні наслідки, характеризуються більш високим індексом маси тіла; у пацієнтів жіночої статі загальний термін лактації за час репродуктивного періоду у 2,3 разу перевищує аналогічні показники жінок без онкологічних наслідків; *prolongatio morbi* є більш характерною для пацієнтів з III стадією процесу, а відсутність онкологічних наслідків та поява других раків — для пацієнтів з II стадією процесу; гірші результати на віддалених термінах після спеціального лікування мали пацієнти з першою групою крові.

**Висновки.** Проведений розвідувальний аналіз катамнестичних даних дозволив виявити низку статистично значущих залежностей, які не були очевидними без відповідної математичної обробки. Розширення обсягу робочого масиву електронних даних дозволить одержати додаткову інформацію, що має клінічну та наукову цінність.

**Ключові слова:** медична інформатика, рак щитоподібної залози, віддалені наслідки лікування, статистичний аналіз.

Реаліями сьогодення є безперервно зростаючий рівень активності у царині біомедичної та медичної інформатики (електронний пропис ліків, електронна медична документація, особисті медичні записи тощо). Ці процеси є джерелом великої кількості даних, які можна використовувати для удосконалення знань і створення доказів з метою поліпшення клінічної та профілактичної допомоги.

Науковці з медичної інформатики розробляють, впроваджують та оцінюють нові біомедично мотивовані методи в таких ділянках, як *data mining* (отримання нових знань з великих баз даних), обробка текстів, когнітивістика, підтримка прийняття рішень, бази даних та алгоритми аналізу великих обсягів даних, що були отримані у сфері охорони здоров'я. Вся робота в галузі інформатики обумовлена необхідністю розробки нових рішень на підґрунті інформаційних технологій, що сприяє збереженню здоров'я населення

та підвищенню якості та безпеки медичного обслуговування хворих [1].

Зокрема, підходи та методи медичної інформатики дозволяють одержувати нові знання завдяки сучасній обробці архівних паперових історій хвороби, які трансформуються в електронні масиви формалізованої медичної інформації.

З огляду на термінологію доказової медицини, дослідження, результати якого пропонуються, належало до ретроспективних досліджень випадок-контроль, в якому аналізувалися оцифровані дані перебігу та післялікувального скринінгу хворих на рак щитоподібної залози (РЩЗ). Метою роботи було проведення розвідувального аналізу над усім масивом даних для виявлення статистично значущих залежностей між катамнестичними даними анамнезу хворих і перебігу захворювання й фактом появи відстрочених онкологічних наслідків лікування захворювання. За умови подальшої адекватної клінічної інтерпретації фахівцями відповідного профілю,

одержані результати можуть надати підґрунтя для удосконалення медичних знань у напрямку «онкологія».

### МЕТОДИКА ДОСЛІДЖЕННЯ

Для формування електронних масивів катанестичних даних використовували власний програмний продукт — систему «База даних хворих», яка дозволяє накопичувати медичну інформацію будь-якого профілю для її подальшої математичної обробки у будь-якому програмному середовищі.

Система відповідає таким вимогам:

- повна формалізація інформації. Ця вимога припускає відсутність у базі даних будь-якої описової інформації. Нечислова інформація попередньо формалізується у вигляді довідкових таблиць і перебуває в базі винятково у вигляді посилання на відповідний рядок таблиці-довідника. Виняток становлять лише анкетні дані пацієнта;

- гнучка структура дозволяє у будь-який час додавати нові види досліджень та модифікувати вже існуючі без утручання у програмний код;

- висока швидкодія, відсутність обмежень на обсяг інформації, яка зберігається.

За період роботи з базою катанестичних даних було проаналізовано 215 паперових історій хвороби та відібрано інформацію про перебіг захворювання 68 пацієнтів для внесення до електронної бази даних. Критерієм відбору була або поява будь-якого злоякісного захворювання через 3 та більше років після проведеного лікування, або поява віддалених терапевтичних ускладнень.

Онкологічні наслідки (ОН) мали 33 пацієнти, в тому числі: метастази — 19 осіб (58%); другі раки — 7 осіб (21%); рецидиви — 7 осіб (21%). Серед хворих налічувалося 5 чоловіків (15%) та 28 жінок (85%). Вік хворих коливався від 23 до 62 років з медіаною 42 роки та інтерквартильним розмахом 40 ÷ 59 років. Загальна тривалість радіонуклідного лікування пацієнтів з ОН становила (медіальні значення) 37,5 міс., під час якого пацієнти одержали 4605 МБк активності протягом 5 курсів.

