Одеський національний університет імені І. І. Мечникова
Факультет математики, фізики та інформаційних технологій
Кафедра оптимального керування і економічної кібернетики

Кваліфікаційна робота

на здобуття ступеня вищої освіти «бакалавр»

«Застосування подвійного машинного навчання для причинно-наслідкового аналізу»

«Double Machine Learning for Causal Inference»

Виконав: здобувач денної форми навчання
спеціальності 113 Прикладна математика
Освітня програма «Прикладна математика»
Шляхов Данііл Віталійович

Керівник: канд. фіз.-мат. наук, доц. Страхов Є. М.
Рецензент: старший викладач Платонов В.В.

Рекомендовано до захисту:
Протокол засідання кафедри
№ ___ від _____ 2026 р.
Завідувач кафедри
_____ _____

Захищено на засіданні ЕК № _____
Протокол № ___ від _____ 2026 р.
Оцінка _____ / _____ / _____
Голова ЕК
_____ _____

Одеса — 2026 р.

ЗМІСТ

Вступ 3

1 Огляд методів причинно-наслідкового аналізу та оцінювання дослідницьких гіпотез 6

1.1 Причинно-наслідковий аналіз та його відмінність від прогнозування 6

1.2 Спостережувальні дані, конфаундинг і контрольні змінні . . 7

1.3 Контрфактичний підхід і поняття причинного ефекту 9

1.4 Метод Double Machine Learning 10

1.5 Оцінювання статистичної значущості та інтерпретація гіпотез 13

2 Аналіз даних та попередній дослідницький аналіз 16

2.1 Опис набору даних та попередня обробка 16

2.2 Загальна характеристика вибірки 18

2.3 Соціально-демографічні та соціальні характеристики 19

2.4 Клінічні та травматологічні характеристики 20

2.5 Афективні та функціональні показники 22

2.6 Когнітивні показники 24

2.7 Асоціативний аналіз соціального профілю клінічних категорій 27

2.8 Аналіз повторного тестування 29

Висновки 32

Список використаних джерел 36

ВСТУП

У сучасних прикладних дослідженнях дедалі більшого значення набуває не лише прогнозування певного стану або показника, а й пояснення того, які чинники можуть бути пов'язані з його формуванням та зміною. Традиційні методи машинного навчання ефективно застосовуються у задачах класифікації, регресії та прогнозування, однак висока точність моделі сама по собі не дає підстав для висновків про причинно-наслідкові зв'язки [2]. Через це в останні роки особливу увагу привертають методи причинно-наслідкового аналізу, які поєднують статистичне оцінювання, контрфактичний підхід і можливості сучасних алгоритмів машинного навчання [3].

Особливо актуальною ця проблема є під час аналізу медико-психологічних даних, де досліджувані показники рідко формуються під впливом одного ізольованого чинника [4]. Психологічний стан, когнітивне функціонування, рівень соціальної адаптації та загальне відновлення пацієнта можуть залежати від поєднання клінічних, демографічних, соціальних і поведінкових характеристик [5]. У таких умовах звичайний кореляційний аналіз або стандартна регресійна модель можуть бути недостатніми, оскільки вони не завжди дозволяють відокремити потенційний самостійний вплив окремого чинника від впливу супутніх змінних [1].

У межах цієї роботи увага зосереджується на даних військовослужбовців із різним клінічним профілем, зокрема з проявами посттравматичного стресового розладу, наслідками черепно-мозкової травми, акубаротравмою, порушеннями зору та іншими особливостями стану. Наявні змінні охоплюють показники тривоги та депресії за шкалою HADS, симптоми ПТСР за PCL, наслідки контузії за NSI, когнітивні показники RBANS, рівень функціонування PSP, а також соціальні та демографічні характеристики, зокрема вік, наявність сім'ї, соціальної підтримки, участь у бойових діях і кількість черепно-мозкових травм. Така структура даних створює потребу в методах, здатних працювати зі складними взаємозв'язками між змінними та зменшувати ризик хибної інтерпретації статистичних залежностей.

Актуальність теми зумовлена необхідністю застосування сучасних методів причинно-наслідкового аналізу до спостережувальних медико-психологічних даних. На відміну від рандомізованих експериментів, такі дані формуються в реальних умовах, де дослідник не контролює розподіл чинників між групами [1]. Це ускладнює оцінювання потенційного впливу окремих клінічних, соціальних і демографічних характеристик на психологічний, когнітивний або функціональний стан пацієнтів, оскільки між цими характеристиками можуть існувати приховані або явні взаємозалежності. У цьому контексті використання Double Machine Learning є доцільним, оскільки цей підхід поєднує гнучкість моделей машинного навчання з фор-

малізованим оцінюванням потенційних причинних ефектів.

Об'єктом дослідження є медико-психологічні дані військовослужбовців із різними клінічними характеристиками та показниками когнітивного, емоційного і функціонального стану.

Предметом дослідження є методи оцінювання потенційних причинних ефектів у спостережувальних даних, зокрема підхід Double Machine Learning, та його застосування для аналізу впливу окремих чинників на психологічні, когнітивні й функціональні показники військовослужбовців.

Метою роботи є застосування підходу Double Machine Learning для оцінювання потенційного впливу клінічних, соціальних і демографічних чинників на психологічний, когнітивний та функціональний стан військовослужбовців на основі наявних медико-психологічних даних.

Для досягнення поставленої мети передбачено виконання таких завдань:

- дослідити теоретичні засади причинно-наслідкового аналізу та підходу Double Machine Learning;
- охарактеризувати структуру наявного набору даних і основні групи змінних;
- провести попередню обробку даних та підготувати їх до моделювання;
- визначити ключові результативні показники, що характеризують психологічний, когнітивний і функціональний стан пацієнтів;
- побудувати моделі для оцінювання потенційного впливу окремих чинників з урахуванням контрольних змінних;
- проаналізувати отримані результати та оцінити можливості й обмеження застосованого підходу.

У роботі використовуються методи описової статистики, попередньої обробки даних, регресійного аналізу, машинного навчання та причинно-наслідкового моделювання. Основним методологічним підходом є Double Machine Learning, який дозволяє оцінювати цільовий ефект за наявності великої кількості контрольних змінних і потенційного конфаундингу [6].

Наукова новизна роботи полягає в застосуванні підходу Double Machine Learning до аналізу медико-психологічних даних військовослужбовців із різними клінічними профілями. На відміну від суто описового або кореляційного аналізу, у роботі здійснюється оцінювання потенційних причинних ефектів окремих чинників із урахуванням контрольних змінних, що дає змогу відокремити статистичні асоціації від ефектів, які зберігають статистичну підтримку після контролю за супутніми характеристиками.

Практичне значення роботи полягає в тому, що застосований підхід дає змогу виділити чинники, які можуть мати самостійне значення для оцінювання психологічного стану військовослужбовців. Зокрема, результати

роботи можуть бути корисними для обґрунтування уважнішого скринінгу тривожної симптоматики у пацієнтів із бойовим досвідом, урахування кількості перенесених черепно-мозкових травм під час оцінювання ризику ПТСР, а також розгляду соціальної підтримки як потенційного напряму зниження депресивної симптоматики. Водночас отримані результати слід інтерпретувати як оцінки потенційних ефектів у спостережувальних даних, а не як остаточне доведення причинності.

РОЗДІЛ 1

ОГЛЯД МЕТОДІВ ПРИЧИННО-НАСЛІДКОВОГО АНАЛІЗУ ТА ОЦІНЮВАННЯ ДОСЛІДНИЦЬКИХ ГІПОТЕЗ

1.1 Причинно-наслідковий аналіз та його відмінність від прогнозування

У сучасному аналізі даних методи машинного навчання найчастіше застосовуються для розв'язання задач прогнозування, класифікації та виявлення складних закономірностей у великих масивах інформації. Такі підходи дають змогу будувати моделі, які на основі наявних ознак передбачають значення певного результативного показника або відносять об'єкт до однієї з визначених категорій. У прикладних дослідженнях це має значну практичну цінність, оскільки дозволяє оцінювати ризики, передбачати ймовірність певного стану або автоматизувати процес ухвалення рішень.

Водночас висока прогнозна точність моделі не означає, що вона пояснює механізм формування досліджуваного явища. Прогнозна модель може ефективно використовувати статистичні асоціації між змінними, однак такі асоціації не завжди мають причинний характер. Наприклад, певна ознака може добре передбачати результат не тому, що вона безпосередньо впливає на нього, а тому, що пов'язана з іншими, не врахованими або опосередкованими чинниками. Саме тому в літературі підкреслюється необхідність розрізняти моделі, орієнтовані на прогнозування, та моделі, призначені для пояснення причинних механізмів [2].

Причинно-наслідковий аналіз має іншу дослідницьку логіку. Якщо прогнозування відповідає переважно на питання про те, яке значення результату можна очікувати за наявних даних, то причинний аналіз спрямований на питання про те, як зміниться результат у разі зміни певного чинника. Іншими словами, у центрі причинного підходу перебуває не лише статистичний зв'язок між змінними, а оцінювання потенційного ефекту впливу. Це особливо важливо у прикладних галузях, де дослідника цікавить не тільки передбачення стану, а й розуміння того, які чинники можуть бути пов'язані з його виникненням, посиленням або зменшенням.

У статистичному моделюванні така різниця має принципове значення. Кореляція між двома змінними може бути високою, але вона не дає достатніх підстав для твердження про причинність. Причиною цього є можливий вплив третіх змінних, зворотна причинність, систематичні від-

мінності між групами спостережень або інші джерела зміщення. Тому причинно-наслідковий аналіз вимагає не лише побудови моделі, а й чіткого визначення досліджуваного чинника, результативної змінної, контрольних змінних та припущень, за яких отриманий ефект може бути інтерпретований як потенційно причинний.

Особливої актуальності це набуває у медичних, психологічних і соціальних дослідженнях, де експериментальний контроль часто є обмеженим або неможливим з етичних та організаційних причин. У таких умовах дослідник працює переважно зі спостережувальними даними, у яких розподіл чинників між групами не є випадковим. Відповідно, просте порівняння груп або використання звичайної регресійної моделі може призводити до хибних висновків, якщо не враховано супутні характеристики, що одночасно пов'язані і з чинником впливу, і з результатом [1, 5].

1.2 Спостережувальні дані, конфаундинг і контрольні змінні

У прикладних дослідженнях часто використовуються спостережувальні дані, тобто такі дані, у яких дослідник не керує розподілом об'єктів між групами та не задає значення досліджуваного чинника. На відміну від рандомізованого експерименту, у спостережувальному дослідженні значення чинників формуються природним чином. Через це порівнювані групи можуть відрізнятися не лише за основною змінною впливу, а й за іншими характеристиками, які також пов'язані з результатом.

Нехай Y позначає результативну змінну, D – досліджуваний чинник впливу, а X – вектор контрольних змінних. У найпростішому випадку зв'язок між результатом і чинником можна описати лінійною моделлю

$$Y_i = \alpha + \theta D_i + \varepsilon_i, \tag{1.1}$$

де Y_i – значення результативного показника для i-го спостереження, D_i – значення досліджуваного чинника, θ – оцінюваний ефект чинника, а ε_i – випадкова похибка.

