

УДК: 616.891:616.85]-07-037-057.875:519.226

## МОДЕЛЬ МУЛЬТИНОМІАЛЬНОЇ ЛОГІСТИЧНОЇ РЕГРЕСІЇ ТА ПОБУДОВА ДЕРЕВА РІШЕНЬ ДЛЯ ВИЗНАЧЕННЯ ФАКТОРІВ РИЗИКУ РОЗВИТКУ НЕПСИХОТИЧНИХ ПСИХІЧНИХ РОЗЛАДІВ У СТУДЕНТІВ ВИЩИХ НАВЧАЛЬНИХ ЗАКЛАДІВ

О. С. Юрценюк

Юрценюк  
Ольга Сидорівна

Вищий державний навчальний заклад України «Буковинський державний медичний університет», пл. Театральна, 2, м. Чернівці, Україна, 58002  
yurtsenyuk.olga@bsmu.edu.ua  
ORCID ID: 0000-0002-1450-1530

**Актуальність.** Важливість сучасних досліджень, пов'язаних з неспсихотичними психічними розладами у студентів вищих навчальних закладів, визначається необхідністю побудови математичної моделі за допомогою якої можна виділити фактори, які достовірно впливають на стан психічного здоров'я молодих людей.

**Мета дослідження.** Побудова математичної моделі впливу різних чинників на рівень психічного здоров'я студентів з неспсихотичними психічними розладами.

Матеріали і методи. Нами, протягом 2015-2017 рр., з дотриманням принципів біоетики та деонтології проведено суцільне комплексне обстеження 1235 студентів. Застосовані методи: клінічний, клініко-психопатологічний, клініко-епідеміологічний, клініко-анамнестичний, експериментально-психологічний та статистичний. Вибірка не мала суттєвих відмінностей за статевим і віковим складом, місцем мешкання, формою навчання. Обстеження проводили в міжсесійний період.

**Результати.** В результаті проведення дисперсійного аналізу ми отримали статистично значимі результати (ANOVA-тесту) між категорійною змінною «група І» та змінними що позначають всі шкали СБОУ: 1. іпохондрії (Hs),  $p < 2,2 \cdot 10^{-16}$ ; 2. депресії (D),  $p < 2 \cdot 10^{-16}$ ; 3. істерії (Hy),  $p < 2 \cdot 10^{-16}$ ; 4. психопатії (Pd),  $p < 2,89 \cdot 10^{-14}$ ; 6. паранояльності (Pa),  $p < 2 \cdot 10^{-16}$ ; 7. психастенії (Pt),  $p < 2 \cdot 10^{-16}$ ; 8. шизоїдності (Sc),  $p < 7,04 \cdot 10^{-16}$ ; 9. гіпотонії [манії] (Ma),  $p < 1,95 \cdot 10^{-10}$ ; шкалою брехні (L),  $p < 2,74 \cdot 10^{-11}$ ; шкалою вірогідності (F),  $p < 4,66 \cdot 10^{-8}$ ; шкалою корекції (K),  $p < 3,6 \cdot 10^{-13}$ .

Встановлено, що для моделі мультиноміальної логістичної регресії, метою якої є оцінка ймовірності належності респондента до однієї з груп (значення змінної «group.I»), значимими предикторами є сімейний стан, наявність дітей, стосунки в сім'ї, нездатність пристосуватися до середовища, якість сну, низька продуктивність, підвищена втомлюваність, тютюнопаління, оцінка власного здоров'я, необхідність консультації лікаря, рівень тривоги, депресії, нейротизму.

Коефіцієнти моделі визначають величину зміни логарифму шансу перебування у тій чи іншій групі («комплексній», «контрольній», «стандартній») проти («порівняльної») групи здорових, якщо значення предиктора збільшиться на одиницю. Точність передбачення на загальній вибірці становила 86,35%, а на тестовій вибірці (основна група, що включала хворих на НПР) - 88,52%.

**Висновки.** Таким чином, проведені  $\chi^2$  тест та ANOVA дали змогу визначити ті змінні, які мають взаємозв'язок за змінною «group.I». Такими змінними є: вік, спеціальність і т.д. (загалом їх 92). Модель дерева рішень, побудована на тих же предикторах, що і модель мультиноміальної регресії, (схематичний рисунок) дає точність прогнозу 89,95% на навчальній вибірці і 87,1% на тестовій вибірці.

**Ключові слова:** студенти, неспсихотичні психічні розлади, множинна лінійна регресія, мультиноміальна логістична регресія, дерево рішень.