По відношенню до активної групи (пацієнтів з ОН) як контрольну розглядали пацієнтів, що мали лише терапевтичні наслідки (ТН) — 35 осіб, 164 різних терапевтичних наслідки. За статевим складом у групі з ТН було 2 чоловіків (5%) та 33 жінки (95%). Вік хворих лежав у межах 13 ÷ 68 років з медіаною 44 роки. Загальна тривалість радіонуклідного лікування пацієнтів з ТН становила (медіальні значення) 30,0 міс., під час якого пацієнти одержали 4685 МБк активності протягом 4 курсів.

Статистичну обробку проводили методами непараметричної статистики за допомогою пакета програм Statistica Basic Academic 13 for Windows.

### РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

На сформованому масиві даних було проведено розвідувальний статистичний аналіз із метою виявлення всіх наявних статистично значущих залежностей.

При аналізі у розрізі «ОН» та лише «ТН» із використанням критерію Манна-Уїтні (КМУ) виявлено дві статистично значущі залежності.

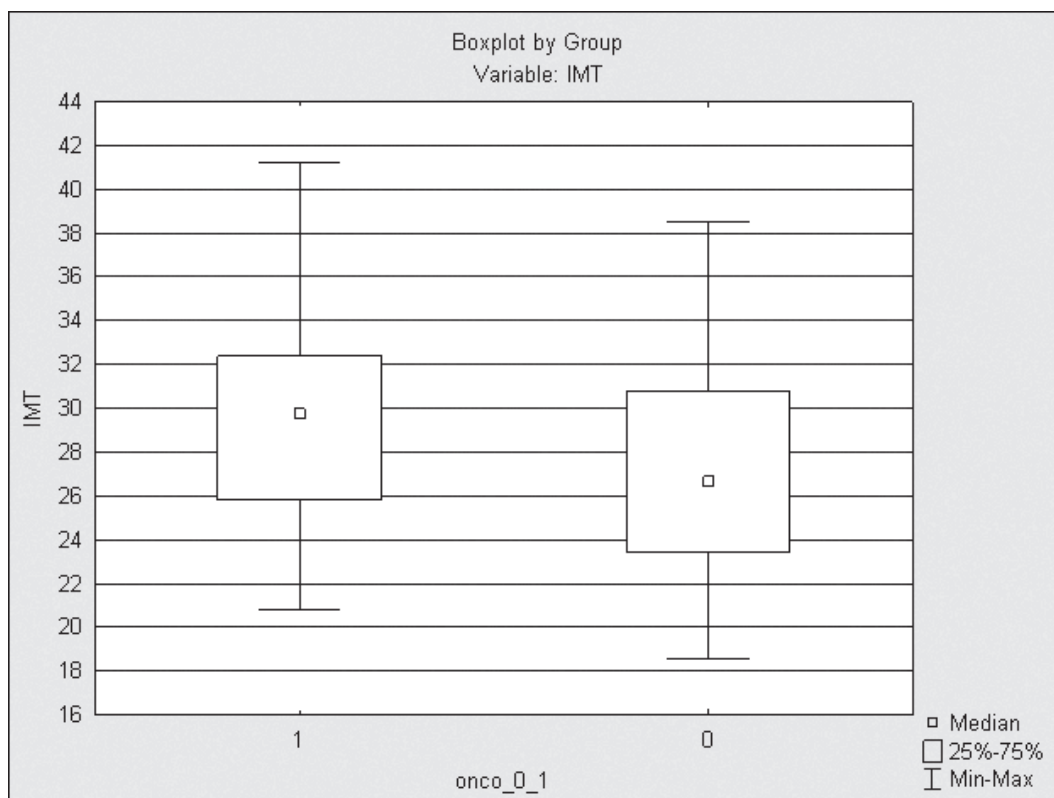


Рис. 1. Коробчасті графіки залежності ІМТ від виду віддалених наслідків лікування хворих на РЦЗ (1 — ОН, 0 — ТН)

По-перше, констатовано наявність зв'язку між індексом маси тіла (ІМТ), що розраховувався за Кетле, та появою ОН — у хворих з ОН показник ІМТ вище у 1,2 разу за аналогічний показник у хворих із ТН (рис. 1).