Однак така модель є коректною лише за умови, що досліджуваний чинник D_i не пов'язаний із неврахованими характеристиками, які входять до похибки ε_i. Формально це можна записати як

$$\mathbb{E}(\varepsilon_i \mid D_i) = 0. \tag{1.2}$$

Якщо ця умова порушується, оцінка параметра θ може бути зміщеною. У спостережувальних даних таке порушення виникає досить часто, оскільки

чинник впливу може бути пов'язаний з іншими характеристиками об'єкта дослідження [1].

Для врахування таких характеристик до моделі вводять контрольні змінні. Тоді модель набуває вигляду

$$Y_i = \alpha + \theta D_i + X_i'\beta + \varepsilon_i, \tag{1.3}$$

де X_i – вектор контрольних змінних для i-го спостереження, а β – вектор відповідних коефіцієнтів. У цьому випадку параметр θ інтерпретується як ефект змінної D_i на Y_i за умови фіксованих значень контрольних змінних X_i.

Ключовою умовою для коректної інтерпретації такого ефекту є умовна незалежність похибки від досліджуваного чинника після врахування контрольних змінних:

$$\mathbb{E}(\varepsilon_i \mid D_i, X_i) = 0. \tag{1.4}$$

Іншими словами, після контролю за X_i у моделі не повинно залишатися неврахованих чинників, які одночасно впливають і на D_i, і на Y_i.

Змінна, яка одночасно пов'язана з досліджуваним чинником і результативною змінною, називається конфаундером. Якщо таку змінну не врахувати, частина її впливу може бути помилково приписана основному чиннику D. У такому випадку оцінка ефекту θ не відображатиме самостійний внесок досліджуваної змінної.

Схематично конфаундинг можна подати так:

$$X \to D, \qquad X \to Y, \tag{1.5}$$

де X впливає як на досліджуваний чинник D, так і на результат Y. Якщо у моделі оцінюється лише зв'язок $D \to Y$, але не враховується X, то отриманий ефект може бути зміщеним [5].

У випадку бінарного чинника впливу, коли $D \in \{0,1\}$, просте порівняння середніх між групами можна записати як

$$\Delta = \mathbb{E}(Y \mid D = 1) - \mathbb{E}(Y \mid D = 0). \tag{1.6}$$

Ця величина показує різницю середніх результатів між групами, але не обов'язково дорівнює причинному ефекту. Причина полягає в тому, що групи з $D = 1$ та $D = 0$ можуть відрізнятися за контрольними змінними X.

Для зменшення такого зміщення розглядають умовне порівняння:

$$\Delta(X) = \mathbb{E}(Y \mid D = 1, X) - \mathbb{E}(Y \mid D = 0, X). \tag{1.7}$$

У цьому випадку порівняння здійснюється між об'єктами з однаковими або

подібними значеннями контрольних змінних. Саме така логіка лежить в основі більшості методів причинно-наслідкового аналізу спостережувальних даних.

На практиці залежності між Y, D та X можуть бути складними, нелінійними та багатовимірними. Тому звичайна лінійна регресія не завжди достатньо точно описує ці зв'язки. Якщо функціональна форма моделі задана неправильно, це також може призвести до зміщення оцінки ефекту. Саме тому в сучасних підходах дедалі частіше використовуються методи машинного навчання, які дозволяють гнучкіше моделювати залежності між змінними.

Основна проблема спостережувальних даних полягає в тому, що досліджуваний чинник може бути пов'язаний із супутніми характеристиками, які також впливають на результат. Для зменшення зміщення необхідно враховувати контрольні змінні та будувати моделі, здатні описувати зв'язки між чинником, результатом і конфаундерами. Це створює методологічне підґрунтя для застосування Double Machine Learning, який поєднує причинно-наслідкову логіку з гнучкими алгоритмами машинного навчання.

1.3 Контрфактичний підхід і поняття причинного ефекту

Одним із базових підходів до формалізації причинно-наслідкових зв'язків є контрфактичний підхід, або підхід потенційних результатів. Його основна ідея полягає в тому, що ефект певного чинника визначається через порівняння двох можливих станів одного й того самого об'єкта: результату за наявності впливу та результату за його відсутності [3].

Нехай D позначає досліджуваний чинник впливу, де $D = 1$ відповідає наявності впливу, а $D = 0$ – його відсутності. Для кожного об'єкта можна теоретично визначити два потенційні результати: $Y(1)$ – результат за умови наявності впливу, та $Y(0)$ – результат за умови його відсутності. Тоді індивідуальний причинний ефект можна записати як:

$$\tau = Y(1) - Y(0). \tag{1.8}$$

Ця формула відображає логіку причинного аналізу: дослідника цікавить не лише те, який результат спостерігається фактично, а й те, яким він міг би бути за іншого значення досліджуваного чинника. Проте на практиці для одного й того самого об'єкта неможливо одночасно спостерігати обидва потенційні результати. Якщо об'єкт належить до групи з наявним впливом, то спостерігається лише $Y(1)$, а $Y(0)$ залишається контрфактичним,

тобто альтернативним і безпосередньо неспостережуваним результатом. Якщо ж вплив відсутній, то спостерігається $Y(0)$, тоді як $Y(1)$ залишається невідомим.

Через цю проблему в прикладних дослідженнях зазвичай оцінюють не індивідуальний ефект для кожного окремого об'єкта, а середній ефект впливу для певної сукупності. Одним із найпоширеніших показників є середній ефект впливу, або Average Treatment Effect:

$$ATE = \mathbb{E}[Y(1) - Y(0)]. \tag{1.9}$$

Цей показник характеризує середню зміну результативної змінної, яка очікується при переході від відсутності впливу до його наявності [7].

Для коректного оцінювання такого ефекту в спостережувальних даних важливо враховувати контрольні змінні. Якщо групи з наявним і відсутнім впливом істотно відрізняються за супутніми характеристиками, то проста різниця між їхніми середніми значеннями може бути зміщеною. Саме тому причинний ефект не слід ототожнювати зі звичайною кореляцією або різницею середніх між групами [1, 5].

У загальному вигляді логіку оцінювання можна подати так: спочатку дослідник визначає чинник впливу D, результативну змінну Y та набір контрольних змінних X, після чого оцінює, як змінюється Y при зміні D з урахуванням X. Такий підхід дозволяє наблизитися до контрфактичного порівняння, хоча повністю відтворити умови рандомізованого експерименту за допомогою спостережувальних даних зазвичай неможливо.

Контрфактичний підхід є важливим для цієї роботи, оскільки він задає логіку перевірки дослідницьких гіпотез. У межах такої логіки гіпотеза формулюється не лише як припущення про наявність статистичного зв'язку між двома змінними, а як припущення про потенційний вплив одного чинника на певний результативний показник. Саме тому подальше оцінювання здійснюється з урахуванням контрольних змінних і обмежень, властивих спостережувальним даним.

1.4 Метод Double Machine Learning

Double Machine Learning, або подвійне машинне навчання, є сучасним підходом до оцінювання потенційних причинних ефектів у задачах, де дослідник працює зі спостережувальними даними та великою кількістю контрольних змінних. Основна ідея методу полягає в тому, щоб використати гнучкість алгоритмів машинного навчання для моделювання складних залежностей між змінними, але при цьому зберегти можливість статистично

коректного оцінювання цільового ефекту [6].

У межах причинно-наслідкового аналізу зазвичай виділяють три основні групи змінних: результативну змінну Y, досліджуваний чинник впливу D та вектор контрольних змінних X. Результативна змінна відображає показник, зміну якого потрібно пояснити або оцінити. Досліджуваний чинник є змінною, потенційний вплив якої аналізується. Контрольні змінні враховують супутні характеристики, які можуть бути пов'язані як із чинником впливу, так і з результатом.

У спрощеному вигляді модель можна записати так:

$$Y = \theta D + g(X) + \varepsilon, \tag{1.10}$$

де θ є основним параметром інтересу, тобто оцінкою ефекту чинника D на результат Y, $g(X)$ описує вплив контрольних змінних на результат, а ε є випадковою похибкою. У класичній лінійній регресії функція $g(X)$ часто задається у лінійній формі. Однак у реальних даних залежності між змінними можуть бути нелінійними та містити складні взаємодії, тому така специфікація не завжди є достатньою.

Метод Double Machine Learning розглядає функцію $g(X)$ як допоміжну функцію, або nuisance function. Вона не є головним об'єктом дослідження, але її коректне оцінювання важливе для отримання надійної оцінки параметра θ. Крім того, у методі оцінюється залежність самого чинника впливу D від контрольних змінних:

$$D = m(X) + v, \tag{1.11}$$

де $m(X)$ описує очікуване значення чинника впливу за наявних контрольних змінних, а v є залишковою частиною, не поясненою цими змінними.

Ідея методу полягає в тому, щоб спочатку за допомогою моделей машинного навчання оцінити дві допоміжні залежності: залежність результату Y від контрольних змінних X та залежність чинника впливу D від контрольних змінних X. Після цього з Y та D вилучається частина, яка пояснюється контрольними змінними. У результаті аналізується зв'язок між залишковою частиною результату та залишковою частиною чинника впливу. Така процедура дозволяє зменшити вплив конфаундингу, пов'язаного з виміряними контрольними змінними.

Формально залишки можна подати так:

$$\tilde{Y} = Y - \hat{g}(X), \tag{1.12}$$

$$\tilde{D} = D - \hat{m}(X), \tag{1.13}$$

де $\hat{g}(X)$ та $\hat{m}(X)$ є оціненими за допомогою машинного навчання функціями.

Після цього параметр θ оцінюється на основі зв'язку між $\tilde{Y}$ та $\tilde{D}$. У такій постановці ефект D на Y оцінюється після усунення тієї частини варіації, яка пояснюється контрольними змінними.

Назва Double Machine Learning пов'язана з тим, що машинне навчання використовується двічі: по-перше, для моделювання результативної змінної через контрольні змінні, по-друге, для моделювання досліджуваного чинника через ці самі контрольні змінні. Завдяки цьому метод дозволяє гнучко враховувати складні залежності в даних, не обмежуючись лише лінійною формою зв'язку.

Важливою складовою Double Machine Learning є процедура cross-fitting. Її суть полягає в тому, що вибірка поділяється на кілька частин. На одній частині даних навчаються допоміжні моделі, а на іншій обчислюються залишки та оцінюється цільовий ефект. Потім ці ролі змінюються, і процедура повторюється для інших частин вибірки. Такий підхід зменшує ризик перенавчання, оскільки оцінювання ефекту здійснюється не на тих самих даних, на яких навчалися допоміжні моделі [6].