**Актуальність.** Відповідно до критеріїв ВООЗ психічне здоров'я визначається як відсутність виражених психічних розладів, наявність певного резерву сил людини, завдяки якому вона може подолати несподівані стреси, труднощі, що виникають у виняткових обставинах, а також стан рівноваги між людиною і навколишнім світом, гармонії між нею і суспільством, співіснування уявлень окремої людини з уявленнями інших людей про «об'єктивну реальність» [1]. Таким чином, діагностика психічного здоров'я повинна проводитися не тільки за принципом виключення (тобто констатації відсутності психічного захворювання), але і враховувати наявність нормального психічного розвитку і сприятливого гармонійного функціонального стану вищих відділів ЦНС [2; 3]. Зрозуміло, що критерії «нормальності» психічного розвитку, певною мірою відносні і не можуть бути зведені лише до поняття статистичної норми, а повинні враховувати все різноманіття психічних проявів особистості, наявність функціональних резервів, визначення яких вимагає здійснення комплексного, динамічного підходу до діагностики психічного здоров'я та його порушень в межах біопсихосоціальної парадигми, що відображає єдність біологічних, психічних і соціальних факторів в формуванні норми, так і у розвитку патологічних змін [4].

**Мета дослідження.** Побудова математичної моделі впливу різних чинників на рівень психічного здоров'я студентів з неспокійними психічними розладами.

**Матеріали і методи.** Нами, протягом 2015–2017 рр., з дотриманням принципів біоетики та деонтології проведено суцільне комплексне обстеження студентів I–V курсів Вищого державного навчального закладу України «Буковинський державний медичний університет» (БДМУ) та студентів I–IV курсів Чернівецького національного університету.

Усіх студентів (1235 осіб) було розподілено на на дві групи – основну (I-шу) та порівняльну (II-гу). До основної групи увійшли 317 (25,67%) студентів, у яких було діагностовано неспокійні психічні розлади (НПР). Для верифікації клінічних особливостей НПР, в якості II порівняльної групи було обстежено 918 (74,33%) студентів без НПР, практично здорових.

Середній вік обстежених  $20,15 \pm 0,05$  років. В гендерному розподілі серед усіх обстежених було 365 чоловіків та 870 жінок, відповідно 29,55 % та 70,45 %. Вибірка не мала суттєвих відмінностей за статевим і віковим складом, місцем мешкання, формою навчання. Обстеження проводили в міжсесійний період. Застосовані методи: клінічний, клініко-психопатологічний, клініко-епідеміологічний, клініко-анамнестичний, експериментально-психологічний та статистичний методи.

## Обговорення результатів дослідження.

### 3.2. Модель мультиноміальної логістичної регресії та побудова дерева рішень для визначення предикторів розвитку, формування неспокійних психічних розладів у студентів вищих навчальних закладів

Для того, щоб перевірити наявність статистично значимого зв'язку між категорійною змінною «Група I» (N=1235) та іншими змінними (N=152), які були використані в розробленій нами «Уніфікованій карті обстеження студента» проведено тест  $\chi^2$  (хі-квадрат тест). Тут ми перевірили нульову гіпотезу:

$H_0$ : {змінні є незалежними (між двома змінними не існує асоціації)}.

при альтернативній гіпотезі

$H_1$ : {змінні не є незалежними (між змінними є зв'язок)}.

Ми отримали статистично значимі результати  $\chi^2$ -тесту, що підтверджували альтернативну гіпотезу, тобто наявність залежності між змінними – категорійною змінною «група I» та віком,  $p=0,0001$ ; спеціальністю,  $p=2,569 \cdot 10^{-8}$ ; місцем проживання,  $p=0,02$ ; сімейним статусом,  $p < 2,2 \cdot 10^{-16}$ ; наявністю дітей,  $p=5,509 \cdot 10^{-5}$ ; стосунків у сім'ї,  $p < 2,2 \cdot 10^{-16}$ ; відсутністю перспективи,  $p=0,0004$ ; нездатністю пристосуватися до середовища,  $p < 2,2 \cdot 10^{-16}$ ; негативним впливом засобів масової інформації,  $p < 4,112 \cdot 10^{-6}$ ; міжособистісними конфліктами в групі,  $p < 5,665 \cdot 10^{-5}$ ; складністю в засвоєнні нових норм і правил,  $p=1,518 \cdot 10^{-8}$ ; відчуттям ізоляції серед оточуючих,  $p < 2,939 \cdot 10^{-8}$ ; звертанням до ворожок, знахарів,  $p < 2,2 \cdot 10^{-16}$ ; проживанням на самоті,  $p=0,006$ ; конфліктами в сім'ї,  $p < 9,642 \cdot 10^{-6}$ ; проблемами з житлом,  $p=0,04$ ; конфліктами з друзями,  $p=9,658 \cdot 10^{-11}$ ; конфліктами в мікрогрупі,  $p=3,715 \cdot 10^{-9}$ ; проблемами в стосунках з дітьми,  $p=0,02$ ; проблеми в стосунках з батьками,  $p < 5,137 \cdot 10^{-5}$ ; вимушеними змінами в сімейному статусі (тимчасові розлучення),  $p=0,02$ ; релігійним фактором (міжконфесійними чварами),  $p=1,847 \cdot 10^{-10}$ ; сексуальними проблемами,  $p < 2,2 \cdot 10^{-16}$ ; конфліктами в навчанні,  $p=0,001$ ; ятрогеніями,  $p=0,05$ ; якістю сну,  $p < 1,414 \cdot 10^{-8}$ ; адаптації до навчання,  $p < 2,2 \cdot 10^{-16}$ ; результатами навчання,  $p=2,559 \cdot 10^{-8}$ ; підготовкою до занять в нічний час,  $p=0,03$ ; самопочуттям після навчання,  $p=1,144 \cdot 10^{-10}$ ; участю у суспільній роботі,  $p=0,03$ ; великого навчального навантаження,  $p=0,0004$ ; невмінням правильно розпорядитися обмеженими фінансами  $p=0,04$ ; конфліктами в групі,  $p=0,0007$ ; відчуттям безпорадності, неможливістю впоратися з проблемами,  $p=1,148 \cdot 10^{-6}$ ; підвищеним відволіканням, поганою концентрацією уваги,  $p=0,004$ ; дратівливістю, образливістю,  $p=2,152 \cdot 10^{-5}$ ;

