

ЕФЕКТИВНІСТЬ МУЛЬТИПАРАМЕТРИЧНОЇ УЛЬТРАЗВУКОВОЇ ДІАГНОСТИКИ З ПРОГРАМАМИ ШТУЧНОГО ІНТЕЛЕКТУ ПРИ ДИФУЗНИХ ЗАХВОРЮВАННЯХ ПЕЧІНКИ У ДІТЕЙ

Віталій О. Бабенко¹, Володимир В. Солодуценко^{1,2}, Марина Б. Діба¹, Марина В. Глоба², Борис А. Тарасюк¹

¹Державна установа «Всеукраїнський центр материнства та дитинства Національної академії медичних наук України», м. Київ, Україна

²Національний університет охорони здоров'я України імені П. Л. Шупика, м. Київ, Україна

Резюме

Вступ. Хронічні дифузні захворювання печінки у дітей потребують безпечної неінвазивної оцінки та повторюваного моніторингу фіброзу. Біопсія печінки залишається референтним методом морфологічної верифікації, однак у педіатричній практиці її застосування обмежують інвазивність, ризик ускладнень та етичні чинники. Мультипараметричне ультразвукове дослідження, зокрема еластографія зсувної хвилі та доплерографічна оцінка кровотоку, придатне для динамічного спостереження, але інтерпретація ультразвукових ознак залишається залежною від оператора. Автоматизований аналіз зображень із застосуванням алгоритмів штучного інтелекту може підвищити об'єктивність та відтворюваність стратифікації фіброзу і підтримати клінічні рішення.

Мета. Оцінити діагностичну точність мультипараметричного ультразвукового дослідження, що включає двовимірну еластографію зсувної хвилі та алгоритми машинного навчання, для стадіювання фіброзу печінки у дітей.

Матеріали та методи. Ретроспективне одноцентрове дослідження включило 185 дітей (140 з хронічними дифузними захворюваннями печінки та 45 дітей контрольної групи). Усім пацієнтам виконано двовимірну еластографію зсувної хвилі, доплерографію печінкових вен та аналіз сірошкальних ультразвукових зображень двома алгоритмами машинного навчання: модель А (алгоритм випадкового лісу, бінарна класифікація) та модель Б (ієрархічний ансамбль, багатокласова класифікація). Референтним підходом слугували клінічна верифікація та стратифікація ризику за результатами еластографії.

Результати. Виявлено виражений позитивний зв'язок між показниками еластографії та прогнозами моделі Б (коефіцієнт рангової кореляції Спірмена 0,77). Модель А продемонструвала специфічність 84 % при диференціації норми від патологічних змін. Модель Б мала вищу чутливість (83 %), однак виявляла тенденцію до завищення оцінки на межі між мінімальними та помірними проявами фіброзу. Доплерографічні патерни, зокрема монофазний спектр кровотоку у печінкових венах, достовірно асоціювалися з вищими стадіями фіброзу.

Висновки. Інтеграція алгоритмів машинного навчання в протокол ультразвукового дослідження підвищує об'єктивність діагностичної оцінки, проте ієрархічні підходи потребують калібрування для зниження частоти хибнопозитивних результатів у зоні ранніх стадій фіброзу.

Ключові слова: педіатрична гепатологія, хронічні дифузні захворювання печінки, фіброз печінки, мультипараметричне ультразвукове дослідження, еластографія зсувної хвилі, доплерографія печінкових вен, штучний інтелект, машинне навчання

ВСТУП

Хоча біопсія печінки залишається золотим стандартом верифікації хронічних дифузних захворювань печінки (ХДЗП), її рутинне застосування

в педіатрії обмежене інвазивністю, ризиком ускладнень та етичними застереженнями [1, 2]. Крім того, гетерогенність розподілу фіброзу може призводити до помилок вибірки. Це зумовлює необхідність впровадження надійних неінвазивних біомаркерів,

здатних забезпечити динамічний моніторинг без седації та втручання. Ці обмеження стимулювали активний розвиток неінвазивних методів діагностики.