Ще однією важливою властивістю методу є ортогоналізація. Вона означає, що оцінка цільового параметра θ стає менш чутливою до помилок в оцінюванні допоміжних моделей. Інакше кажучи, навіть якщо моделі $\hat{g}(X)$ та $\hat{m}(X)$ не є ідеальними, правильно побудована процедура Double Machine Learning дозволяє зменшити вплив цих помилок на кінцеву оцінку потенційного ефекту. Саме ця властивість відрізняє DML від простого використання машинного навчання для прогнозування результату.

У практичних задачах як допоміжні моделі можуть використовуватися різні алгоритми машинного навчання: лінійна регресія, Lasso-регресія, випадковий ліс, градієнтний бустинг, нейронні мережі та інші методи. Вибір конкретного алгоритму залежить від структури даних, кількості спостережень, кількості змінних та очікуваного характеру залежностей між ними. Основна вимога полягає в тому, щоб ці моделі достатньо добре описували зв'язки між контрольними змінними, чинником впливу та результативною змінною.

Після оцінювання параметра θ дослідник аналізує не лише саме значення ефекту, а й його статистичну невизначеність. У цій роботі для цього використовуються оцінене значення ефекту та 95% довірчий інтервал. Якщо довірчий інтервал не включає нуль, то ефект розглядається як такий, що отримав статистичну підтримку в межах обраної специфікації моделі. Водночас у спостережувальних даних такий результат слід інтерпретувати обережно, оскільки він залежить від коректності вибору контрольних змінних і припущення про відсутність суттєвого неврахованого конфаундингу.

Метод Double Machine Learning є особливо корисним у тих випадках, коли дослідника цікавить не побудова найточнішого прогнозу, а оцінювання

окремого цільового ефекту. На відміну від звичайного машинного навчання, де головним критерієм часто є точність передбачення, у DML основна увага приділяється коректному відокремленню ефекту досліджуваного чинника від впливу контрольних змінних. Саме тому цей підхід добре узгоджується із завданнями причинно-наслідкового аналізу в прикладних дослідженнях.

1.5 Оцінювання статистичної значущості та інтерпретація гіпотез

У межах причинно-наслідкового аналізу дослідницька гіпотеза формулюється як припущення про можливий вплив певного чинника на результативний показник. На відміну від звичайного кореляційного аналізу, де оцінюється лише сила та напрям статистичного зв'язку між двома змінними, у причинно-наслідковому підході основна увага приділяється оцінюванню ефекту чинника після врахування контрольних змінних. Тому перевірка гіпотези передбачає не лише встановлення наявності асоціації, а й аналіз того, чи зберігається ефект після контролю за супутніми характеристиками.

У загальному вигляді нульова та альтернативна гіпотези можуть бути записані так:

$$H_0 : \theta = 0, \tag{1.14}$$

$$H_1 : \theta \neq 0, \tag{1.15}$$

де θ позначає оцінюваний ефект досліджуваного чинника на результативну змінну. Нульова гіпотеза означає, що після врахування контрольних змінних досліджуваний чинник не має статистично значущого ефекту на результат. Альтернативна гіпотеза, навпаки, передбачає, що такий ефект існує.

Основним кількісним результатом оцінювання є значення ефекту $\hat{\theta}$. Його знак показує напрям зв'язку між досліджуваним чинником і результативною змінною. Якщо $\hat{\theta} > 0$, то досліджуваний чинник пов'язаний зі збільшенням результативного показника. Якщо $\hat{\theta} < 0$, то наявність або збільшення чинника пов'язане зі зменшенням результативної змінної. Водночас сам знак ефекту не є достатньою підставою для висновку про статистичну підтримку гіпотези, оскільки необхідно враховувати невизначеність оцінки.

Для оцінювання невизначеності використовується стандартна похибка:

$$SE(\hat{\theta}), \tag{1.16}$$

яка характеризує варіативність оцінки ефекту. На основі стандартної похибки будується довірчий інтервал. Для 95% довірчого інтервалу викори-

стовується формула:

$$CI_{95\%} = \hat{\theta} \pm 1.96 \cdot SE(\hat{\theta}). \tag{1.17}$$

Довірчий інтервал показує діапазон значень, у якому з певним рівнем довіри може знаходитися істинне значення ефекту. У практичній частині цієї роботи статистична підтримка гіпотез оцінюється саме за 95% довірчим інтервалом. Якщо такий інтервал не включає нуль, результат розглядається як статистично підтриманий у межах обраної специфікації моделі. Якщо інтервал включає нуль, отриманих даних недостатньо для висновку про наявність статистично значущого ефекту.

У контексті Double Machine Learning оцінка ефекту інтерпретується як середній потенційний вплив досліджуваного чинника на результативну змінну після врахування контрольних змінних. Якщо досліджуваний чинник є бінарним, то значення $\hat{\theta}$ можна розглядати як середню різницю в результаті між групою з наявним чинником і групою без нього після статистичного контролю за супутніми характеристиками. Якщо чинник є неперервним, то $\hat{\theta}$ показує очікувану зміну результативного показника при збільшенні чинника на одну одиницю за інших рівних умов [6].

Для інтерпретації дослідницької гіпотези важливо враховувати три елементи: напрям ефекту, його величину та статистичну підтримку. Напрям ефекту показує, чи пов'язаний чинник зі збільшенням або зменшенням результативного показника. Величина ефекту дає змогу оцінити його змістовне значення. Статистична підтримка показує, чи можна вважати отриманий ефект достатньо надійним з погляду вибіркової невизначеності.

У роботах зі спостережувальними даними статистично підтриманий результат не слід інтерпретувати як абсолютне доведення причинності. Такий результат означає, що отримані дані підтримують дослідницьку гіпотезу за умови коректного вибору контрольних змінних, правильної специфікації процедури оцінювання та відсутності суттєвих неврахованих конфаундерів. Тому в інтерпретації доцільно використовувати обережні формулювання: «ефект отримав статистичну підтримку», «результати свідчать про можливий вплив», «гіпотеза отримала часткову підтримку в межах обраної специфікації» [1, 5].

Якщо оцінка ефекту має очікуваний напрям, але довірчий інтервал включає нуль, таку гіпотезу не можна вважати статистично підтвердженою. Водночас подібний результат може розглядатися як попередня тенденція, яка потребує перевірки на більшій вибірці або з використанням додаткових змінних. Це особливо важливо для прикладних досліджень, де обсяг вибірки може бути обмеженим, а частина ефектів не досягає формального рівня статистичної значущості через високу варіативність даних.

Оцінювання дослідницьких гіпотез у межах причинно-наслідкового аналізу ґрунтується не лише на наявності статистичного зв'язку, а на аналізі оціненого ефекту, його невизначеності та змістовної інтерпретації. У практичній частині роботи для кожної гіпотези розглядаються значення ефекту, його напрям, 95% довірчий інтервал та відповідність отриманого результату теоретично очікуваному напряму впливу.

РОЗДІЛ 2

АНАЛІЗ ДАНИХ ТА ПОПЕРЕДНІЙ ДОСЛІДНИЦЬКИЙ АНАЛІЗ

2.1 Опис набору даних та попередня обробка

Практична частина роботи виконана з використанням набору даних, наданого співробітниками Інституту очних хвороб та тканинної терапії ім. В.П. Філатова НАМН України. До дослідження залучили пацієнтів із легкою черепно-мозковою травмою та супутньою вибуховою травмою ока зі зниженням зору (за висновками невропатолога й офтальмолога). Обстеження проводили у віддаленому періоді (через 3 і більше місяців після травми згідно з класифікацією Міжнародної асоціації з вивчення травм головного мозку) з дотриманням принципів Гельсінської декларації та за інформованої згоди пацієнтів.

Комплексна оцінка включала:

- Формалізоване інтерв'ю: біографічні відомості, самооцінка рівня стресу та соціальної підтримки.
- Клінічні шкали та опитувальники: шкала особистісного і соціального функціонування (PSP), шпитальна шкала тривоги та депресії (HADS), шкала ПТСР (з опитувальника MINI) та опитувальник болю (PROMIS).
- Когнітивні та нейропсихологічні тести: прогресивні матриці Равена, батарея RBANS (оцінка пам'яті, уваги, мовлення, зорово-просторових функцій) та субтест «Лабіринти» (NAB: Mazes) для дослідження виконавчих функцій планування й прийняття рішень.

Початковий набір даних містив 148 спостережень та 59 змінних. Перевірка структури набору даних показала відсутність дубльованих рядків, що дозволило використовувати всі спостереження для подальшого аналізу. Також було проведено первинну перевірку кількості заповнених і пропущених значень у кожній змінній. Більшість основних показників, необхідних для аналізу, були заповнені достатньо повно. Водночас частина технічних змінних не містила змістовної інформації, а показники повторного обстеження були доступні лише для окремої підгрупи пацієнтів.

На етапі попередньої обробки з набору даних було вилучено колонки `V49`, `V50`, `V51`, `V52` та `V53`, оскільки вони не містили корисних даних для дослідження. Після цього змінні було перейменовано у стандартизований формат англійською мовою. Такий підхід спростив подальшу роботу з даними, зробив код більш читабельним і дозволив чітко групувати показники

за змістом. Повна відповідність між початковими назвами змінних і назвами, використаними в програмному коді, наведена в додатку А разом із фрагментами попередньої обробки даних.

У наборі даних було виділено кілька основних груп змінних. До соціально-демографічних характеристик належать вік пацієнта, вікова група, наявність сім'ї, дітей, друзів або колег, життєвих планів, а також суб'єктивна оцінка стосунків із партнером і рівня соціальної підтримки. Клінічні та травматологічні характеристики включають участь у бойових діях, наявність порушень сну, акубаротравми, кількість перенесених черепно-мозкових травм, строк після останньої травми, вживання психоактивних речовин і наявність діагнозу ПТСР.

Окрему групу становлять психологічні та функціональні показники. До них належать рівень тривоги та депресії за шкалою HADS, показник симптомів ПТСР за шкалою PCL, вираженість постконтузійної симптоматики за шкалою NSI, рівень психосоціального функціонування за PSP, а також показник PROMIS. Ці змінні використовувалися для оцінювання афективного стану, функціонування та загального психологічного профілю пацієнтів.

Нейропсихологічний блок даних включав загальний когнітивний індекс RBANS, окремі когнітивні домени RBANS, показники невербального інтелекту за тестом Равена, робочу пам'ять, виконання лабіринту, а також тести TMT-A і TMT-B. Ці показники дозволяють оцінити пам'ять, увагу, мовні функції, просторові функції, виконавчі функції та швидкість обробки інформації.

У наборі даних також містилися показники повторного тестування, позначені як T2. До них належать повторні значення HADS, TMT, PSP, PROMIS, RBANS та окремих когнітивних доменів. Наявність цих змінних дала змогу окремо проаналізувати динаміку стану пацієнтів у підвибірці тих осіб, які проходили повторне обстеження.

Для частини категоріальних змінних було використано порядкове кодування. Це стосувалося змінних, які мають природний порядок значень, зокрема категорій рівня функціонування, тривоги, депресії, PCL та NSI. Наприклад, рівні стану можуть бути впорядковані як норма, субклінічний рівень і клінічний рівень, а рівні вираженості симптомів як низький, середній і високий. Використання порядкового кодування дозволяє зберегти змістову ієрархію категорій і зробити ці змінні придатними для подальшого статистичного аналізу.