поганим настроєм,  $p=5,447 \cdot 10^{-8}$ ; страхом, тривогою,  $p=4,147 \cdot 10^{-8}$ ; втратою впевненості, зниження самооцінки,  $p=1,048 \cdot 10^{-10}$ ; поганим сном,  $p=9,523 \cdot 10^{-6}$ ; порушенням соціальних контактів, проблемами в спілкуванні,  $p < 2,2 \cdot 10^{-16}$ ; прискореним серцебиттям, болями в серці,  $p=0,01$ ; утрудненим диханням,  $p=1,313 \cdot 10^{-5}$ ; проблемами з шлунково-кишковим трактом,  $p=0,05$ ; напругою або тремтінням м'язів,  $p=0,03$ ; низькою працездатністю, підвищеною стомлюваністю,  $p=0,0008$ ; рівня постійного стресу за останні три місяці навчання,  $p=0,0004$ ; кількістю вкурених цигарок за день,  $p=5,787 \cdot 10^{-7}$ ; оцінкою власного здоров'я,  $p < 2,2 \cdot 10^{-16}$ ; спеціальністю лікарів для консультації,  $p < 2,2 \cdot 10^{-16}$ .

В якості засобів подолання стресу: отримали достовірні результати  $\chi^2$  тесту, що підтверджували альтернативну гіпотезу, тобто наявність залежності між категорійною змінною «група І» та прийомом алкоголю,  $p=1,826 \cdot 10^{-6}$ ; тютюнопалінням,  $p=7,461 \cdot 10^{-14}$ ; вживанням наркотиків,  $p=6,444 \cdot 10^{-6}$ ; сексом,  $p=0,03$ ; спілкуванням з батьками,  $p=0,03$ .

Також, в результаті дослідження отримали статистично значимі результати  $\chi^2$  тесту, що підтверджували альтернативну гіпотезу, тобто наявність залежності між категорійною змінною «група І» та змінами, що характеризували як преморбідні фактори формування НПП – нещасливим дитинством,  $p=0,0005$ ; шкільними проблемами,  $p=0,0001$ ; емоційними проблемами,  $p=8,54 \cdot 10^{-5}$ ; поведінкові проблеми,  $p=0,006$ ; проблеми в стосунках з матір'ю,  $p=1,267 \cdot 10^{-6}$ ; проблеми в стосунках з бабусею, дідусем,  $p=4,748 \cdot 10^{-14}$ ; правопорушеннями,  $p=0,001$ ; сильними релігійними переконаннями,  $p=0,04$ ; проблемами зі здоров'ям,  $p=0,01$ ; вживанням алкоголю,  $p=0,0001$ ; відчуттям ізоляції серед однолітків,  $p=0,002$ ; суворими покараннями,  $p=2,46 \cdot 10^{-7}$ ; надмірною батьківською опікою,  $p=0,0002$ ; сексуальним насильством,  $p=0,0009$ ; емоційним ігноруванням в родині,  $p=3,638 \cdot 10^{-11}$ ; порушеним харчуванням,  $p=1,643 \cdot 10^{-6}$ .

Якщо одна із змінних є категорійною, а інша числовою (інтервальною) яка у групах за категорійною змінною є нормально розподіленою (нормальність перевіряється за допомогою тесту Шапіро-Уїлка), то для виявлення залежностей між змінними ми використовували ANOVA-тест, який перевіряє нульову гіпотезу  $H_0: \{ \text{усі групові середні рівні} \}$  при альтернативній гіпотезі  $H_1: \{ \text{не всі групові середні рівні} \}$ .