З рис. 1 можна бачити, що у хворих на РЦЗ без ОН інтерквартильний проміжок належав діапазону преожиріння (табл. 1) з медіальним значенням 27 кг/м<sup>2</sup>, а у пацієнтів з ОН — преожиріння та ожиріння I класу з медіаною 29,5 кг/м<sup>2</sup>. Таким чином, надмірне підвищення маси тіла є фактором ризику *prolongatio morbi* або появи других раків.

Таблиця 1

**Класифікація у дорослих індивідуумів зниженої і підвищеної маси тіла та ожиріння по відношенню до індексу маси тіла (довідкова інформація)**

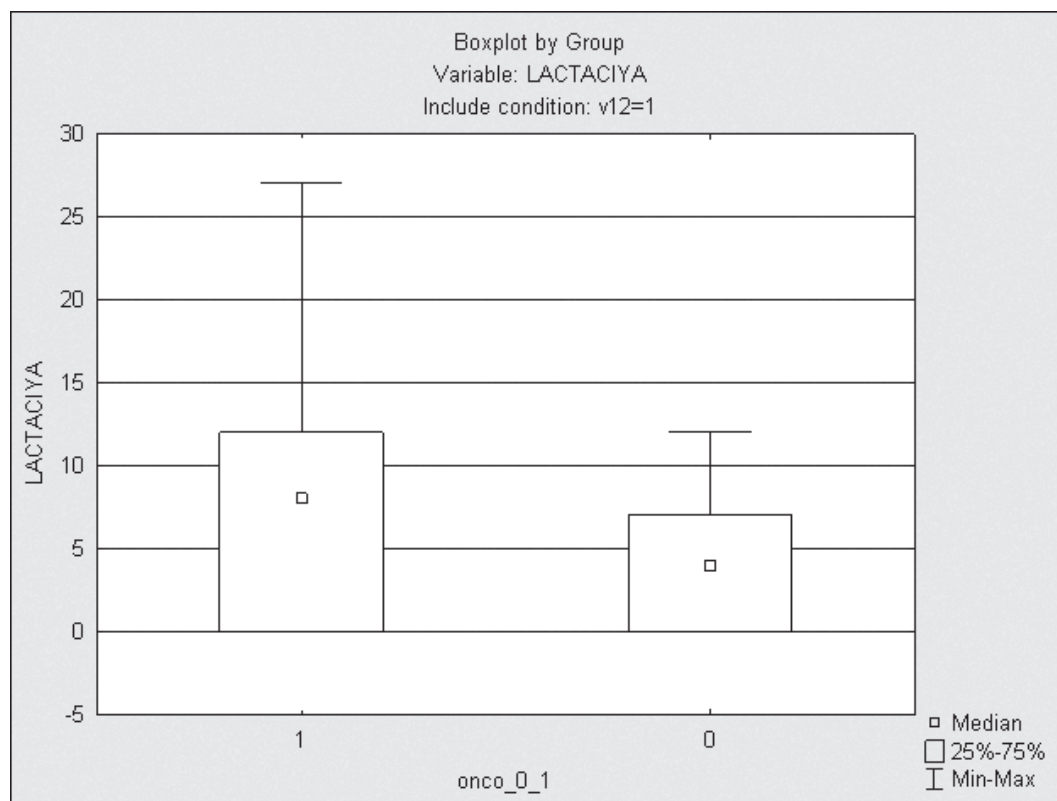
Класифікація	Значення ІМТ
Недостатність ваги	< 18,50
- виражена	< 16
- помірна	16,00–16,99
- легка	17,00–18,49
Нормальна вага	18,50–24,99
Підвищена вага	≥ 25,00
- преожиріння	25,00–29,99
- ожиріння	≥ 30,00
- ожиріння класу I	30,00–34,99
- ожиріння класу II	35,00–39,99
- ожиріння класу III	≥ 40,00

Доволі несподіваною стала наявність статистично значущої залежності у пацієнтів жіночої статі зв'язків між віддаленими результатами лікування та сумарною тривалістю лактаційного періоду за репродуктивний проміжок життя як складової гінекологічного анамнезу: у хворих із ОН термін лактації доходив до 27 міс. (медіальне значення 8 міс.), а за відсутності ОН — не перевищував 12 міс. із медіаною 3,5 міс. (рис. 2).