Серед неінвазивних підходів вивчення структури печінки провідну роль посідає мультипараметрична ультразвукова діагностика (УЗД), зокрема еластографія зсувної хвилі (SWE) для кількісної оцінки жорсткості печінки [3, 4]. Висока кореляція результатів еластографії та гістологічного дослідження печінки у визначенні стадій фіброзу доведена в низці біопсійно контрольованих досліджень 2D-SWE, зокрема в когорті MASLD [5] та при хронічному гепатиті В [6]. Попри значний прогрес, чутливість і специфічність стандартного В-режиму УЗД для диференціації ранніх стадій фіброзу (F0–F2) залишаються обмеженими, а інтерпретація зображень залежить від досвіду оператора [7]. Це формує запит на об'єктивний аналіз ультразвукових даних.

Системи штучного інтелекту (ШІ) на основі машинного і глибинного навчання відкривають можливості для підвищення відтворюваності та точності діагностики, зокрема через аналіз текстурних патернів (радіоміку) ультразвукових зображень. За даними систематичних оглядів, діагностична точність для виявлення значущого фіброзу ($F \geq 2$) та цирозу в дослідженнях на дорослих пацієнтах досягала 80-90% [8]. Водночас більшість досліджень стосуються дорослих пацієнтів і потребують стандартизованої валідації [9]. У педіатрії доказів суттєво менше; це принципова прогалина, зважаючи на особливості дитячих когорт і акустичних властивостей тканин, які вимагають адаптованих моделей і протоколів [10].

МЕТА

Оцінити та порівняти діагностичну ефективність мультипараметричної УЗД із використанням алгоритмів штучного інтелекту та еластографії зсувної хвилі для неінвазивної класифікації стадій фіброзу печінки у дітей із дифузними захворюваннями печінки.

МАТЕРІАЛИ ТА МЕТОДИ

Проведено одноцентрове ретроспективне дослідження діагностичної точності. Дизайн та звітність узгоджено з рекомендаціями STARD 2015 для досліджень діагностичної точності та TRIPOD+AI для опису моделей машинного навчання [11, 12]. Дослідження схвалено локальним етичним комітетом ДУ «ВІЦМД НАМН України», від усіх учасників або їхніх законних представників отримано інформовану згоду.

Учасники дослідження. До дослідження було включено 185 дітей, які проходили обстеження в період 2021-2024 рр., розділених на дві групи:

- **Основна група ($n = 140$),** середній вік $11 \pm 3,4$ роки, з клінічно встановленими (верифікованими за прийнятими клініко-лабораторними та інструментальними критеріями) хронічними дифузними захворюваннями печінки (ХДЗП), зокрема: автоімунний гепатит, хвороба Вільсона, хронічні вірусні гепатити В/С та метаболічно асоційована стеатотична хвороба печінки (MASLD) [13].

- **Контрольна група ($n = 45$),** середній вік $10 \pm 2,6$ років, без клінічних та анамнестичних ознак захворювання печінки, прийому гепатотоксичних препаратів чи ожиріння.

Критерії включення: вік від 5 до 18 років; встановлений діагноз ХДЗП (для основної групи); виконане УЗД; наявність попередніх результатів вимірів жорсткості печінки.

Критерії виключення: гострий гепатит або загострення хронічного (активність АЛТ/АСТ $> 5 \times$ від верхньої межі норми), ознаки холестазу, недотримання протоколу підготовки до УЗД (порушення режиму голодування), невалідні вимірювання при проведенні еластографії ($IQR/median > 30\%$). Критерії гармонізовано з рекомендаціями та протоколами WFUMB 2024 [3, 4].