Дисперсійний аналіз є сукупністю статистичних методів, призначених для перевірки гіпотез про зв'язок між певною ознакою та досліджуваними факторами, які не мають кількісного опису, а також для встановлення ступеня впливу факторів та їх взаємодії. У спе-

ціальній літературі дисперсійний аналіз часто називають ANOVA (від англословної назви Analysis of Variations). Вперше цей метод було розроблено Р. Фішером в 1925 р. Факторами називають контрольовані чинники, що впливають на кінцевий результат. Рівнем фактора, або способом обробки, називають значення, що характеризують конкретний прояв цього фактора. Ці значення зазвичай подають у номінальній або порядковій шкалі вимірювань. Значення вимірюваної ознаки називають відгуком. Часто вихідні значення факторів вимірюють у кількісних або порядкових шкалах. Тоді постає проблема групування вихідних даних у ряди спостережень, що відповідають приблизно однаковим значенням фактора. Якщо кількість груп взяти надмірно великою, то кількість спостережень у них може виявитися недостатньою для отримання надійних результатів. Якщо її взяти надмірно малою, це може призвести до втрати суттєвих особливостей впливу досліджуваного фактора на систему. Вибір конкретного способу групування даних залежить від їх обсягу і характеру варіювання значень фактора. Кількість і розміри інтервалів при однофакторному аналізі найчастіше визначають за принципом рівних інтервалів або за принципом рівних частот. При багатфакторному аналізі застосовують три типи групування:

- групи з рівною кількістю спостережень;
- групи з різною кількістю спостережень;
- групи, кількості спостережень у яких відповідають певній пропорції

#### Тест Шапіро-Уїлка

Нульова гіпотеза  $H_0$  тесту Шапіро-Уїлка полягає в тому, що випадкова величина, вибірка  $x$  якої відома, розподілена за нормальним законом. Альтернативна гіпотеза  $H_1$  полягає в тому, закон розподілу не є нормальним.

Для виконання тесту Шапіро-Уїлки призначена функція `shapiro.test(x)`, приймаюча на вхід вибірку  $x$  об'єму не менше 3 і не більше 5000.

Функція повертає список з наступними компонентами:

- `statistic` - значення статистики тесту, яку прийнято позначати буквою  $W$ ;
- `p.value` - апроксимация  $p$ -value для отриманого значення статистики;
- `method` - рядок з назвою тесту;
- `data.name` – ім'я змінної, що містить вибірку, яка була передана функції `shapiro.test` як аргумент.

Слід зазначити, що незважаючи на те, що значення яке повертається є списком, його висновок на екрані обробляється особливим чином, дозволяючи більш компактно представити дані про виконаний тест.

В результаті проведення дисперсійного аналізу ми отримали статистично значимі результати (ANOVA-тесту)

між категорійною змінною «група І» та змінними що позначають всі шкали СБОУ: 1. іпохондрії (Hs),  $p < 2,2 \cdot 10^{-16}$ ; 2. депресії (D),  $p < 2 \cdot 10^{-16}$ ; 3. істерії (Hy),  $p < 2 \cdot 10^{-16}$ ; 4. психопатії (Pd),  $p < 2,89 \cdot 10^{-14}$ ; 6. паранояльності (Pa),  $p < 2 \cdot 10^{-16}$ ; 7. психастенії (Pt),  $p < 2 \cdot 10^{-16}$ ; 8. шизоїдності (Sc),  $p < 7,04 \cdot 10^{-16}$ ; 9. гіпотонії [манії] (Ma),  $p < 1,95 \cdot 10^{-10}$ ; шкалою брехні (L),  $p < 2,74 \cdot 10^{-11}$ ; шкалою вірогідності (F),  $p < 4,66 \cdot 10^{-8}$ ; шкалою корекції (K),  $p < 3,6 \cdot 10^{-13}$ .

Також, ми отримали статистично значимі результати (ANOVA-тесту) між категорійною змінною «група І» та абсолютними значеннями рівня ситуативної тривожності,  $p < 2 \cdot 10^{-16}$ ; рівнем ситуативної тривожності,  $p < 2 \cdot 10^{-16}$ ; абсолютними значеннями рівня особистісної тривоги,  $p < 2 \cdot 10^{-16}$ ; ROT  $p < 2 \cdot 10^{-16}$ ; абсолютними значеннями алекситимії,  $p < 2 \cdot 10^{-16}$ ; рівнем алекситимії,  $p < 2 \cdot 10^{-16}$ ; абсолютними значеннями депресії,  $p < 2 \cdot 10^{-16}$ ; рівнем депресії,  $p < 2 \cdot 10^{-16}$ ; абсолютними значеннями нейротизму,  $p < 2 \cdot 10^{-16}$ ; рівнем нейротизму,  $p < 2 \cdot 10^{-16}$ .

З метою прогнозування значень «Групи І» ми побудували модель мультиноміальної логістичної регресії. Мультиноміальна логістична регресія використовується для моделювання змінних номінальних результатів, в яких коефіцієнти журналу результатів моделюються як лінійна комбінація змінних предиктора.