Клінічна категорія пацієнта представлена змінною `group_category`. Для зручності інтерпретації числові значення цієї змінної були зіставлені з текстовими назвами груп: ПТСР, ПТСР з контузією, контузія, без діагнозу та без зору. Така класифікація використовувалася в подальшому дослідни-

цькому аналізі для порівняння соціальних, травматологічних, афективних і когнітивних характеристик між групами.

Попередня обробка даних включала перевірку структури набору, вилучення технічних змінних, перейменування колонок, створення зрозумілих назв клінічних категорій та підготовку змінних до подальшого дослідницького аналізу. Після цього набір даних був використаний для побудови графіків, описової статистики, асоціативного аналізу та подальшої перевірки дослідницьких гіпотез у третьому розділі.

Програмний код, використаний для попередньої обробки даних, побудови графіків і оцінювання моделей Double Machine Learning, наведено в додатку А. Посилання на використаний набір даних наведено в додатку Б.

2.2 Загальна характеристика вибірки

Досліджувана вибірка включала 148 військовослужбовців із різним клінічним профілем. Для подальшого аналізу пацієнтів було розподілено на п'ять клінічних категорій: ПТСР, ПТСР з контузією, контузія, без діагнозу та без зору. Крім того, для загального опису вибірки було використано поділ за віковими групами: 25–30 років, 31–37 років та 38–45 років.

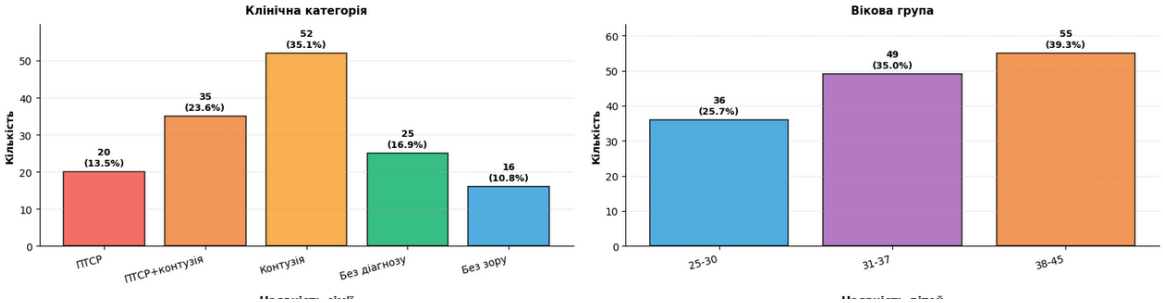


Рис. 2.1. Розподіл пацієнтів за клінічною категорією та віковою групою

Рисунок 2.1 демонструє, що найбільш чисельною клінічною категорією була група пацієнтів із контузією, до якої належали 52 особи, що становить 35.1% вибірки. Другою за чисельністю була група ПТСР з контузією – 35 осіб, або 23.6%. До групи без діагнозу увійшли 25 пацієнтів, що становить 16.9%, тоді як група ПТСР налічувала 20 осіб, або 13.5%. Найменшою була група пацієнтів без зору – 16 осіб, що становить 10.8% вибірки.

Вікова структура вибірки також була нерівномірною. Найбільшу частку становили пацієнти віком 38–45 років – 55 осіб, або 39.3%. До вікової групи 31–37 років належали 49 осіб, що становить 35.0%, а до групи 25–30 років – 36 осіб, або 25.7%. Вибірка представлена переважно військовослужбовцями середнього віку, що є очікуваним для досліджуваної категорії

пацієнтів.

Наведений розподіл свідчить про клінічну та вікову неоднорідність вибірки. Це важливо враховувати під час подальшого аналізу, оскільки відмінності між клінічними групами та віковими категоріями можуть бути пов'язані з психологічними, когнітивними й функціональними показниками пацієнтів.

2.3 Соціально-демографічні та соціальні характеристики

Окрім клінічної категорії та віку, у наборі даних було представлено низку соціальних характеристик пацієнтів. До них належали наявність сім'ї, дітей, друзів або колег, а також життєвих планів. Ці змінні дають змогу оцінити рівень соціальної включеності пацієнтів і можуть бути важливими для подальшого аналізу психологічного стану, адаптації та функціонування.

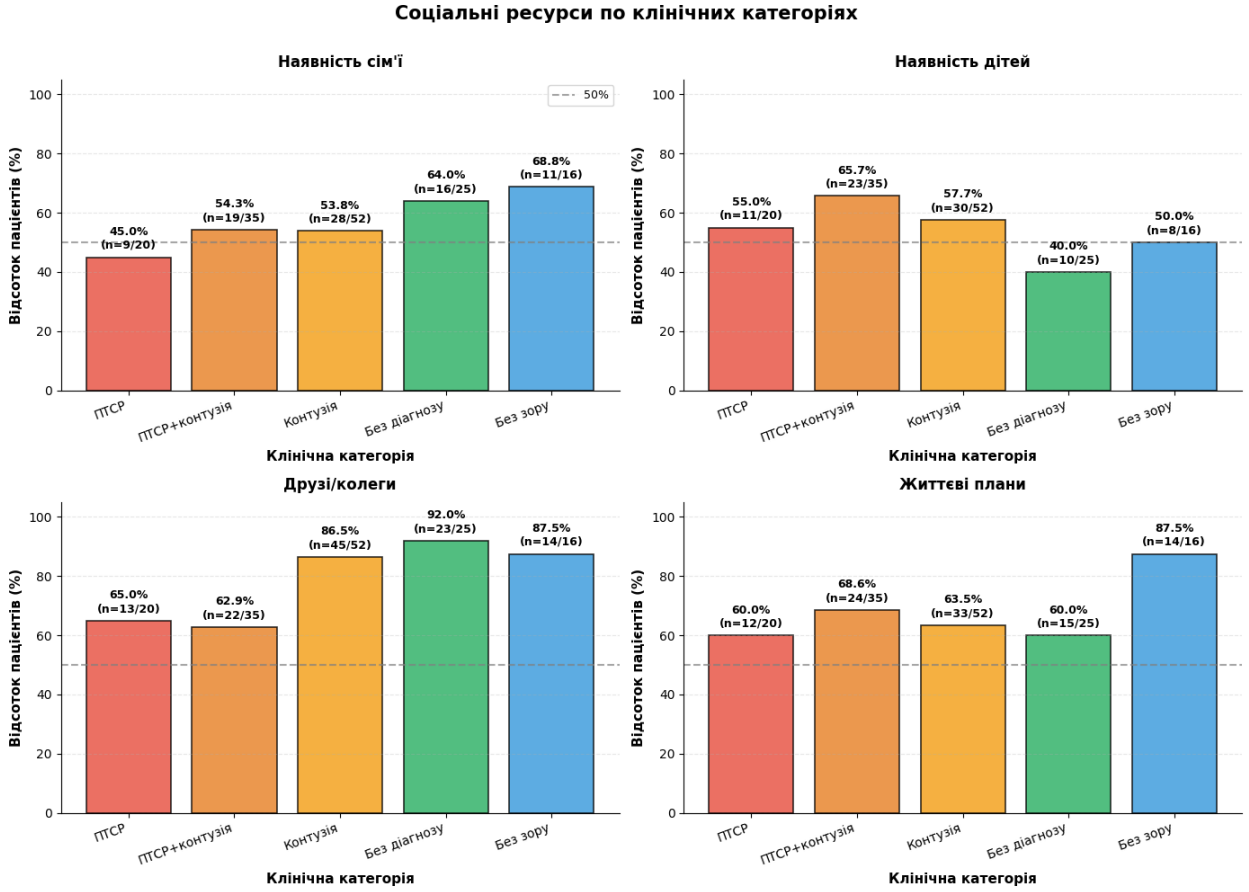


Рис. 2.2. Соціальні ресурси пацієнтів за клінічними категоріями

За даними рис. 2.2, частка пацієнтів із сім'єю відрізнялася між клінічними категоріями. Найвищий показник спостерігався у групі пацієнтів

без зору, де сім'ю мали 68.8% осіб. У групі без діагнозу цей показник становив 64.0%, у групі ПТСР з контузією – 54.3%, у групі контузії – 53.8%, а найнижчим він був у групі ПТСР – 45.0%.

Наявність дітей також мала нерівномірний розподіл між групами. Найбільша частка пацієнтів із дітьми була у групі ПТСР з контузією – 65.7%, а також у групі контузії – 57.7%. У групі ПТСР цей показник становив 55.0%, у групі без зору – 50.0%, тоді як найнижче значення було зафіксовано у групі без діагнозу – 40.0%.

Найвищі значення серед соціальних характеристик були характерні для показника наявності друзів або колег. У групі без діагнозу цей показник становив 92.0%, у групі без зору – 87.5%, а у групі контузії – 86.5%. Нижчі значення спостерігалися у групах ПТСР та ПТСР з контузією, де наявність друзів або колег була зазначена відповідно у 65.0% та 62.9% пацієнтів. Це може свідчити про менш виражену соціальну включеність у групах із посттравматичною симптоматикою.

Життєві плани були наявні у більшості пацієнтів у всіх клінічних категоріях, однак найвищий показник спостерігався у групі без зору – 87.5%. У групі ПТСР з контузією життєві плани мали 68.6% пацієнтів, у групі контузії – 63.5%, а у групах ПТСР і без діагнозу цей показник становив по 60.0%.

Загальний розподіл соціальних ресурсів показує, що соціальні характеристики не є однаковими для всіх клінічних категорій. Найбільш помітні відмінності спостерігаються за показником наявності друзів або колег, де групи з ПТСР мають нижчі значення порівняно з групами без діагнозу, контузією та без зору. Це може мати значення для подальшого аналізу, оскільки соціальні ресурси потенційно пов'язані з рівнем психологічної адаптації, депресивної симптоматики та функціонування пацієнтів.

2.4 Клінічні та травматологічні характеристики

Наступним етапом дослідницького аналізу було вивчення клінічних і травматологічних характеристик вибірки. До цього блоку було віднесено участь у бойових діях, наявність проблем зі сном, акубаротравму, вживання психоактивних речовин, а також наявність діагнозу ПТСР. Ці змінні є важливими для подальшого аналізу, оскільки вони можуть бути пов'язані як із психологічним станом пацієнтів, так і з когнітивним функціонуванням та рівнем загальної адаптації.