Еталонний стандарт. Ступінь фіброзу печінки класифікували в категорії F0–F4 за аналогією до шкали METAVIR, використовуючи фіксовані порогові значення жорсткості печінки (кПа), визначені в локальному протоколі на основі опублікованих даних. У цьому дослідженні дані біопсії печінки були недоступні, тому ступінь фіброзу, визначений за SWE, розглядали як референтний метод для порівняння з моделями А та Б. У додатковому аналізі «норма vs. патологія» еталонним критерієм була належність до клінічної групи (ХДЗП проти контролю). Оцінку проводили два незалежних лікарі УЗД анонімно. У разі розбіжностей рішення приймалося консенсусно [14].

Індекс-тести. Усі інструментальні дослідження печінки проводили на УЗ-системах експертного класу Aplio 500/Aplio 800. УЗД проводилося згідно загальноприйнятого протоколу. В усіх випадках воно доповнювалося скануванням високочастотним (ВЧ) лінійним датчиком частотою від 5,3 до 17 МГц, з реєстрацією доплерівських спектрів кровотоку в печінкових венах.

- **Еластографія зсувної хвилі (2D-SWE):** виконувалась згідно з протоколом WFUMB [3]. Після мінімум 4-годинного голодування проводили не менше 10 валідних вимірювань жорсткості печінки у кПа. Результат вважався надійним при співвідношенні міжквартильного розмаху до медіани ($IQR/median \leq 30\%$).

- **В-режим УЗД для аналізу ШІ:** для кожного пацієнта зберігали 5-10 нативних зображень паренхіми правої частки печінки в сагітальній та

поперечній проєкціях. Зображення архівували без стиснення у вихідній роздільній здатності (800×608 пікселів). З цих зображень двома експертами було розмічено 582 області інтересу (ROI) з подальшим узгодженням у разі розбіжностей: 312 у контрольній групі та 270 у пацієнтів з ХДЗП.

Додатково у частини пацієнтів оцінювали ступінь стеатозу печінки за контрольованим параметром ослаблення (controlled attenuation parameter, CAP) [15]. Значення CAP наводилися у децибелах на метр (дБ/м) у діапазоні 100-400 дБ/м з подальшою градацією ступеня стеатозу: S0 – стеатоз <10% гепатоцитів, S1 – $\geq 10\%$, S2 – $\geq 33\%$, S3 – $\geq 66\%$ гепатоцитів.

Інтерпретацію лабораторних та клінічних даних проводили згідно з рекомендаціями NASPGHAN щодо неінвазивної оцінки фіброзу та стеатозу у дітей [16].

Розробка та валідація моделей III. Для класифікації УЗ-зображень було використано дві моделі машинного навчання, розроблені раніше нашою командою [17, 18]. Навчання моделей проводилося з використанням перехресної перевірки (k-fold cross-validation).

- Модель А (Текстурний класифікатор): Random Forest. Вхідний вектор ознак: статистики першого порядку (середнє арифметичне, дисперсія, коефіцієнт асиметрії) та текстурні дескриптори Хараліка (на основі матриці суміжності відтінків сірого), розраховані для ROI. Цільова змінна: бінарна мітка (Норма vs. Патологія).

- Модель Б (Ієрархічний класифікатор): каскадний ансамбль бінарних класифікаторів. Архітектура передбачає послідовне відсікання класів: спочатку відділення F0 від $F \geq 1$, потім F1 від $F \geq 2$ тощо. Як еталонне значення (ground truth) для навчання використовували стадіювання за шкалою METAVIR, отримане методом 2D-SWE.

Весь набір даних було розділено на навчальну (80 %), валідаційну (10 %) та відкладену тестову

(10 %) вибірки на рівні пацієнтів, щоб уникнути перетікання даних. Підбір гіперпараметрів моделей виконувався виключно на валідаційній вибірці. Фінальна оцінка ефективності проводилася один раз на тестовій вибірці, яка не брала участі у навчанні.

