

УДК 311.21:[614.1:616-072/-078]

## СТАТИСТИЧНИЙ АНАЛІЗ ЗВ'ЯЗКУ МІЖ СТАТТЮ ПАЦІЄНТІВ ТА ДІАГНОСТИКОЮ ГЕПАТИТУ С

Черненко Д.І., аспірант

*Харківський національний університет ім. В.Н. Каразіна*

У статті розглянуто методику та викладено результати статистичного аналізу зв'язку між статтю пацієнтів та наявністю антитіл до гепатиту С за різними віковими групами на основі даних, отриманих в результаті вибіркового обстеження пацієнтів діагностичних лабораторій м. Харкова з метою виявлення груп населення, які мають найбільший ризик інфікування.

**Ключові слова:** захворюваність на гепатит С, таблиця спряженості, тіснота зв'язку, якісні ознаки.

Вірусні гепатити на сьогодні становлять надзвичайно важливу проблему для охорони здоров'я. Питанню поширеності вірусного гепатиту різної етіології серед населення як в Україні, так і в інших країнах світу, приділяють все більше уваги.

Вчені виділяють п'ять основних типів вірусного гепатиту, що визначаються літерами А, В, С, D і Е. Гепатит С входить до числа найпоширеніших вірусів, що інфікують печінку, призводять до розвитку хронічної хвороби у сотень мільйонів людей і в цілому є найпоширенішою причиною цирозу і раку печінки [8].

За даними інформаційного бюлетеня Всесвітньої організації охорони здоров'я [6], у 2013 році близько 150 мільйонів людей були хронічно інфіковані вірусом гепатиту С, і щорічно більше 350 000 чоловік вмирають від пов'язаних з гепатитом С хвороб печінки. Інкубаційний період гепатиту С триває від 2 тижнів до 6 місяців. Після інфікування приблизно 80% людей не відчувають будь-яких симптомів, тому гепатит С часто діагностується випадково. У 25% пацієнтів з раком печінки основоположною причиною є гепатит С [6, 7].

Раннє виявлення і лікування випадків захворювання на гепатит С здійснюється у рамках програм раннього діагностування і скринінгу. Програми раннього діагностування особливо важливі в умовах обмежених ресурсів, де у більшості пацієнтів хвороба діагностується і лікується на дуже пізніх стадіях. Програми скринінгу, що ґрунтуються на застосуванні тестів серед здорового населення для виявлення ознак захворювання, дозволяють швидко направляти людей з такими ознаками на діагностування і лікування [8].

В Україні з 2001 р. спостерігалась загальна тенденція до зниження захворюваності на вірусний гепатит, проте слід відмітити зростання захворюваності в 2010 р. майже вдвічі у порівнянні з 2009 р. (з 13,8 до 26,8 випадків на 100 тис. населення) (рис.1) [2].

Статистичний аналіз захворюваності населення України на вірусний гепатит доцільно провести також у регіональному розрізі. Ранжування регіонів України за кількістю випадків захворюваності на вірусний гепатит на 100 тис. населення наведено на рис. 2 [5]. Так, у 2011 р. «лідерами» серед регіонів по кількості випадків захворюваності на 100 тис. населення були Миколаївська, Запорізька та Закарпатська області. Найменший рівень захворюваності на вірусний гепатит у 2011 р. можна спостерігати у Чернігівській області.