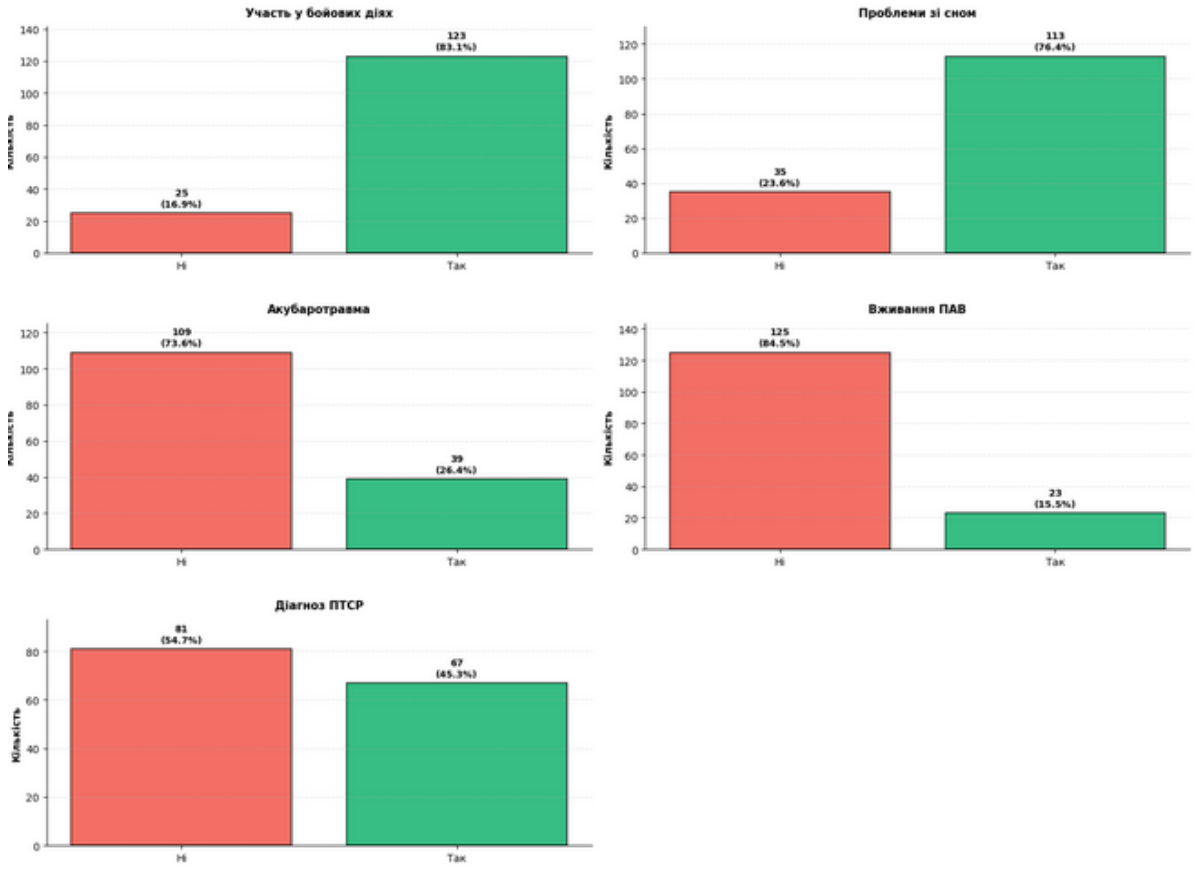


Рис. 2.3. Розподіл основних клінічних та травматологічних характеристик пацієнтів

Рисунок 2.3 показує, що переважна більшість пацієнтів мали досвід участі у бойових діях. Такий досвід був зафіксований у 123 осіб, що становить 83.1% вибірки, тоді як 25 пацієнтів, або 16.9%, не брали участі у бойових діях. Такий розподіл є очікуваним з огляду на специфіку досліджуваної вибірки та вказує на значну представленість бойового досвіду серед пацієнтів.

Проблеми зі сном також були поширеною характеристикою. Їх наявність зазначена у 113 пацієнтів, тобто у 76.4% вибірки. Відсутність проблем зі сном була зафіксована у 35 осіб, або у 23.6%. Висока частка пацієнтів із порушеннями сну є важливою для подальшого аналізу, оскільки сон може бути пов'язаний як з афективними симптомами, так і з когнітивними показниками.

Акубаротравма була наявна у 39 пацієнтів, що становить 26.4% вибірки. Водночас у 109 осіб, або 73.6%, акубаротравма не була зафіксована. Такий розподіл свідчить про те, що акубаротравма представлена меншою, але достатньо помітною частиною вибірки, що дозволяє розглядати її як окремий чинник у подальшому аналізі.

Вживання психоактивних речовин було зафіксовано у 23 пацієнтів,

або у 15.5% вибірки. У більшості пацієнтів, а саме у 125 осіб, що становить 84.5%, така характеристика була відсутня. Через відносно невелику частку позитивних випадків цей показник слід інтерпретувати обережно, особливо в моделях, де він може використовуватися як чинник або контрольна змінна.

Діагноз ПТСР був наявний у 67 пацієнтів, що відповідає 45.3% вибірки. Водночас у 81 пацієнта, або 54.7%, діагноз ПТСР не був зазначений. Такий розподіл показує, що вибірка містить як пацієнтів із вираженим посттравматичним профілем, так і пацієнтів без цього діагнозу, що створює підґрунтя для порівняльного аналізу психологічних і когнітивних показників.

Клінічні та травматологічні характеристики вибірки демонструють високу представленість чинників, пов'язаних із бойовим досвідом, порушеннями сну та посттравматичною симптоматикою. Водночас окремі чинники, зокрема акубаротравма та вживання психоактивних речовин, були наявні лише в частини пацієнтів. Це важливо враховувати під час подальшої інтерпретації результатів, оскільки нерівномірність розподілу окремих змінних може впливати на стабільність подальшого оцінювання їхнього потенційного ефекту.

2.5 Афективні та функціональні показники

Окремий блок дослідницького аналізу становили афективні та функціональні показники пацієнтів. До нього було включено показник симптомів посттравматичного стресового розладу за шкалою PCL, рівень тривоги та депресії за шкалою HADS, вираженість постконтузійної симптоматики за шкалою NSI, рівень психосоціального функціонування за PSP, а також показник PROMIS. Ці змінні дають змогу описати психологічний і функціональний профіль пацієнтів у різних клінічних категоріях.

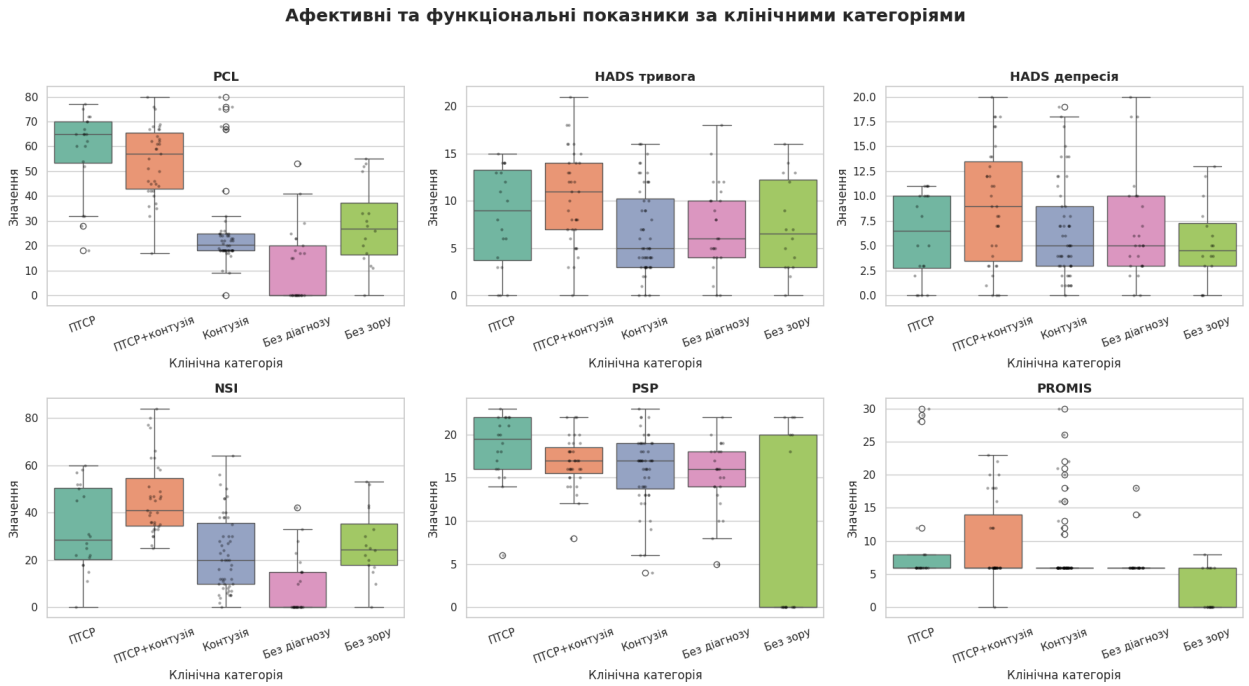


Рис. 2.4. Афективні та функціональні показники за клінічними категоріями

За даними рис. 2.4, показник симптомів ПТСР за шкалою PCL суттєво відрізнявся між клінічними категоріями. Найвищі значення спостерігалися у групах ПТСР та ПТСР з контузією, що відповідає їхньому клінічному профілю. У групі контузії значення PCL були помітно нижчими, хоча окремі пацієнти мали підвищені показники. У групі без діагнозу більшість значень була близькою до нуля, тоді як у групі без зору спостерігався ширший розподіл із помірними значеннями PCL.

Показники тривоги за шкалою HADS також демонстрували міжгрупову варіативність. Найвищі медіанні значення спостерігалися у групі ПТСР з контузією, тоді як у групах контузії, без діагнозу та без зору рівень тривоги був нижчим, але залишався неоднорідним. Наявність окремих високих значень у різних групах свідчить про те, що тривожна симптоматика не обмежується лише пацієнтами з формальним діагнозом ПТСР.

Розподіл депресивної симптоматики за шкалою HADS мав подібну структуру. Вищі значення були характерні для груп із ПТСР, особливо для пацієнтів із поєднанням ПТСР та контузії. Водночас у групах без діагнозу та без зору також спостерігалися окремі підвищені значення, що вказує на індивідуальну варіативність афективного стану в межах клінічних категорій.

Постконтузійна симптоматика за шкалою NSI була найбільш вираженою у групі ПТСР з контузією. Це може свідчити про поєднання травматологічного та посттравматичного компонентів у цій категорії пацієнтів. У групі контузії значення NSI також були підвищеними порівняно з групою без діагнозу, однак загалом мали нижчий рівень, ніж у групі ПТСР з

контузією.

Показник PSP, який характеризує рівень психосоціального функціонування, мав відносно високі значення у більшості клінічних категорій, але в окремих групах спостерігався широкий розкид. Особливо це помітно у групі без зору, де частина пацієнтів мала дуже низькі значення PSP. Це може бути пов'язано з неоднорідністю функціонального стану всередині цієї категорії.

Показник PROMIS також демонстрував значну варіативність між пацієнтами. У частини спостережень значення були сконцентровані на нижчому рівні, тоді як в окремих випадках фіксувалися вищі показники. Такий розподіл свідчить про неоднорідність загального психофізичного стану пацієнтів і потребує обережної інтерпретації при подальшому аналізі.