Встановлено, що для моделі мультиноміальної логістичної регресії, метою якої є оцінка ймовірності належності респондента до однієї з груп (значення змінної «group.I»), значимими предикторами є сімейний стан, наявність дітей, стосунки в сім'ї, нездатність пристосуватися до середовища, якість сну, низька продуктивність, підвищена втомлюваність, тютюнопаління, оцінка власного здоров'я, необхідність консультації лікаря, рівень тривоги, депресії, нейротизму.

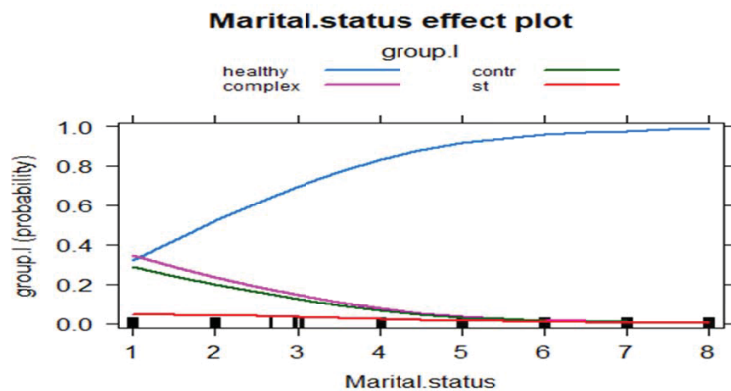
Коефіцієнти моделі визначають величину зміни логарифму шансу перебування у тій чи іншій групі («комплексній», «контрольній», «стандартній») проти («порівняльній») групи здорових, якщо значення предиктора збільшиться на одиницю.

Точність передбачення на загальній вибірці становила 86,35%, а на тестовій вибірці (основна група, що включала хворих на НПП) – 88,52%.

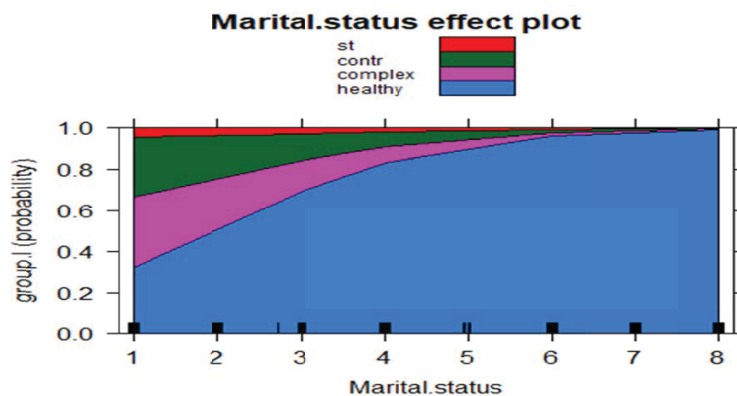
Отримано графіки та гістограми прогнозів з накопиченням за кожним предиктором прогнозних ймовірностей для кожного значення «group.I» (категорійної змінної «Група І» (N=1235)):

Примітка: графік та гістограма взяті із створеної нами комп'ютерної програми, тому переклад надаємо в примітці. Marital status – сімейний стан. Тут і далі\*. Group

```
plot(Effect("Marital.status",rd),multiline=T) #shows multiple lines of predictions in the same plot
```



```
plot(Effect("Marital.status",rd),style="stacked") # shows a stacked vertical bar chart of predictions
```