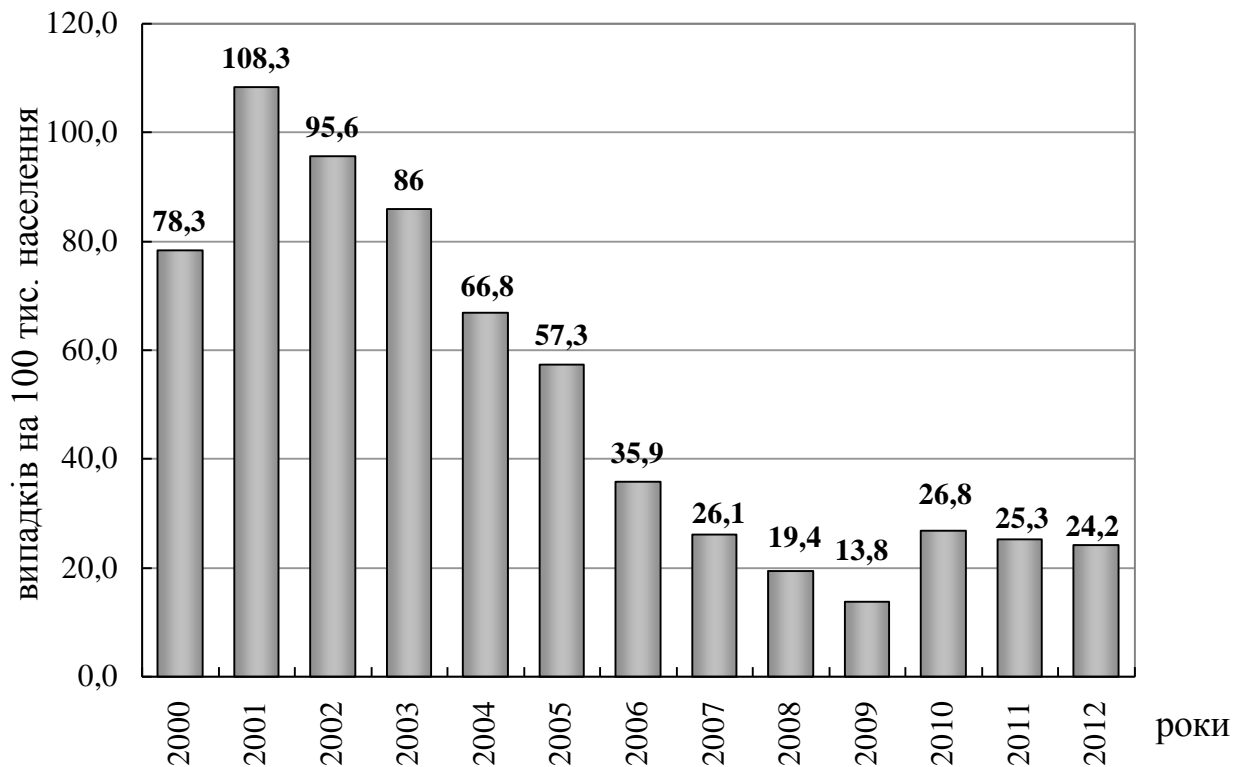


Рис. 1. Динаміка захворюваності на вірусний гепатит на території України в 2000–2012 рр.

Звернувши увагу на показники захворюваності на вірусний гепатит серед населення Харківської області за останні кілька років, можна спостерігати поступове збільшення кількості випадків на 100 тис. населення з 2009 р. (рис. 3) [3].

Також слід зауважити, що офіційні дані не завжди відображають справжній рівень поширеності вірусного гепатиту, оскільки вони базуються на реєстрації головним чином клінічно виражених форм цього захворювання.

З метою статистичного забезпечення попередження захворюваності на вірусний гепатит С, було проведено безповторне 50 % механічне вибіркове обстеження осіб, які звернулись до діагностичних лабораторій м. Харкова в 2012 та 2013 рр. для проведення аналізу на наявність антитіл до гепатиту С (Anti-HCV).

Антитіла до гепатиту С визначаються за допомогою імуноферментного аналізу (ІФА). Цей аналіз найчастіше використовується для скринінгу (первинного виявлення) вірусу гепатиту С. Присутність антитіл до вірусу гепатиту С вказує на те, що людина інфікована або була інфікована. Можливий хибно-позитивний аналіз на Anti-HCV, що зустрічається у пацієнтів з імунодефіцитом, наприклад, у хворих на ВІЛ першого типу, при нирковій недостатності. Тому для підтвердження діагнозу проводиться подальша діагностика з використанням лабораторних методів, наприклад рекомбінантний імуноблотинг (RIBA) на антитіла до гепатиту С і тестування на РНК вірусу гепатиту С.

Автор вважає за доцільне провести дослідження тісноти зв'язку між статтю пацієнтів, віком пацієнтів та результатом аналізу на наявність антитіл до гепатиту С (позитивний чи негативний) з метою виявлення груп населення, які мають найбільший ризик інфікування.

В ході дослідження було розроблено анонімну анкету, до уваги пацієнтів діагностичних лабораторій було запропоновано ряд запитань, які стосувались віку, статі та результату аналізу на наявність антитіл до гепатиту С (Anti-HCV).

Залежно від природи результативної й факторних ознак використовують різні статистичні методи вимірювання зв'язку: метод аналітичних групувань, дисперсійний аналіз, регресійний аналіз, рангову кореляцію та інші методи.

Для набору обраних ознак доцільно застосувати методи аналізу таблиць взаємної спряженості.