Загальний розподіл афективних та функціональних показників свідчить про те, що найбільш виражені посттравматичні та постконтузійні прояви спостерігаються у групах із ПТСР, особливо при поєднанні ПТСР з контузією. Водночас наявність розкиду та окремих високих значень у різних клінічних категоріях показує, що клінічна група не повністю визначає психологічний або функціональний стан пацієнта. Це обґрунтовує необхідність подальшого аналізу окремих чинників і їхнього потенційного впливу на результативні показники.

2.6 Когнітивні показники

Окремий напрям дослідницького аналізу становило вивчення когнітивних показників пацієнтів. До цього блоку було включено загальний когнітивний індекс RBANS, показник невербального інтелекту за тестом Равена, обсяг робочої пам'яті, показники виконання тестів TMT-A та TMT-B, а також результат виконання лабіринту. Ці змінні дозволяють оцінити загальний рівень когнітивного функціонування, увагу, швидкість обробки інформації, виконавчі функції та окремі аспекти інтелектуального профілю.

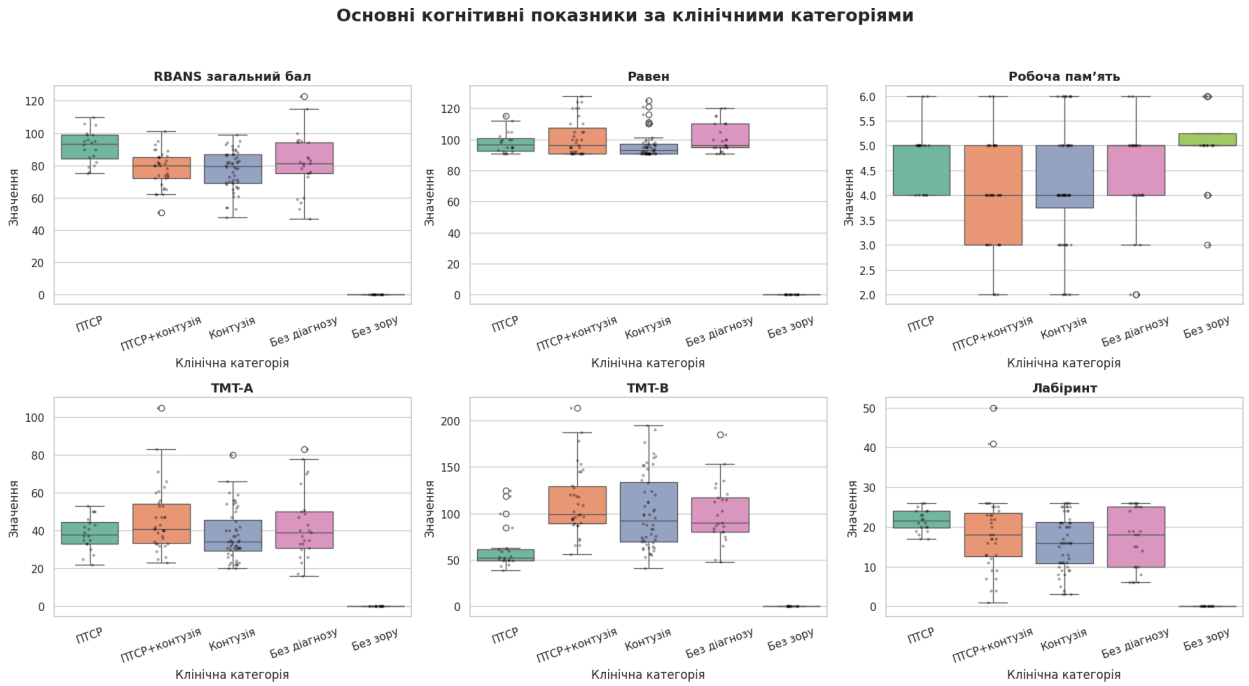


Рис. 2.5. Основні когнітивні показники за клінічними категоріями

Рисунок 2.5 показує, що когнітивні показники мали помітну варіативність як між клінічними категоріями, так і всередині окремих груп. Загальний індекс RBANS відображає інтегральну оцінку когнітивного функціонування. Його розподіл дозволяє порівняти клінічні категорії за загальним рівнем нейропсихологічних показників, однак наявність широкого розкиду значень свідчить про неоднорідність когнітивного стану пацієнтів у межах кожної групи.

Показник за тестом Равена використовувався як характеристика невербального інтелекту. Його розподіл між клінічними категоріями дає змогу оцінити, чи відрізняються групи за базовим рівнем інтелектуальних можливостей. Це важливо для подальшого аналізу, оскільки когнітивні результати можуть залежати не лише від клінічного стану або травматичного досвіду, а й від вихідного рівня інтелектуального функціонування.

Обсяг робочої пам'яті відображає здатність утримувати та опрацьовувати інформацію протягом короткого часу. Цей показник є важливим для оцінки когнітивної продуктивності, оскільки робоча пам'ять пов'язана з увагою, навчанням, плануванням і виконанням складних завдань.

Тести TMT-A та TMT-B характеризують різні аспекти уваги, швидкості обробки інформації та виконавчого контролю. TMT-A переважно пов'язаний із швидкістю зорового пошуку та базовою увагою, тоді як TMT-B більшою мірою відображає здатність до переключення уваги та когнітивної гнучкості. Оскільки в цих тестах результат зазвичай вимірюється часом виконання, вищі значення можуть вказувати на повільніше виконання завдання.

Показник виконання лабіринту використовувався як один із індикаторів виконавчих функцій. Він може бути пов'язаний із плануванням, контролем дій, просторовою орієнтацією та здатністю утримувати стратегію виконання завдання.

Окремо було проаналізовано домени RBANS, які відображають різні компоненти когнітивного функціонування: негайну пам'ять, зорово-просторові функції, мовні функції, увагу, відстрочену пам'ять та зорову пам'ять. Такий поділ є важливим, оскільки загальний когнітивний індекс може приховувати відмінності між окремими когнітивними сферами.

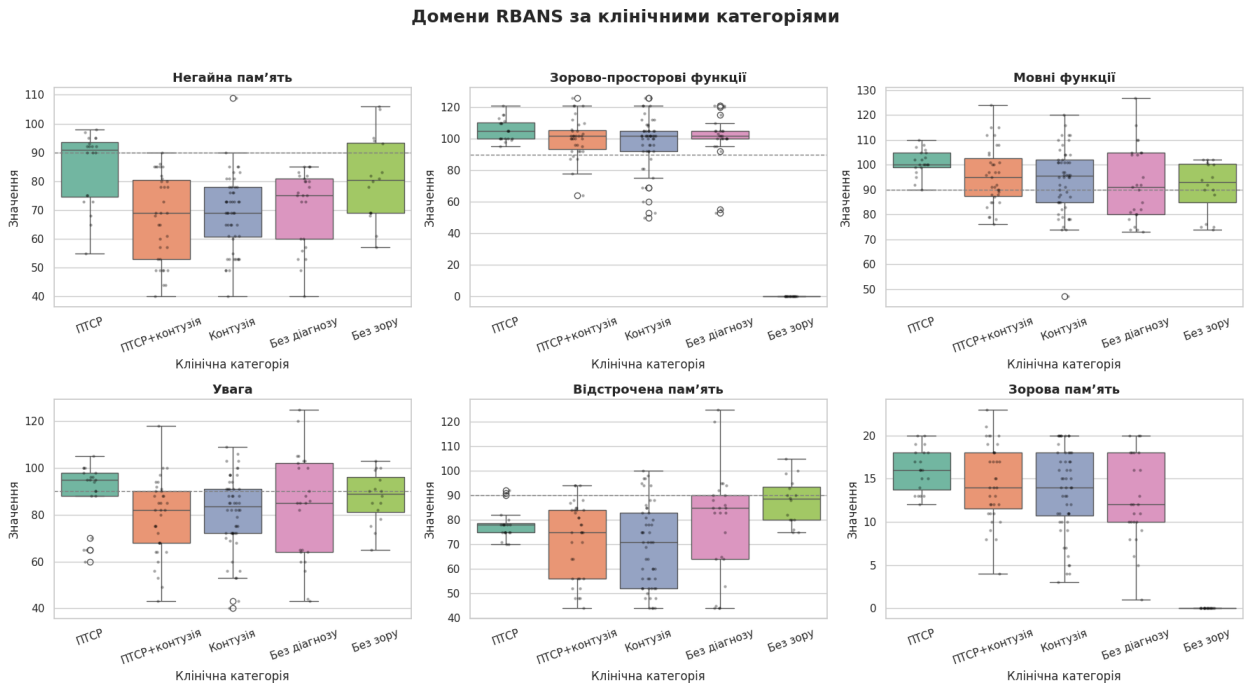


Рис. 2.6. Домени RBANS за клінічними категоріями

За даними рис. 2.6, для більшості доменів RBANS спостерігалася помітна міжгрупова та внутрішньогрупова варіативність. Пунктирна горизонтальна лінія на рівні 90 позначає орієнтовну нижню межу вікової норми для стандартних показників RBANS. У домені негайної пам'яті нижчі значення частіше спостерігалися у групах ПТСР з контузією, контузії та без діагнозу, тоді як у групах ПТСР і без зору частина пацієнтів мала показники, близькі до нормативного рівня.

Зорово-просторові функції загалом мали вищі значення порівняно з іншими доменами RBANS. У більшості груп медіанні значення перебували вище рівня 90, що може свідчити про відносно збережене функціонування цієї когнітивної сфери. Водночас у групі без зору цей показник не може бути інтерпретований аналогічно до інших груп, оскільки сенсорні обмеження впливають на можливість виконання відповідних завдань.

Мовні функції також демонстрували відносно широкий розподіл зна-

чень. У частини пацієнтів показники перебували нижче умовної межі норми, особливо у групах ПТСР з контузією, контузії та без діагнозу. Це свідчить про те, що мовний компонент когнітивного функціонування може бути неоднорідним навіть у межах однієї клінічної категорії.

Показник уваги мав значну варіативність між пацієнтами. У групі ПТСР медіанні значення були близькими до нормативного рівня, тоді як у групах ПТСР з контузією та контузії спостерігалися нижчі значення та ширший розкид. Це може вказувати на можливі труднощі уваги у пацієнтів із травматологічним компонентом, однак для остаточної інтерпретації потрібне подальше статистичне оцінювання.

Відстрочена пам'ять у більшості груп мала нижчі значення порівняно із зорово-просторовими та мовними функціями. Найбільша варіативність спостерігалася у групах ПТСР з контузією, контузії та без діагнозу. Це свідчить про те, що пам'яттєві процеси можуть бути однією з найбільш неоднорідних когнітивних сфер у досліджуваній вибірці.

Показник зорової пам'яті подано окремо від інших доменів, оскільки він має іншу шкалу вимірювання. У групі без зору значення цього показника були фактично відсутніми або нульовими, що є очікуваним з огляду на сенсорні обмеження пацієнтів. Тому порівняння цієї групи з іншими за зоровою пам'яттю потребує особливо обережної інтерпретації.