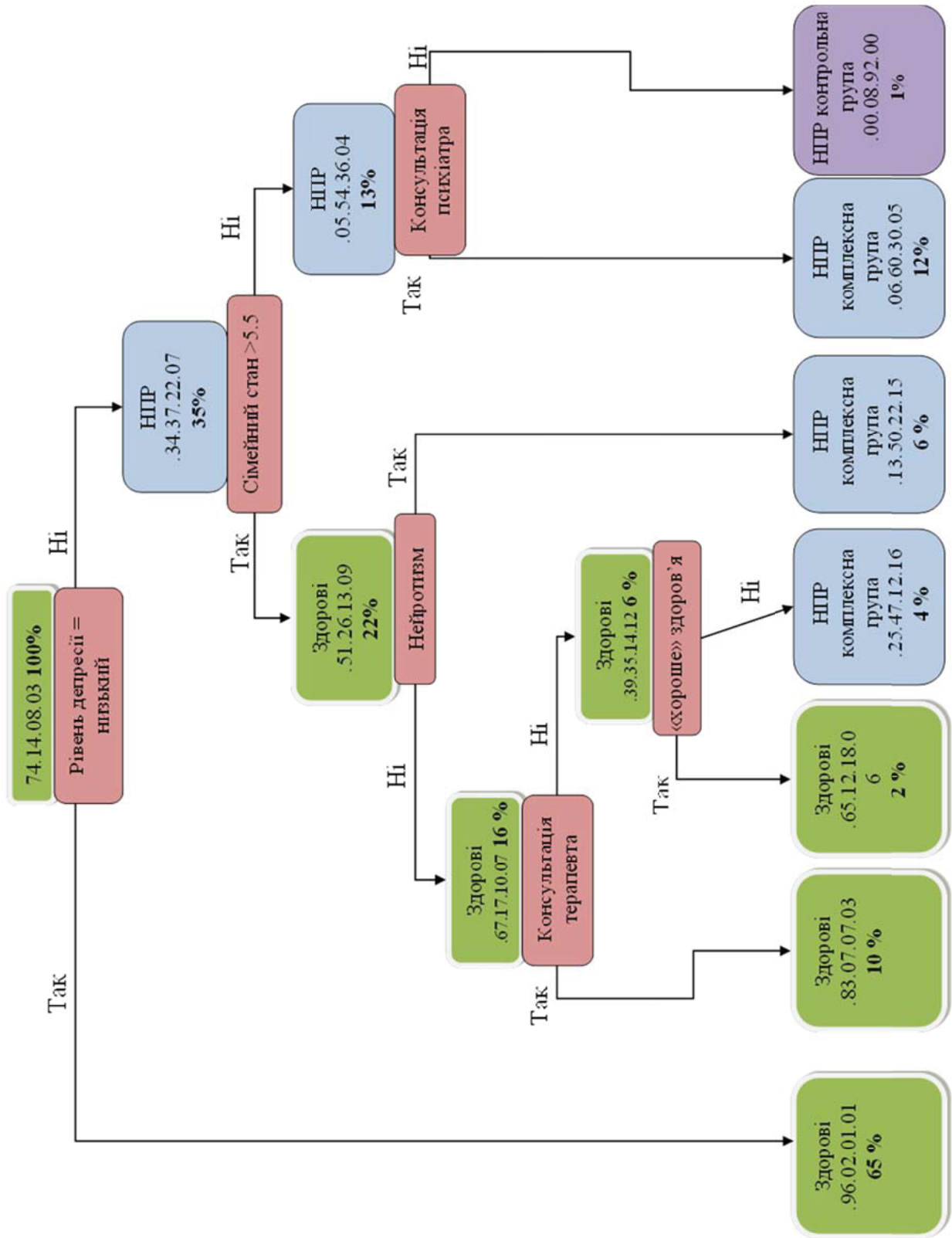


I – Група I. healthy – здорові, complex – комплексна, contr – контрольна, st – стандартна групи. Probability – ймовірність. 1 шлюб; 2 стосунки на відстані; 3 розлучені; 4 віртуальні стосунки; 5 громадянський шлюб; 6 самотні; 7 серйозні дошлюбні стосунки; 8 свій варіант.

Модель дерева рішень, побудована на тих же предикторах, що і модель мультиноміальної регресії, (схематичний рисунок) дає точність прогнозу 89,95% на навчальній вибірці і 87,1% на тестовій вибірці.

**Висновки.** Таким чином, проведені  $\chi^2$  тест та ANOVA дали змогу визначити ті змінні, які мають взаємозв'язок за змінною «group.I». Такими змінними є: вік, спеціальність і т.д. (загалом їх 92).





Встановлено, що для моделі мультиноміальної логістичної регресії, метою якої є оцінка ймовірності належності респондента до однієї з груп (значення змінної «group»), значимими предикторами є сімейний стан, наявність дітей, стосунки в сім'ї, нездатність пристосуватися до середовища, якість сну, низька продуктивність, підвищена втомлюваність, тютюнопаління, оцінка власного здоров'я, необхідність консультації лікаря, рівень тривоги, депресії, нейротизму. Коефіцієнти моделі визначають величину зміни логарифму шансу перебування у тій чи іншій групі («комплексній», «контрольній», «стандартній») проти («порівняльній») групи

здорових, якщо значення предиктора збільшиться на одиницю. Точність передбачення на загальній вибірці становила 86,35%, а на тестовій вибірці (основна група, що включала хворих на НПР) – 88,52%. Модель дерева рішень, побудована на тих же предикторах, що і модель мультиноміальної регресії, (схематичний рисунок) дає точність прогнозу 89,95% на навчальній вибірці і 87,1% на тестовій вибірці. Зважаючи на досить високу точність, обидві моделі (мультиноміальної логістичної регресії та дерева рішень) можуть бути рекомендовані для використання при комплексному оцінюванні анкет респондентів.