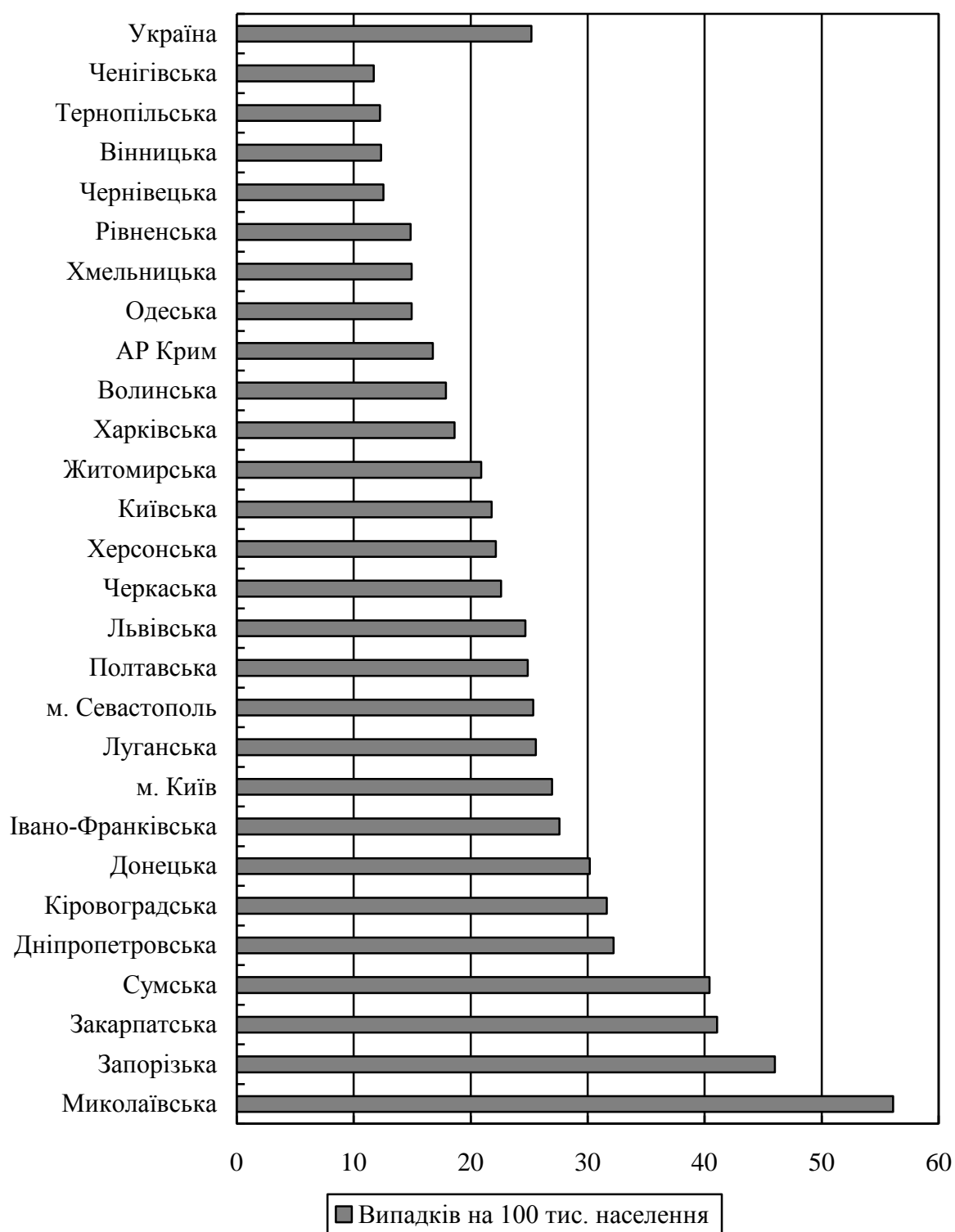


Рис. 2. Захворюваності на вірусний гепатит в регіонах України в 2011 р.

Теоретичні положення визначення тісноти зв'язку між ознаками, з використанням таблиць взаємної спряженості, викладено у працях таких учених, як: Г. Аптон [1], Л.І. Трошин, В.А. Балаш, О.С. Балаш [4] та інших.