Розподіл доменів RBANS показує, що когнітивний профіль пацієнтів не є однорідним. Окремі когнітивні сфери можуть бути збереженими, тоді як інші демонструють зниження або значну варіативність. Це обґрунтовує необхідність аналізувати не лише загальний індекс RBANS, а й окремі домени, особливо в контексті клінічної категорії, травматологічного анамнезу та сенсорних обмежень.

2.7 Асоціативний аналіз соціального профілю клінічних категорій

Для узагальнення соціальних характеристик пацієнтів було побудовано інтегральний соціальний профіль клінічних категорій. На цьому етапі враховувалися як бінарні соціальні показники, зокрема наявність сім'ї, дітей, друзів або колег і життєвих планів, так і кількісні оцінки стосунків із партнером та соціальної підтримки. Такий підхід дозволяє не лише описати окремі соціальні змінні, а й порівняти загальний соціальний контекст різних клінічних груп.

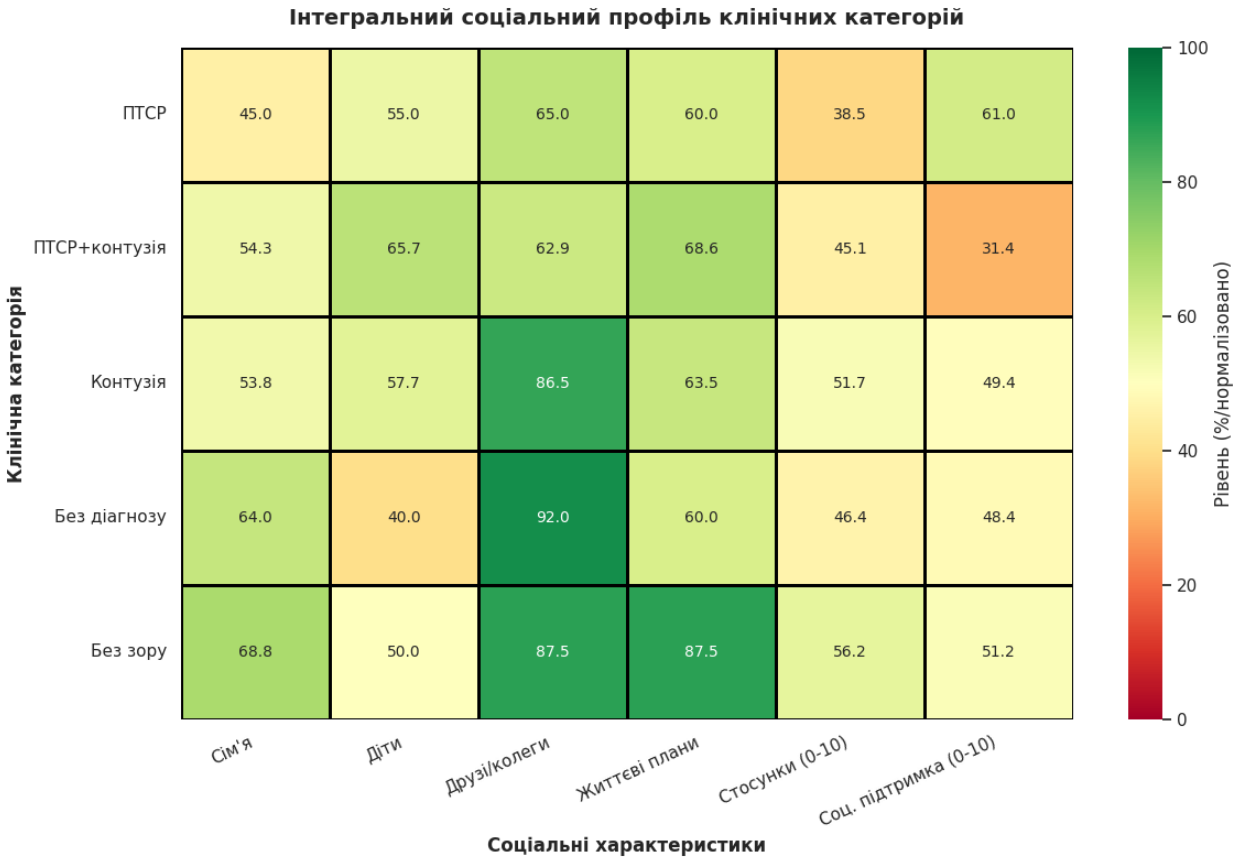


Рис. 2.7. Інтегральний соціальний профіль клінічних категорій

Рисунок 2.7 демонструє, що соціальний профіль клінічних категорій є неоднорідним. Найбільш помітні відмінності спостерігаються за показниками наявності друзів або колег, життєвих планів, суб'єктивної соціальної підтримки та оцінки стосунків із партнером. Зокрема, групи з ПТСР та ПТСР з контузією мають нижчі значення за окремими показниками соціального включення порівняно з групами без діагнозу, контузією та без зору.

Окрему увагу становить суб'єктивна соціальна підтримка, оскільки вона може бути пов'язана не лише з формальною наявністю соціального оточення, а й з тим, як пацієнт оцінює доступність допомоги та підтримки. Це важливо для подальшого аналізу, оскільки соціальні характеристики можуть виступати потенційними чинниками, пов'язаними з депресивною симптоматикою, тривогою та рівнем психосоціального функціонування.

Водночас результати, наведені на рис. 2.7, мають описовий характер. Вони демонструють відмінності соціального профілю між клінічними категоріями, але самі по собі не дозволяють зробити причинно-наслідкові висновки. Для оцінювання того, чи мають окремі соціальні характеристики статистично підтриманий потенційний ефект на психологічні або функціональні показники, у подальшому розділі застосовується метод Double Machine Learning.

2.8 Аналіз повторного тестування

Окремим етапом дослідницького аналізу було вивчення динаміки показників у пацієнтів, для яких були доступні результати повторного тестування. Для цього було розраховано різницю між другим і першим вимірюванням:

$$\Delta = T2 - T1. \tag{2.1}$$

Додатні значення Δ свідчать про зростання відповідного показника при повторному тестуванні, тоді як від'ємні значення вказують на його зниження. Інтерпретація напряму змін залежить від змісту конкретної шкали: для RBANS і PSP зростання можна розглядати як сприятливу динаміку, тоді як для HADS, TMT-A і TMT-B зниження показника відображає зменшення симптоматики або швидше виконання завдання.

Повторні вимірювання були доступні не для всієї вибірки, а лише для частини пацієнтів. До аналізу динаміки увійшли три клінічні категорії: без діагнозу, контузія та ПТСР з контузією. Розмір підгруп був невеликим: 5 пацієнтів у групі без діагнозу, 10 пацієнтів у групі контузії та 5 пацієнтів у групі ПТСР з контузією. У зв'язку з цим результати повторного тестування мають описовий характер і потребують обережної інтерпретації.

Табл. 2.1. Середня динаміка показників повторного тестування за клінічними категоріями, $\Delta = T2 - T1$

Показник	**Без діагнозу**	**Контузія**	**ПТСР+контузія**
RBANS	-11.6	4.4	5.0
HADS тривога	-7.0	-5.8	1.6
HADS депресія	-3.2	-3.8	1.2
PSP	2.8	0.7	2.0
TMT-A	-4.4	-5.8	-6.0
TMT-B	-8.8	-24.2	-74.6
Лабіринт	0.2	3.2	0.4
Робоча пам'ять	0.0	-0.8	0.4

Примітка: розмір підгруп становив $n = 5$ для групи без діагнозу, $n = 10$ для групи контузії та $n = 5$ для групи ПТСР з контузією.

За даними табл. 2.1, динаміка загального когнітивного індексу RBANS відрізнялася між групами. У групах контузії та ПТСР з контузією було

зафіксовано помірне зростання цього показника: середня зміна становила відповідно +4.4 та +5.0 бала. Натомість у групі без діагнозу спостерігалося зниження RBANS у середньому на −11.6 бала. Через малий розмір підгруп цей результат не слід інтерпретувати як остаточну закономірність, однак він вказує на різний характер когнітивної динаміки між клінічними категоріями.

Афективні показники продемонстрували різні напрями змін залежно від клінічної категорії. У групі без діагнозу рівень тривоги за HADS зменшився в середньому на −7.0 бала, а рівень депресії – на −3.2 бала. У групі контузії також спостерігалося зниження афективної симптоматики: тривога зменшилася на −5.8 бала, а депресія – на −3.8 бала. На відміну від цих двох груп, у групі ПТСР з контузією було зафіксовано зростання афективних показників: тривога збільшилася в середньому на +1.6 бала, а депресія – на +1.2 бала. Це може свідчити про більш несприятливу короткострокову емоційну динаміку в пацієнтів із поєднанням посттравматичної симптоматики та контузії.

Показники уваги та швидкості виконання завдань мали переважно позитивну динаміку. Для TMT-A у всіх трьох групах спостерігалося зменшення часу виконання: у групі без діагнозу на −4.4 секунди, у групі контузії на −5.8 секунди, у групі ПТСР з контузією на −6.0 секунд. Більш виражені зміни було отримано для TMT-B, який пов'язаний із переключенням уваги та когнітивною гнучкістю. Найбільше покращення спостерігалося у групі ПТСР з контузією, де середня зміна становила −74.6 секунди. У групі контузії середнє зниження часу виконання дорівнювало −24.2 секунди, тоді як у групі без діагнозу – лише −8.8 секунди.

Рівень психосоціального функціонування за PSP покращився в усіх трьох групах. Найбільше середнє зростання було зафіксовано у групі без діагнозу (+2.8), дещо менше – у групі ПТСР з контузією (+2.0), а найменше – у групі контузії (+0.7). Така динаміка свідчить про певне покращення функціонального стану, хоча її вираженість відрізнялася між клінічними категоріями.

Показники виконавчих функцій за лабіринтом та робочої пам'яті залишалися відносно стабільними. Для лабіринту найпомітніше середнє зростання спостерігалося у групі контузії (+3.2), тоді як у групах без діагнозу та ПТСР з контузією зміни були мінімальними: відповідно +0.2 та +0.4. Робоча пам'ять також не продемонструвала вираженої узгодженої динаміки: у групі без діагнозу середня зміна становила 0.0, у групі контузії −0.8, а у групі ПТСР з контузією +0.4.

Описовий аналіз повторного тестування вказує на те, що різні клінічні категорії могли мати неоднакові патерни змін. У групі контузії спостерігалося поєднання зниження тривоги й депресії з помірним покращенням RBANS та TMT-B, що може вказувати на відносно сприятливу короткострокову

динаміку. Група ПТСР з контузією демонструвала суперечливий профіль: з одного боку, було зафіксовано покращення RBANS і значне зменшення часу виконання ТМТ-В, з іншого – рівні тривоги й депресії зросли. У групі без діагнозу, навпаки, спостерігалося зниження афективної симптоматики та покращення PSP, але водночас зменшення загального когнітивного індексу RBANS. Такі результати мають розглядатися як описові тенденції, що потребують подальшої перевірки на більшій вибірці.