### ЛІТЕРАТУРА

1. Европейская декларация по охране психического здоровья. Проблемы и пути решения. – ВОЗ, 2005. – 9 с.
2. Баевский Р.М. Прогнозирование состояний на грани нормы и патологии. М.: Медицина, 1979. – 295 с.
3. Абрамов В.А., Лебедев Д.С. Психическая дезадаптация //Журнал психиатрии и мед. психологии. - 1996. - № 1(2). – С. 45-55.
4. Марута Н.О., Панько Т.В., Явдак І.О., Семікіна О.Є., Стадник А.В. Показник якості життя у хворих на афективні розлади та його динаміка в процесі лікування //Український вісник психоневрології. – Т.10, вип. 2(31). – 2002. – С. 121 – 123.

5. Ведяев Ф.П., Воробьева Т.М. Модели и механизмы эмоциональных стрессов. – К.: Здоров'я, 1983. – 136 с.
6. Соколова И.М. Методы исследования адаптации студентов. – Харьков, 2001. – 276 с.
7. Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. – К.: Морион, 2000. – 320 с.
8. Кулачев А.П. Методы и средства анализа данных в среде Windows. STADIA. – М.:Информатика и компьютеры, 1999. – 341 с.

### REFERENCES

1. European Declaration on Mental Health. Problems and solutions. - WHO, 2005, 9 p. [In Russ.]
2. Bayevsky P.M. Prediction of conditions on the verge of norm and pathology. Medicine. 1979, 295 p.[In Russ.]
3. Abramov V.A., Lebedev D.S. Mental maladaptation. Journal of Psychiatry and Medical Psychology. 1996, no. 1 (2), pp. 45-55. [In Russ.]
4. Maruta N.O., Panko T.V., Yavdak I.O., Semikina O.Ye., Stadnik A.V. The indicator of life's suffering in ailments on an affective path and that dynamics in the process of recovery. Ukrainian Bulletin of Psychoneurology. 2002, no.10, VIP. 2 (31), 2002, pp. 121-123. [In Ukr.]

5. Vedyayev F.P., Vorobyova T.M. Models and mechanisms of emotional stress. Zdorovya. 1983, 136 p. [In Russ.]
6. Sokolova I.M. Research methods for students' adaptation. 2001, 276 p. [In Rus.]
7. Lapach S.N., Chubenko A.V., Babich P.N. Statistical methods in biomedical research using Excel. Morion. 2000, 320 p. [In Russ.]
8. Kulachev A.P. Methods and tools for data analysis in a Windows environment. STADIA. Informatics and computers. 1999, 341 p. [In Russ.]

## МОДЕЛЬ МУЛЬТИНОМИАЛЬНОЙ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ И ПОСТРОЕНИЕ ДЕРЕВА РЕШЕНИЙ ДЛЯ ОПРЕДЕЛЕНИЯ ФАКТОРОВ РИСКА РАЗВИТИЯ НЕПСИХОТИЧЕСКИХ ПСИХИЧЕСКИХ РАССТРОЙСТВ У СТУДЕНТОВ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

**Юрценюк  
Ольга Сидоровна**

Высшее государственное учебное заведение Украины «Буковинский государственный медицинский университет», пл. Театральная, 2, г. Черновцы, Украина, 58002  
yurtsenyuk.olga@bsmu.edu.ua  
ORCID ID: 0000-0002-1450-1530

**Актуальность.** Важность современных исследований, связанных с непсихотическими психическими расстройствами у студентов высших учебных заведений, определяется необходимостью построения математической модели с помощью которой можно выделить факторы, которые достоверно влияют на состояние психического здоровья молодых людей.

**Цель исследования.** Построение математической модели влияния различных факторов на уровень психического здоровья студентов с непсихотическими психическими расстройствами.

**Материалы и методы.** Нами, в течение 2015-2017 гг., с соблюдением принципов биоэтики и деонтологии проведено сплошное комплексное обследование 1235 студентов. Применены методы: клинический, клинико-психопатологический, клинико-эпидемиологический, клинико-анамнестический, экспериментально-психологический и статистический. Выборка не имела существенных различий по половым и возрастным признакам, местом проживания, формой обучения. Обследование проводили в межсессионный период.

**Результаты.** В результате проведения дисперсионного анализа мы получили статистически значимые результаты (ANOVA-теста) между категориальной переменной «группа I» и сменными обозначаемыми все шкалы СБОУ 1. ипохондрии (Hs),  $p < 2,2 * 10^{-16}$ ; 2. депрессии (D),  $p < 2 * 10^{-16}$ ; 3. истерии (Hy),  $p < 2 * 10^{-16}$ ; 4. психопатии (Pd),  $p < 2,89 * 10^{-14}$ ; 6. параноальности (Pa),  $p < 2 * 10^{-16}$ ; 7. психастении (Pt),  $p < 2 * 10^{-16}$ ; 8. шизоидности (Sc),  $p < 7,04 * 10^{-16}$ ; 9. гипотонии [мании] (Ma),  $p < 1,95 * 10^{-10}$ ; шкале лжи (L),  $p < 2,74 * 10^{-11}$ ; шкале вероятности (F),  $p < 4,66 * 10^{-8}$ ; шкале коррекции (K),  $p < 3,6 * 10^{-13}$ .