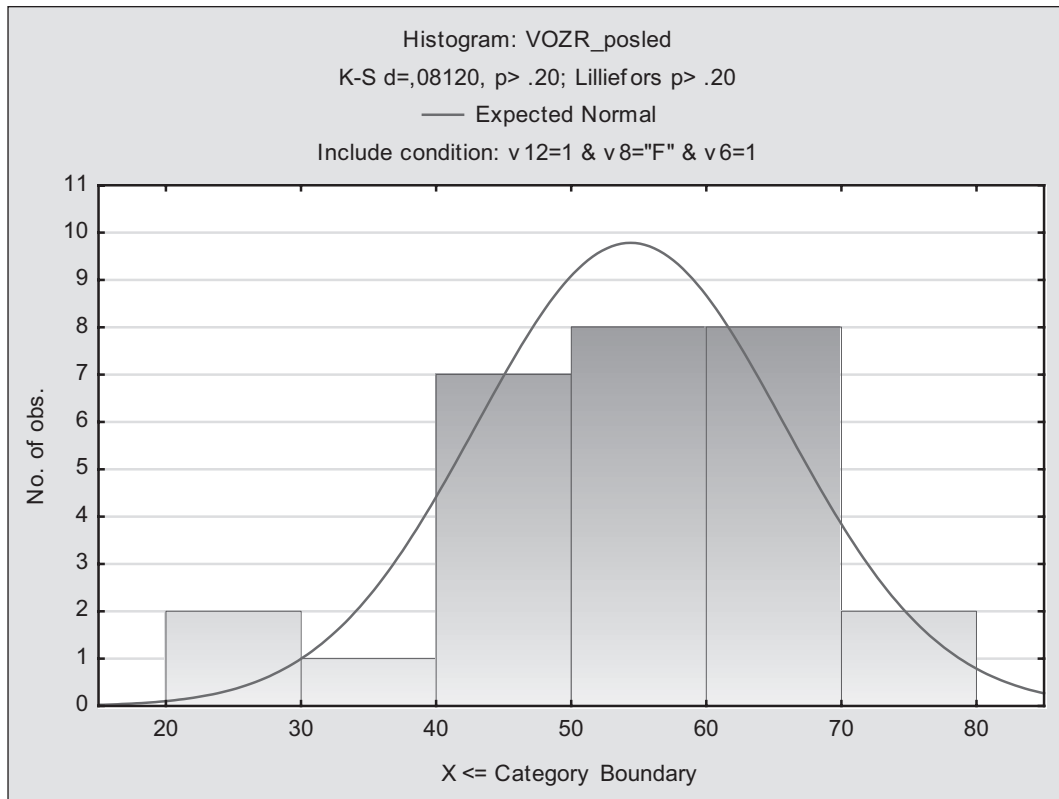
Нами не знайдено досліджень, повністю аналогічних нашому, проте зв'язок між грудним годуванням та перебігом захворювання у жінок, хворих на РЦЗ, у світовій літературі дискутується доволі часто. При цьому результати часом є суперечливими, а інколи навіть протилежними.

Так, наприклад, Xingyang Yi зі співавт. [2] стверджують, що грудне вигодовування потенційно знижує ризик розвитку раку щитоподібної залози, та чим більш тривалим є грудне вигодовування, тим більше знижується ризик. Протилежної думки додержуються Christopher W Rowe зі співавт. [3]. На підставі своїх досліджень вони доводять, що епідеміологічні та клітинні дані підтверджують роль естрогену як фактора росту доброякісних і злоякісних тироцитів, але визначення величини цього ризику для вагітних пацієнток і осіб, які здійснюють грудне вигодовування, залишається складним завданням.