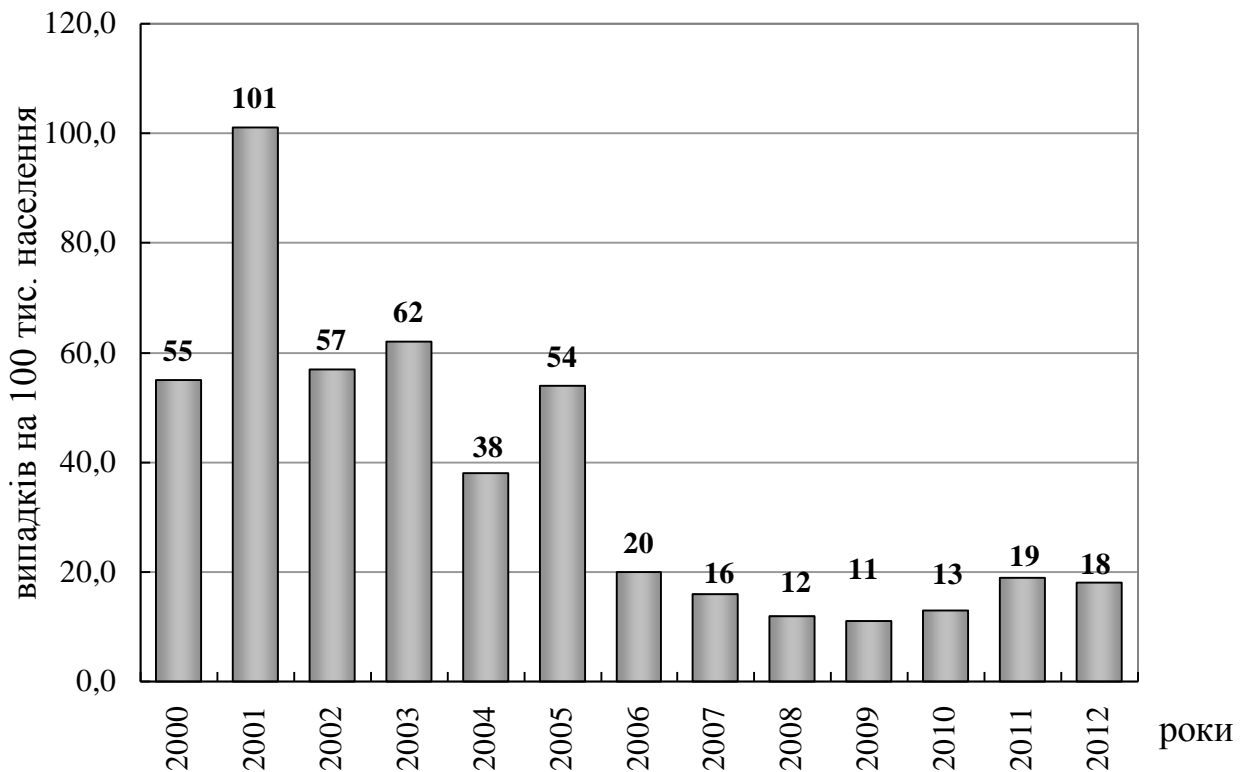


Рис. 3. Динаміка захворюваності на вірусний гепатит на території Харківської області в 2000–2012 рр.

Для проведення аналізу взаємозв'язку між результатом аналізу і статтю пацієнта за допомогою пакету прикладних програм «Statistica» версії 6.0 в ході реалізації модуля Crosstabulation Tables була сформована таблиця спряженості за наступним виділеними групами опитаних пацієнтів (рис. 4):

- 1) чоловіки з негативним результатом аналізу на наявність Anti-HCV ( $R < 1$ );
- 2) чоловіки з позитивним результатом аналізу на наявність Anti-HCV ( $R \geq 1$ );
- 3) жінки з негативним результатом аналізу на наявність Anti-HCV ( $R < 1$ );
- 4) жінки з позитивним результатом аналізу на наявність Anti-HCV ( $R \geq 1$ ).

При розрахунках було використано референтні значення результату, прийняті в діагностичних лабораторіях ( $R$ -коефіцієнт позитивності) для аналізу на наявність Anti-HCV:  $R \geq 1$  – позитивний,  $R < 1$  – негативний.

Слід зауважити, що до діагностичних лабораторій частіше звертаються жінки. Так і в даному випадку - частка жінок серед опитаних пацієнтів значно вища, ніж чоловіків (75,4% відсотків жінок проти 24,6% чоловіків). Цей факт необхідно взяти до уваги при формулюванні висновків і розробці рекомендацій щодо формування груп ризику, з метою попередження захворюваності на вірусний гепатит С.

Наглядне уявлення закономірностей розподілу елементів досліджуваної сукупності було отримано після застосування графічного методу в ході реалізації модуля Crosstabulation Tables прикладного пакету програм «Statistica» версії 6.0 (рис. 5).

Серед опитаних пацієнтів діагностичних лабораторій м. Харкова частка чоловіків з позитивним результатом аналізу на наявність Anti-HCV у 3,8 р. більша, ніж жінок.

STATISTICA - Workbook11-1\* - [2-входная итоговая: наблюдаемые частоты (Гепатит.sta)]

File Правка Вид Вставка Формат Статистика Графики Инструменты Данные Workbook Окно Помощь

Workbook11-1\* - 2-входная итоговая: наблюдаемые частоты (Гепатит.sta)

Результаты кросстабуляц

Итоговая таблица час  
Категориз.гистограмм  
График взаимод.: Ст  
3М Гистограмма:  
2-входная итоговая:  
Статистики: Статъ(2)  
Итоговая таблица час  
Итоговая таблица час  
Категориз.гистограмм  
График взаимод.: Вік  
3М Гистограмма:  
2-входная итоговая:

2-входная итоговая: наблюдаемые частоты (Гепатит.sta)  
Частоты выделенных ячеек > 10

Стать	2-входная итоговая: наблюдаемые частоты (Гепатит.sta)		Всего по стр.
	Результат анализу на наявність Anti-HCV R<1	Результат анализу на наявність Anti-HCV R>=1	
Жіноча	1462	46	1508
Чоловіча	435	57	492
Всього	1897	103	2000

2-входная итоговая: наблюдаемые частоты (Гепатит.sta) | Статистики: Статъ(2) x Результат анализу на наявність Anti-HC... | Итоговая таблица ч

Рис. 4. Таблица сопряженности признаков «Стать» и «Anti-HCV»