ВИСНОВКИ

У межах кваліфікаційної роботи було застосовано метод Double Machine Learning для аналізу медико-психологічних даних військовослужбовців із різним клінічним профілем. Основна увага приділялася не лише опису статистичних зв'язків між змінними, а й оцінюванню потенційних ефектів окремих клінічних, соціальних і демографічних чинників на психологічний, когнітивний та функціональний стан пацієнтів. Такий підхід був обраний з огляду на спостережувальний характер даних, де просте порівняння груп або кореляційний аналіз не завжди дозволяє відокремити вплив досліджуваного чинника від супутніх характеристик.

У теоретичній частині роботи було розглянуто основні положення причинно-наслідкового аналізу, проблему конфаундингу, відмінність між прогнозуванням і причинною інтерпретацією, контрфактичний підхід, а також метод Double Machine Learning. Було показано, що цей підхід є доцільним у задачах, де досліджуваний чинник і результативний показник можуть одночасно залежати від багатьох контрольних змінних. У практичній частині було проведено попередню обробку даних, описано основні групи змінних, виконано дослідницький аналіз вибірки, сформульовано низку дослідницьких гіпотез і побудовано моделі для оцінювання потенційних ефектів.

Першою було перевірено гіпотезу про те, що участь у бойових діях може бути пов'язана з вищим рівнем тривожної симптоматики. Ця гіпотеза отримала статистичну підтримку в межах обраної специфікації Double Machine Learning. За результатами моделювання участь у бойових діях мала позитивний ефект на показник тривоги за шкалою HADS: середній ефект становив $ATE = 2.35$ при 95% довірчому інтервалі [0.58; 4.12]. Оскільки довірчий інтервал не включав нуль, результат можна розглядати як статистично значущий. Це означає, що після контролю за віком, наявністю ПТСР і порушеннями сну бойовий досвід залишався пов'язаним із вищим рівнем тривожної симптоматики. Водночас цей результат слід інтерпретувати як оцінку потенційного ефекту за умови обраного набору контрольних змінних, а не як остаточне доведення причинності.

Друга гіпотеза стосувалася зв'язку між кількістю перенесених черепно-мозкових травм та вираженістю симптомів ПТСР. Вона також отримала статистичну підтримку. Було отримано позитивний ефект кількості ЧМТ на показник PCL: $ATE = 2.12$, 95% довірчий інтервал [0.41; 3.84]. Це свідчить про те, що більша кількість черепно-мозкових травм була пов'язана з вищою вираженістю симптомів ПТСР після врахування віку, бойового досвіду, порушень сну та часу після останньої травми. Отриманий результат є змістовно важливим, оскільки вказує на можливе значення накопичення

травматичних ушкоджень для психологічного стану військовослужбовців.

Окремо перевірялася гіпотеза про протективну роль соціальної підтримки щодо депресивної симптоматики. У межах Double Machine Learning статистично підтриманим виявився ефект соціальної підтримки на рівень депресії: $ATE = -0.43$, 95% довірчий інтервал $[-0.68;\ -0.17]$. Негативний напрям ефекту означає, що вищий рівень соціальної підтримки був пов'язаний із нижчим рівнем депресивної симптоматики після контролю за віком, бойовим досвідом, кількістю ЧМТ і порушеннями сну. У практичному сенсі це означає, що збільшення соціальної підтримки на один бал асоціювалося зі зменшенням показника депресії приблизно на 0.43 бала за шкалою HADS. Отже, соціальна підтримка може розглядатися як потенційний протективний чинник щодо депресивної симптоматики.

Разом із тим не всі початкові гіпотези отримали статистичну підтримку. Зокрема, гіпотеза про зв'язок між наявністю ПТСР і загальним когнітивним індексом RBANS не була статистично підтверджена. За результатами моделювання було отримано $ATE = 6.97$ при 95% довірчому інтервалі $[-1.91;\ 15.85]$. Оскільки довірчий інтервал включав нуль, немає достатніх підстав стверджувати про статистично значущий окремий ефект ПТСР на загальний когнітивний індекс у межах обраної моделі. Такий результат не означає, що ПТСР загалом не пов'язаний із когнітивним функціонуванням. Радше він свідчить про те, що за наявного обсягу вибірки та обраних контрольних змінних окремий внесок ПТСР у зміну загального індексу RBANS не був статистично підтверджений.

Також не отримала статистичної підтримки гіпотеза про окремий зв'язок акубаротравми з вираженістю постконтузійної симптоматики за шкалою NSI. На повній вибірці було отримано $ATE = -0.15$, 95% довірчий інтервал $[-7.66;\ 7.36]$, а в підгрупі пацієнтів із легкою ЧМТ $ATE = -1.39$, 95% довірчий інтервал $[-9.69;\ 6.91]$. В обох випадках довірчі інтервали були широкими і включали нуль. Це свідчить про значну невизначеність оцінки та про те, що за наявного обсягу даних неможливо надійно відокремити потенційний ефект акубаротравми від інших травматологічних характеристик, зокрема кількості ЧМТ або особливостей бойового досвіду.

Гіпотеза про негативний зв'язок порушень сну з когнітивним функціонуванням також не отримала статистичної підтримки, хоча напрям ефекту відповідав теоретичному очікуванню. Для впливу порушень сну на RBANS було отримано $ATE = -1.88$, 95% довірчий інтервал $[-6.62;\ 2.85]$. Це означає, що наявність порушень сну була пов'язана зі зниженням когнітивного індексу, однак у цій вибірці ефект не був статистично значущим. Подібний результат було отримано і для гіпотези щодо зв'язку наявності сім'ї з психосоціальним функціонуванням: $ATE = 0.97$, 95% довірчий інтервал $[-0.84;\ 2.77]$. Напрям ефекту був позитивним, але довірчий інтервал

включав нуль. Це дозволяє припустити, що сама по собі формальна наявність сім'ї не є достатнім показником соціального ресурсу; більш важливими можуть бути якість стосунків, реальний рівень підтримки та суб'єктивне сприйняття соціальної допомоги.

Окремим етапом роботи був аналіз динаміки показників у пацієнтів, які проходили повторне тестування. Повторні вимірювання були доступні для підвибірки з 20 пацієнтів, тому результати цього аналізу мають описовий характер. Порівняння значень $T2 - T1$ показало, що різні клінічні категорії могли мати неоднакові патерни змін. У групах із контузією та ПТСР з контузією спостерігалося помірне зростання загального когнітивного індексу RBANS, тоді як у групі без діагнозу було зафіксовано його зниження. За афективними показниками групи без діагнозу та з контузією демонстрували зменшення тривоги й депресії, тоді як у групі ПТСР з контузією, навпаки, спостерігалося зростання цих показників. Особливо помітним було зменшення часу виконання TMT-B у групі ПТСР з контузією, що може вказувати на покращення окремих аспектів когнітивної гнучкості. Однак через малий розмір підгруп ці результати слід розглядати лише як попередні тенденції.

Загалом результати роботи показали, що статистичну підтримку отримали три основні гіпотези: участь у бойових діях була пов'язана з вищим рівнем тривоги; більша кількість ЧМТ була пов'язана з вищою вираженістю симптомів ПТСР; соціальна підтримка була пов'язана з нижчим рівнем депресивної симптоматики. Інші гіпотези, зокрема щодо зв'язку ПТСР із когнітивним індексом, акубаротравми з постконтузійними симптомами, порушень сну з RBANS і наявності сім'ї з PSP, не отримали статистично значущого підтвердження. У частині випадків напрям ефекту був змістовно очікуваним, однак довірчі інтервали включали нуль, що не дозволяє вважати ці результати статистично підтриманими.

Практичне значення роботи полягає в тому, що застосований підхід дозволив виділити чинники, які можуть мати значення для оцінювання психологічного стану військовослужбовців. Зокрема, результати підкреслюють доцільність уважного скринінгу тривоги у пацієнтів із бойовим досвідом, урахування кількості перенесених ЧМТ при оцінюванні ризику ПТСР, а також посилення соціальної підтримки як потенційного напряму профілактики та зниження депресивної симптоматики. Водночас робота показала, що не всі клінічно очікувані зв'язки залишаються статистично значущими після контролю за супутніми змінними, що підтверджує доцільність використання причинно-наслідкового підходу в аналізі спостережувальних медико-психологічних даних.

Основними обмеженнями дослідження є відносно невеликий обсяг вибірки, нерівномірність клінічних груп, обмежена кількість повторних

вимірювань, а також ризик неврахованого конфаундингу. Оскільки дані мають спостережувальний характер, отримані результати не слід інтерпретувати як остаточне доведення причинності. Крім того, частина контрольних змінних може виступати не лише конфаундерами, а й медіаторами або наслідками досліджуваних чинників, що потребує додаткового каузального обґрунтування. Тому результати роботи доцільно розглядати як попередні оцінки потенційних ефектів, які потребують перевірки на більшій вибірці та з використанням розширеної каузальної специфікації.

Перспективи подальших досліджень полягають у розширенні вибірки, збільшенні кількості пацієнтів у кожній клінічній групі, уточненні змінних, що характеризують тяжкість травматичного досвіду, а також у ширшому використанні повторних вимірювань. Доцільним є подальше порівняння Double Machine Learning з іншими методами причинно-наслідкового аналізу, зокрема propensity score matching, causal forests та байєсівськими моделями. Окремої уваги потребує дослідження довгострокової динаміки стану пацієнтів, оскільки отримані результати свідчать, що психологічні та когнітивні показники можуть змінюватися з різною швидкістю і мати різні патерни відновлення залежно від клінічного профілю.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Gvozdenović, E., Malvisi, L., Cinconze, E., Vansteelandt, S., Nakanwagi, P., Aris, E., & Rosillon, D. (2021). *Causal inference concepts applied to three observational studies in the context of vaccine development: from theory to practice*. BMC Medical Research Methodology, 21, Article 35. https://doi.org/10.1186/s12874-021-01220-1
2. Shmueli, G. (2010). *To Explain or to Predict?* Statistical Science, 25(3), 289–310. https://doi.org/10.1214/10-STS330
3. Pearl, J. (2009). *Causal inference in statistics: An overview*. Statistics Surveys, 3, 96–146. https://doi.org/10.1214/09-SS057
4. Prosperi, M., Guo, Y., Sperrin, M., Koopman, J. S., Min, J. S., He, X., Rich, S., Wang, M., Buchan, I. E., & Bian, J. (2020). *Causal inference and counterfactual prediction in machine learning for actionable healthcare*. Nature Machine Intelligence, 2, 369–375. https://doi.org/10.1038/s42256-020-0197-y
5. Hammerton, G. L., & Munafò, M. R. (2021). *Causal inference with observational data: the need for triangulation of evidence*. Psychological Medicine, 51(4), 563–578. https://doi.org/10.1017/S0033291720005127
6. Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). *Double/debiased machine learning for treatment and structural parameters*. The Econometrics Journal, 21(1), C1–C68. https://doi.org/10.1111/ectj.12097
7. Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. https://doi.org/10.1017/CBO9781139025751