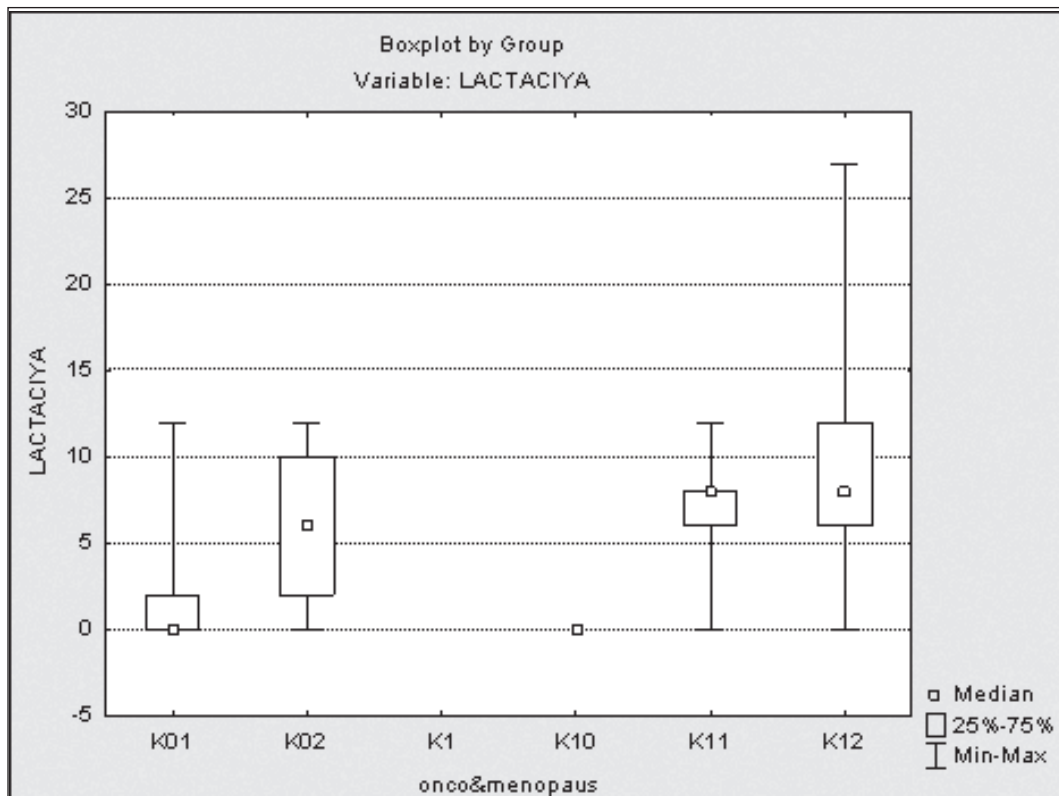
У дослідженні, що описується, вік пацієнтів жіночої статі на початок лікування коливався в межах 23 ÷ 62 роки з медіаною 62 роки, а на момент встановлення діагнозу про появу *prolongatio morbi* або другого раку — в межах 29 ÷ 71 рік з медіаною 54 роки (рис. 3).



**Рис. 2.** Коробчасті графіки залежності між сумарною тривалістю лактаційного періоду за репродуктивний проміжок життя та видом віддалених наслідків лікування хворих на РЦЗ (1 — ОН, 0 — ТН)



**Рис. 3.** Гістограма розподілу за віком пацієнтів жіночої статі, хворих на рак щитоподібної залози, які мали *prolongatio morbi* або другі раки на віддалених термінах після лікування



**Рис. 4.** Коробчасті графіки залежності між сумарною тривалістю лактаційного періоду за репродуктивний проміжок життя та видом віддалених наслідків лікування хворих на РЩЗ у поєднанні з менструальним статусом на початок лікування: K01 — ТН + пременопауза; K02 — ТН + перименопауза; K1, K10 — ОН + статус не описано; K11 — ОН + пременопауза; K12 — ОН + перименопауза

2-Way Summary Table: Observed Frequencies (ITOGTABL1)					
Marked cells have counts > 10					
Include condition: v15=1					
mts&rec1	STADI 1	STADI 2	STADI 3	STADI 4	Row Totals
0	10	21	4	4	39
Row %	25,64%	53,85%	10,26%	10,26%	
1	4	6	9	2	21
Row %	19,05%	28,57%	42,86%	9,52%	
2	2	5	0	0	7
Row %	28,57%	71,43%	0,00%	0,00%	
Totals	16	32	13	6	67

Рис. 5. Аналіз сполученості між стадією процесу та наслідками лікування (протокол середовища STATISTICA v.13)

Для більш глибокого висвітлення ситуації вибірку було розшаровано за комплексною ознакою — вид наслідків (ОН або ТН) та менструальним статусом на початок лікування (пременопауза та перименопауза) (рис. 4).

Як можна бачити, на фоні цілком очевидної залежності загальної тривалості лактації від менструального статусу (у жінок позарепродуктивного, перименопаузального статусу (тобто у більш літніх жінок) тривалість є більш високою — групи K01 і K02 та групи K11 і K12) чітко візуалізується тенденція перевищення загальної тривалості лактації у пацієток з ОН над пацієтками з ТН.