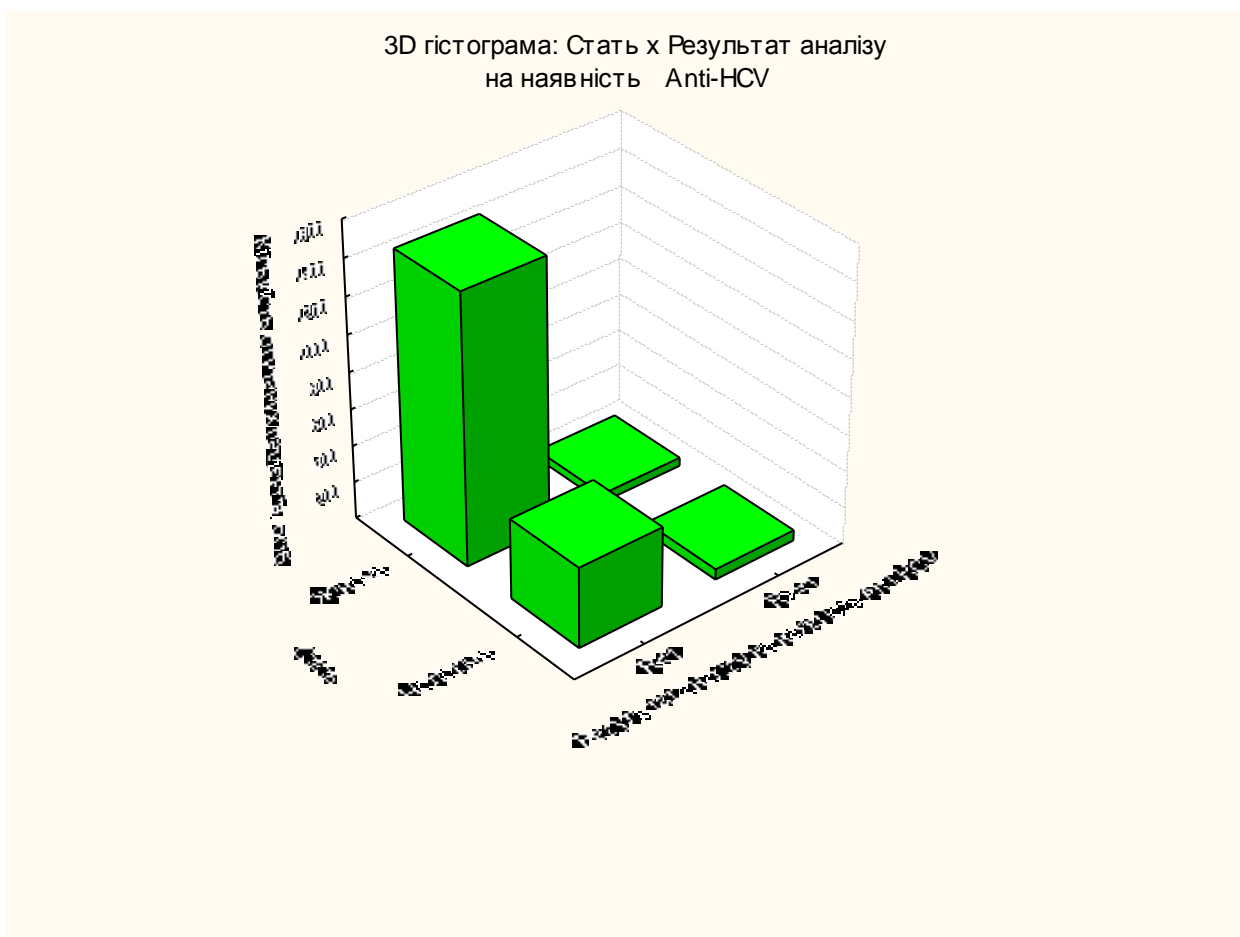


Рис. 5. 3D гистограмма: «Стать» x «Результат анализу на наявність Anti-HCV»

Щоб підтвердити припущення про наявність зв'язку між ознаками «Результат анализу» і «Стать» пацієнтів, необхідно скористатися критерієм Пірсона  $\chi^2$  [1, 4]. Критерій Пірсона  $\chi^2$  дозволяє судити про випадковість (або не випадковість) розподілу в таблицях взаємної

спряженості, а отже і про відсутність або наявності залежності між ознаками групування в таблиці.

Критерій Пірсона  $\chi^2$  засновано на зіставленні емпіричних і теоретичних частот [1, 4]:

$$\chi_{набл.}^2 = \sum_j^{m_2} \sum_i^{m_1} \frac{(n_{ij} - n_{ij}^T)^2}{n_{ij}^T}, \quad (1)$$

$$\text{або } \chi_{набл.}^2 = \sum_j^{m_2} \sum_i^{m_1} \frac{n_{ij}^2}{n_{ij}^T} - n, \quad (2)$$

де  $n_{ij}$  і  $n_{ij}^T$  – відповідно емпіричні й теоретичні частоти за групами;

$$n = \sum_j^{m_2} \sum_i^{m_1} n_{ij} - \text{загальне число одиниць сукупності.}$$

Якщо розраховане  $\chi_{набл.}^2 > \chi_{табл.}^2$ , то висунута нульова гіпотеза про випадковий розподіл відкидається, тобто розподіл не випадковий, та існує стохастичний зв'язок між ознаками.