Установлено, что для модели мультиномиальной логистической регрессии, целью которой является оценка вероятности принадлежности респондента к одной из групп (значение переменной «group.I»), значимыми предикторами является семейное положение, наличие детей, отношения в семье, неспособность приспособиться к среде, качество сна, низкая производительность, повышенная утомляемость, курение, оценка собственного здоровья, необходимости консультации врача, уровень тревоги, депрессии, нейротизма.

Коэффициенты модели определяют величину изменения логарифма шанса пребывания в той или иной группе («комплексной», «контрольной», «стандартной») против («сравнительной») группы здоровых, если значение предиктора увеличится на единицу. Точность предсказания на общей выборке составила 86,35%, а на тестовой выборке (основная группа, включавшая больных НПР) - 88,52%.

**Выводы.** Таким образом, проведенные тест и ANOVA позволили определить те переменные, которые имеют взаимосвязь с переменной «group.I». Такими переменными являются: возраст, специальность и т.д. (Всего их 92). Модель дерева решений, построенная на тех же предикторах, что и модель мультиномиальной регрессии, (схематический рисунок) дает точность прогноза 89,95% в обучающей выборке и 87,1% на тестовой выборке.

**Ключевые слова:** студенты, непсихотические психические расстройства, множественная линейная регрессия, мультиномиальная логистическая регрессия, дерево решений.

## MODEL OF MULTINOMIAL LOGISTIC REGRESSION AND CONSTRUCTION OF A DECISION TREE FOR DETERMINATION OF RISK FACTORS IN THE DEVELOPMENT OF NON-PSYCHOTIC MENTAL DISORDERS AMONG STUDENTS OF HIGHER EDUCATIONAL INSTITUTIONS

**Yurtsenyuk Olha**

Higher State Educational Institution of Ukraine "Bukovynian State Medical University",  
Teatralna sq., 2, Chernivtsi, Ukraine, 58002  
yurtsenyuk.olga@bsmu.edu.ua  
ORCID ID: 0000-0002-1450-1530

**Topicality.** The importance of modern research related to non-psychotic mental disorders in students of higher education is determined by the need to build a mathematical model that can identify factors that significantly affect the mental health of young people.

The aim of the study. Construction of a mathematical model of the influence of various factors on the level of mental health of students with non-psychotic mental disorders.

**Materials and methods.** During 2015-2017, we conducted a continuous comprehensive survey of 1,235 students in compliance with the principles of bioethics and deontology. Applied methods: clinical, clinical-psychopathological, clinical-epidemiological, clinical-anamnestic, experimental-psychological and statistical. The sample did not have significant differences in gender and age, place of residence, form of education. The survey was conducted in the intersessional period.

**Results.** As a result of analysis of variance, we obtained statistically significant results (ANOVA-test) between the category variable "group I" and variables denoting all scales mini mult test: 1. hypochondria (Hs),  $p < 2,2 * 10^{-16}$ ; 2. depression (D),  $p < 2 * 10^{-16}$ ; 3. hysteria (Hy),  $p < 2 * 10^{-16}$ ; 4. psychopathy (Pd),  $p < 2,89 * 10^{-14}$ ; 6. paranoia (Pa),  $p < 2 * 10^{-16}$ ; 7. psychasthenia (Pt),  $p < 2 * 10^{-16}$ ; 8. schizophrenia (Sc),  $p < 7,04 * 10^{-16}$ ; 9. hypotension [mania] (Ma),  $p < 1,95 * 10^{-10}$ ; lie scale (L),  $p < 2,74 * 10^{-11}$ ; probability scale (F),  $p < 4,66 * 10^{-8}$ ; correction scale (K),  $p < 3,6 * 10^{-13}$ .

It is established that for the model of multinomial logistic regression, the purpose of which is to assess the probability of belonging of the respondent to one of the groups (value of the variable "group.I"), significant predictors are marital status, presence of children, family relations, inability to adapt to the environment. quality of sleep, low productivity, increased fatigue, smoking, assessment of their own health, the need to consult a doctor, the level of anxiety, depression, neuroticism.

The coefficients of the model determine the magnitude of the change in the logarithm of the chance of being in a group ("complex", "control", "standard") against the ("comparative") group of healthy people, if the value of the predictor increases by one. The prediction accuracy in the general sample was 86.35%, and in the test sample (the main group, which included patients with NPR) - 88.52%.

**Conclusions.** Thus, the test and ANOVA made it possible to identify those variables that are related to the variable "group.I". Such variables are: age, specialty, etc. (a total of 92). The decision tree model, built on the same predictors as the multinomial regression model (schematic drawing), gives an accuracy of 89.95% on the training sample and 87.1% on the test sample.

**Key words:** students, nonpsychotic mental disorders, multiple linear regression, multinomial logistic regression, decision tree.