Таким чином, залежність є не зовсім зрозумілою та потребує подальшого вивчення на розширеній вибірці.

Другим етапом розвідувального аналізу було дослідження статистично значущих відмінностей у розрізі

«неонкологічні наслідки», «*prolongatio morbi*» (рецидиви або віддалені за часом метастази) та «другі раки» (інша нозологічна форма).

За таких умов, статистично значущою виявилася залежність між видом віддалених наслідків та стадією процесу (рис. 5).

Аналіз сполученості демонструє, що у більшості пацієнтів (21 особа, 61,5%), які у майбутньому мали *prolongatio morbi*, було діагностовано III стадію процесу. У двох інших групах — II: у групі без онкологічних наслідків — 21 особа (58,85%), у групі з другими пухлинами — 5 пацієнтів (71,43%).

Високу статистичну значущість має залежність віддалених наслідків від групи крові пацієнтів (КХП,  $p = 0,017$ ). Відповідні дані наведено у табл. 2.

Таблиця 2

**Розподіл частоти зустрічальності різних груп крові у пацієнтів без онкологічних наслідків, з *prolongatio morbi* та пацієнтів із другими раками у порівнянні з даними про середньоєвропейську популяцію та популяцію мешканців України**

Популяція або група	Група крові, % (абс.)			
	0 (1)	2	3	4
Середньоєвропейська	34,00%	38,00%	21,00%	8%
Мешканці України	43,00%	42,00%	11,00%	4%
Група без ОН	23% (9)	65% (26)	10% (4)	3% (1)
Група з <i>prolongatio morbi</i>	38% (8)	38% (8)	5% (1)	19% (4)
Група з другими раками	57% (4)	14% (1)	29% (2)	0%

Результати, наведені у табл. 2, демонструють, що більшість пацієнтів без ОН мали 2-гу групу крові, у пацієнтів з *prolongatio morbi* найбільш численними були 1-ша та 2-га групи крові, у пацієнтів з другими раками — перша група. Таким чином, гірші результати щодо віддалених наслідків мали пацієнти з першою групою крові. Одержані результати потребують уточнення на вибірці більшого обсягу.

## ВИСНОВКИ

На підставі розвідувального аналізу катamnестичних даних 68 пацієнтів, хворих на РЦЗ, серед яких 33 особи на віддалених термінах мали онкологічні наслідки, а 35 — лише терапевтичні, було виявлено низку статистично значущих залежностей, а саме:

1. Хворі, які на віддалених термінах після лікування мають ОН, характеризуються більш високим ІМТ за Кетле.

2. У пацієнтів жіночої статі з ОН загальний термін лактації за час репродуктивного періоду у 2,3 разу перевищує аналогічні показники жінок групи з ТН.

3. Продовження захворювання на віддалених термінах є більш характерним для пацієнтів з III стадією процесу, а відсутність ОН та поява других раків — для пацієнтів з II стадією процесу.

4. Гірші результати на віддалених термінах після спеціального лікування мали пацієнти з першою групою крові.

У зв'язку з обмеженим обсягом вибірки одержані результати мають статус гіпотез і потребують уточнення на більш численній вибірці.

**СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ**

1. *Informatics Association health policy conference. Journal of the American Medical Informatics Association: JAMIA.* — Vol. 17, N 2. — P. 115–123. <http://doi.org/10.1136/jamia.2009.001370>.
2. *Breastfeeding and thyroid cancer risk in women: a dose-response meta-analysis of epidemiological studies / X. Yi, J. Zhu, X. Zhu et al. // Clinical nutrition (Edinburgh, Scotland).* — 2016. — Vol. 35, N 5. — P. 1039–1046.
3. *Management of metastatic thyroid cancer in pregnancy: risk and uncertainty [Електронний ресурс] / C. W. Rowe, K. Murray, A. Woods et al. // Endocrinology, Diabetes & Metabolism Case Reports.* — 2016. — Режим доступу: <https://www.edmcasereports.com/articles/endocrinology-diabetes-and-metabolism-case-reports/10.1530/EDM-16-0071>. — Назва з екрану.

Стаття надійшла до редакції 22.02.2018.