Критерій Пірсона було розраховано за допомогою пакету прикладних програм «Statistica» версії 6.0 (рис. 6)

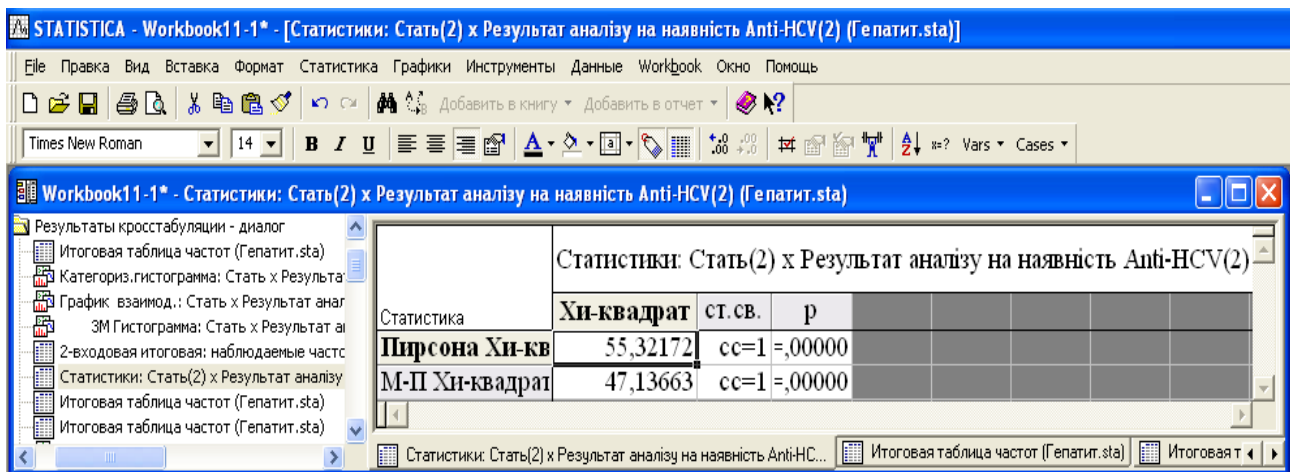


Рис. 6. Розрахунок критерію Пірсона (ознаки: «Стать» та «Результат аналізу на наявність Anti-HCV») за допомогою пакету прикладних програм «Statistica»

Розраховане (фактичне) значення  $\chi^2$  зіставимо з табличним (критичним) для заданого рівня значущості  $\alpha$  й числа ступенів свободи  $\nu = (m_1 - 1)(m_2 - 1)$ , де  $m_1$  й  $m_2$  – число груп за ознаками групування або, що те ж саме, число рядків і число граф у таблиці:

$$\chi_{кр.}^2 (\alpha = 0,05; \nu = 1) = 3,84.$$

Оскільки  $\chi_{емп.}^2 = 55,32 > \chi_{кр.}^2 = 3,84$ , то існує стохастична залежність між такими ознаками, як «Результат аналізу» і «Стать» пацієнта.

Дослідження наявності зв'язку між результатом аналізу та віком пацієнта за 4 віковими групами: 0–14 років, 15–35 років, 36–50 років, 51 рік і старше було проведено також із використанням критерію Пірсона  $\chi^2$ .

Критичне значення:  $\chi_{кр.}^2 (\alpha = 0,05; \nu = 3) = 7,815$ .

Порівняння розрахованого (емпіричного) значення  $\chi_{емп.}^2 = 12,36$  (рис. 7) з табличним (критичним) дозволило зробити висновок про відсутність стохастичного зв'язку між віком пацієнтів та результатом аналізу.

Статистика		Хи-квадрат	ст. св.	р
Пирсона Хи-квадрат		12,36048	сс=3	р=,00625
М-П Хи-квадрат		9,817990	сс=3	р=,02018

Рис. 7. Розрахунок критерію Пірсона (ознаки: «Вікова група» та «Результат аналізу на наявність Anti-HCV») за допомогою пакету прикладних програм «Statistica»

За кожною статтю було виділено вікові групи: 0–14 років, 15–35 років, 36–50 років, 51 рік і старше та розраховано критерій Пірсона за віковими групами пацієнтів (табл. 1).

Таблиця 1.

**Розрахунок критерію Пірсона для різних вікових груп опитаних пацієнтів за допомогою пакету прикладних програм «Statistica»**

Вікова група, років	Розрахунок критерію Пірсона															
0–14	<table border="1"> <thead> <tr> <th colspan="2">Статистика</th> <th>Хи-квадрат</th> <th>ст. св.</th> <th>р</th> </tr> </thead> <tbody> <tr> <td colspan="2">Пирсона Хи-квадрат</td> <td>,0798319</td> <td>сс=1</td> <td>р=,77753</td> </tr> <tr> <td colspan="2">М-П Хи-квадрат</td> <td>,0794766</td> <td>сс=1</td> <td>р=,77801</td> </tr> </tbody> </table>	Статистика		Хи-квадрат	ст. св.	р	Пирсона Хи-квадрат		,0798319	сс=1	р=,77753	М-П Хи-квадрат		,0794766	сс=1	р=,77801
Статистика		Хи-квадрат	ст. св.	р												
Пирсона Хи-квадрат		,0798319	сс=1	р=,77753												
М-П Хи-квадрат		,0794766	сс=1	р=,77801												
15–35	<table border="1"> <thead> <tr> <th colspan="2">Статистика</th> <th>Хи-квадрат</th> <th>ст. св.</th> <th>р</th> </tr> </thead> <tbody> <tr> <td colspan="2">Пирсона Хи-квадрат</td> <td>31,92831</td> <td>сс=1</td> <td>р=,00000</td> </tr> <tr> <td colspan="2">М-П Хи-квадрат</td> <td>26,75457</td> <td>сс=1</td> <td>р=,00000</td> </tr> </tbody> </table>	Статистика		Хи-квадрат	ст. св.	р	Пирсона Хи-квадрат		31,92831	сс=1	р=,00000	М-П Хи-квадрат		26,75457	сс=1	р=,00000
Статистика		Хи-квадрат	ст. св.	р												
Пирсона Хи-квадрат		31,92831	сс=1	р=,00000												
М-П Хи-квадрат		26,75457	сс=1	р=,00000												
36–50	<table border="1"> <thead> <tr> <th colspan="2">Статистика</th> <th>Хи-квадрат</th> <th>ст. св.</th> <th>р</th> </tr> </thead> <tbody> <tr> <td colspan="2">Пирсона Хи-квадрат</td> <td>15,15002</td> <td>сс=1</td> <td>р=,00010</td> </tr> <tr> <td colspan="2">М-П Хи-квадрат</td> <td>13,16197</td> <td>сс=1</td> <td>р=,00029</td> </tr> </tbody> </table>	Статистика		Хи-квадрат	ст. св.	р	Пирсона Хи-квадрат		15,15002	сс=1	р=,00010	М-П Хи-квадрат		13,16197	сс=1	р=,00029
Статистика		Хи-квадрат	ст. св.	р												
Пирсона Хи-квадрат		15,15002	сс=1	р=,00010												
М-П Хи-квадрат		13,16197	сс=1	р=,00029												

51 і старше	Статистики: Результат аналізу на наявність Anti-HCV(2) x Стать(2) (Гепатит. Подтабл. внутрі: Вікова група:51 і старше			
	Статистика	Хи-квадрат	ст. св.	p
	Пірсона Хи-квадрат	9,283397	сс=1	p=,00231
	М-П Хи-квадрат	8,028874	сс=1	p=,00460
Статистики: Результат аналізу на наявність Anti-HCV(2) x Стат...				Категориз. гистограма: Вікова група x Ре

Розраховані  $\chi_{емт.}^2 > \chi_{кр.}^2 = 3,84$  для вікових груп: 15–35 років, 36–50 років, 51 рік і старше, отже для цих груп з імовірністю 0,95 зв'язок між ознаками «Результат аналізу» і «Стать» пацієнта визнається істотним. Для групи пацієнтів, що не досягли 15 років, зв'язок між результатом ознаками «Результат аналізу» і «Стать» відсутній, оскільки  $\chi_{аіі}^2 = 0,07983 < \chi_{еđ}^2 = 3,84$ .

Слід зазначити, що величина  $\chi^2$  лише вказує на наявність зв'язку, а оцінювання сили зв'язку здійснюють за допомогою коефіцієнтів взаємної спряженості.

У нашому випадку має сенс виміряти (оцінити) тісноту зв'язку за допомогою таких коефіцієнтів спряженості [4]:

– коефіцієнту асоціації (коефіцієнт Юла):

$$K_{ac.} = \frac{n_{11} \cdot n_{22} - n_{12} \cdot n_{21}}{n_{11} \cdot n_{22} + n_{12} \cdot n_{21}}, \quad (3)$$

– коефіцієнту коллігації:

$$K_{колл.} = \frac{\sqrt{n_{11} \cdot n_{22}} - \sqrt{n_{12} \cdot n_{21}}}{\sqrt{n_{11} \cdot n_{22}} + \sqrt{n_{12} \cdot n_{21}}} \quad (4)$$

– коефіцієнту контингенції:

$$K_{конт.} = \frac{n_{11} \cdot n_{22} - n_{12} \cdot n_{21}}{\sqrt{(n_{11} + n_{12})(n_{21} + n_{22})(n_{11} + n_{21})(n_{12} + n_{22})}} = \frac{n_{11} \cdot n_{22} - n_{12} \cdot n_{21}}{\sqrt{n_{1*} \cdot n_{2*} \cdot n_{*1} \cdot n_{*2}}} \quad (5)$$

– коефіцієнту спряженості Пірсона:

$$K_{П} = \sqrt{\frac{\chi^2}{\chi^2 + n}} \quad (6)$$

– коефіцієнту Чупрова-Крамера:

$$K_{ч} = \sqrt{\frac{\chi^2}{n}} \quad (7)$$

–  $\tau$ -коефіцієнту, що запропонований Гудменом і Краскалом :

$$\tau = K_{краск.} = \frac{(n_{11} \cdot n_{22} - n_{12} \cdot n_{21})^2}{n_{1*} \cdot n_{2*} \cdot n_{*1} \cdot n_{*2}} = K_{конт.}^2 \quad (9)$$

Розраховані значення коефіцієнтів свідчать про помірний зв'язок між результатом аналізу та статтю пацієнта для групи 51 року і старше ( $K_{ac.} = 0,389$ ) та тісний зв'язок для вікових груп 15–35 років та 36–50 років (табл. 2).