Е. Б. РАДЗИШЕВСКАЯ<sup>1,2</sup>

<sup>1</sup> ГУ «Институт медицинской радиологии им. С. П. Григорьева НАМН Украины», Харьков

<sup>2</sup> Харьковский национальный медицинский университет

**РЕЗУЛЬТАТЫ РАЗВЕДЫВАТЕЛЬНОГО АНАЛИЗА ОТДАЛЕННЫХ ПОСЛЕДСТВИЙ ЛЕЧЕНИЯ БОЛЬНЫХ РАКОМ ЩИТОВИДНОЙ ЖЕЛЕЗЫ**

**Цель работы.** Проведение разведывательного анализа базы данных течения и послелечебного скрининга больных раком щитовидной железы (РЩЖ) с целью выявления статистически значимых зависимостей между данными анамнеза и течения заболевания и фактом возникновения отдаленных онкологических последствий лечения.

**Материалы и методы.** Проанализированы 215 бумажных историй болезни и внесена в электронную базу данных информация о течении заболевания 68 пациентов, отобранных по критериям или появления любого злокачественного заболевания через 3 и более лет после проведенного лечения, или появления отдаленных терапевтических осложнений. На момент начала обработки данных онкологические последствия были у 33 пациентов, терапевтические — у 35. Статистическую обработку проводили методами непараметрической статистики с помощью пакета программ Statistica Basic Academic 13 for Windows.

**Результаты.** Больные РЩЖ, у которых возникли отдаленные онкологические последствия, характеризуются более высоким индексом массы тела; у пациентов женского пола общий срок лактации за время репродуктивного периода в 2,3 раза превышает аналогичные показатели у женщин без онкологических последствий; *prolongatio morbi* более характерна для пациентов с III стадией процесса, а отсутствие онкологических последствий и появление вторых раков — для пациентов со II стадией процесса; худшие результаты на отдаленных сроках после специального лечения имели пациенты с первой группой крови.

**Выводы.** Проведенный разведывательный анализ катамнестических данных позволил выявить ряд неочевидных статистически значимых зависимостей. Расширение объема рабочего массива электронных данных позволит получить дополнительную информацию, имеющую клиническую и научную ценность.

**Ключевые слова:** медицинская информатика, рак щитовидной железы, отдаленные результаты лечения, статистический анализ.

Y. RADZISHEVSKA<sup>1,2</sup>

<sup>1</sup> SI «Grigoriev Institute for Medical Radiology of National Academy of Medical Sciences of Ukraine», Kharkiv

<sup>2</sup> Kharkiv National Medical University

**INTELLIGENCE ANALYSIS OUTCOMES OF REMOTE EFFECTS OF TREATMENT OF PATIENTS WITH THYROID CANCER**

**Purpose.** To carry out an intelligence analysis of the database containing information about the course and post-treatment screening of patients with thyroid cancer (TC) in order to reveal statistically significant relationships between past medical history data, disease course and the fact of appearing remote oncological effects after treatment.

**Materials and methods.** The study was focused on the analysis of 215 paper medical records. The information concerning disease course of 68 patients, who were selected in accordance with the criteria or due to appearing of any kind of malignancy 3 and more years after provided treatment or in case of remote therapeutic complications, was added to the electronic database. At the time of data processing, oncological consequences were revealed in 33 patients, therapeutic ones - in 35 patients. Statistical processing was carried out using nonparametric statistics by means of Statistica Basic Academic 13 for Windows software package.

**Outcomes.** The patients with TC, who have had remote oncological effects after the treatment, are characterized by a higher body mass index; in female patients, the overall lactation term in the reproductive period is 2.3 times higher than that one in women without remote oncological effects; *prolongatio morbi* is more typical for patients with stage 3 of the process, and the absence of oncological effects and appearing of second cancers - for patients with stage 2 of the process; the patients with the first group of blood had worse findings in remote terms after special treatment.

**Conclusions.** The intelligence analysis of catamnestic data have made it possible to reveal a lot of statistically significant dependencies that were not obvious without an adequate mathematical processing. Expanding the volume of the electronic array of data will allow us to obtain additional information of clinical and scientific value.

**Keywords:** medical informatics, thyroid gland cancer, treatment remote effects, statistical analysis.

**Контактна інформація:**

Радзішевська Євгенія Борисівна

канд. фіз.-мат. наук, доцент кафедри медичної та біологічної фізики і медичної інформатики ХНМУ; керівник групи медичної інформатики ДУ «ІМР НАМН України»

тел.: +38 (067) 799-36-63

e-mail: radzishevaska@ukr.net