Результати проведених розрахунків наведено в табл. 2.



Таблиця 2

## Значення коефіцієнтів спряженості для вікових груп 15–35 років, 36–50 років

Вікова група	$K$	$K$	$K$	$\tau$	$K_{II}$	$K_{\varphi}$
	<i>ас.</i>	<i>колл.</i>	<i>конт.</i>			
15–35	0,5 97	0,33 1	0,158	0,0 25	0,1 56	0,1 58
36–50	0,6 45	0,36 5	0,193	0,0 37	0,1 89	0,1 92

В результаті проведеного аналізу можна зробити висновок, в групу найбільшого ризику входять чоловіки віком 15–50 років.

Результат проведеного статистичного аналізу дозволив зробити висновок про необхідність проведення регулярних медичних обстежень з використанням лабораторних методів, особливо серед чоловіків у віковій групі 15–35 років та 36–50 років, з метою попередження захворюваності на вірусний гепатит С (НСV). Систематичне обстеження населення (особливо осіб, що входять до груп ризику) дозволяє здійснити ранню діагностику вірусного гепатиту С і запобігти розвитку ускладнень. Раннє діагностування може запобігти проблеми зі здоров'ям, які можуть виникати в результаті інфекції, а також передачу інфекції членам сім'ї та іншим людям, що мають тісні контакти з інфікованою людиною.

**Література:**

1. Аптон Г. Анализ таблиц сопряженности; пер. с англ./ Г. Аптон. – М.: Финансы и статистика, 1982. – 143 с.
2. Статистичний щорічник України за 2012 рік/ Державна служба статистики України; за ред. О.Г. Осауленка. – К., 2013. – 552 с.
3. Статистичний щорічник Харківська область у 2012 році/ Державна служба статистики України. Головне управління статистики у Харківській області; за ред. О.Г. Мамонтової. – Х.: Головне управління статистики у Харківській області, 2013. – 555 с.
4. Трошин Л.И. Статистический анализ нечисловой информации/ Л.И. Трошин, В.А. Балаш, О.С. Балаш. – Московский государственный университет экономики, статистики и информатики. – М., 2003. – 67 с.
5. Українська база медико-статистичної інформації [Електронний ресурс]// Центр медичної статистики Міністерства охорони здоров'я України. – Режим доступу: <http://medstat.gov.ua/ukr/normdoc.html>.
6. Fact sheet N°164. Hepatitis C [Electronic resource]// The World Health Organization Official Website. – Mode of access: <http://www.who.int/mediacentre/factsheets/fs164>.
7. Mauss S. Short guide to hepatitis C. – Flying Publisher [Electronic resource]/ S. Mauss, T. Berg, J. Rockstroh et al. 2013. – 178 p. – Mode of access: [http://pdf.flyingpublisher.com/HepatitisC\\_Guide\\_2013.pdf](http://pdf.flyingpublisher.com/HepatitisC_Guide_2013.pdf).
8. What is hepatitis? [Electronic resource]// The World Health Organization Official Website. – Mode of access: <http://www.who.int/features/qa/76/en>.

СТАТИСТИЧЕСКИЙ АНАЛИЗ СВЯЗИ МЕЖДУ ПОЛОМ ПАЦИЕНТОВ  
И ДИАГНОСТИКОЙ ГЕПАТИТА С

Черненко Д.И., аспирант

*Харьковский национальный университет имени В.Н. Каразина*

В статье рассмотрена методика и представлены результаты статистического анализа связи между полом пациентов и наличием антител к гепатиту С (положительный или отрицательный результат) по различным возрастным группам на основе данных, полученных

в результате выборочного обследования пациентов диагностических лабораторий г. Харькова, с целью выявления групп населения, которые имеют наибольший риск инфицирования.

**Ключевые слова:** заболеваемость гепатитом С, таблица сопряженности, теснота связи, качественные признаки.

STATISTICAL ANALYSIS OF THE RELATIONSHIP BETWEEN SEX  
OF PATIENTS AND DIAGNOSIS OF HEPATITIS C

Chernenko D.I., post-graduate student  
*V.N. Karazin Kharkiv National University*

The article reviews the methodology and presents the results of the statistical analysis of the relationship between sex of patients and the presence of antibodies to HCV (positive or negative) in different age groups on the basis of data derived from a sample survey of patients diagnostic laboratories Kharkov to identify groups people who have the greatest risk of infection.

**Keywords:** incidence of hepatitis C, crosstabulation table, distress communication, quality attributes.

Надійшла до редколегії 17.02.2014 р.

Науковий керівник: Чала Т.Г., к. е. н., доцент