Статистичний аналіз. Для основного завдання (класифікація стадій фіброзу F0–F4) оцінювали узгодженість між SWE та моделями А і Б. Аналізували:

- середній ступінь фіброзу за кожною нозологією,
- розподіл пацієнтів за категоріями F0–F4,
- частку пацієнтів, яких модель Б «перекласифікувала» з F0–F1 (за SWE) у F2–F4.

Узгодженість кількісних оцінок (середній ступінь фіброзу, частка зрізів з «патологією» за моделлю А) з показниками SWE оцінювали за коефіцієнтом кореляції Спірмена (ρ) з розрахунком 95 % довірчих інтервалів.

У додатковому аналізі «норма vs. патологія» референтним критерієм була належність до клінічної групи (основна група з ХДЗП проти контрольної групи). Для кожного методу (SWE, модель А, модель Б) розраховували частоту виявлення патології (detection rate) та показники чутливості, специфічності, прогностичної цінності позитивного і негативного результатів щодо виявлення ХДЗП при пороговому значенні $F \geq 1$ (F0 проти F1–F4).

Статистичний аналіз виконували мовою Python v3.9 із використанням бібліотек scikit-learn та statsmodels. Рівень статистичної значущості встановлювали як $p < 0,05$ (двобічні тести).

РЕЗУЛЬТАТИ

Характеристика когорти. До аналізу було включено 185 дітей. Основні демографічні та клінічні характеристики представлено в таблиці 1.

Таблиця 1

Базові характеристики учасників дослідження

Характеристика	Основна група (n = 140)	Контрольна група (n = 45)	Уся когорта (n = 185)
Вік, років (середнє $\pm$ СВ)	11,0 $\pm$ 3,4	10,0 $\pm$ 2,6	10,8 $\pm$ 3,3
Стать (хлопчики / дівчата), n (%)	55 (39,3%) хлопчиків 85 (60,7%) дівчат	13 хлопчиків (28,8%) 32 дівчини (71,2%)	68 хлопчиків (36,8%) 117 дівчат (63,2%)
Розподіл за нозологіями, n (%)	-	-	-
MASLD	65 (46,4%)	-	65 (35,1%)
Автоімунний гепатит (АІГ)	23 (16,4%)	-	23 (12,4%)
Хронічний гепатит С (ХГС)	11 (7,9%)	-	11 (5,9%)
Хронічний гепатит В (ХГВ)	16 (11,4%)	-	16 (8,6%)
Хвороба Вільсона	25 (17,9%)	-	25 (13,5%)
Розподіл стадій фіброзу за еластографією (METAVIR), n (%)	-	-	-
F0	51 (27,6%)	21 (46,7%)	72 (38,9%)
F1	45 (24,3%)	19 (42,2%)	64 (34,6%)
F2	22 (11,9%)	5 (11,1%)	27 (14,6%)
F3	17 (9,2%)	0 (0,0%)	17 (9,2%)
F4	5 (2,7%)	0 (0,0%)	5 (2,7%)

За даними еластографії, значущий фіброз (стадії $F \geq 2$) було виявлено у 49 (26,5 %) пацієнтів, а поширений фіброз ($F \geq 3$) – у 22 (11,9 %) пацієнтів.

Результати моделі А. Після обчислення текстурних ознак на виділених ROI їх подавали на вхід моделі А, завдяки чому була виконана попередня діагностика текстури паренхіми на ознаки того, чи є вона незмінною та чи має патологічні ознаки.

Узагальнені результати бінарної класифікації «норма vs. патологія» для еластографії, моделі А та моделі Б наведено в табл. 2.

Результати моделі Б. Так само були розраховані текстурні ознаки, які подавали на вхід ієрархічної моделі Б для визначення стадії фіброзу. Розподіл прогнозованих стадій фіброзу за моделлю Б наведено в табл. 3.

Таблиця 2

Діагностична ефективність «норма vs. патологія»

Метод (індекс-тест)	Чутливість	Специфічність	Точність
Еластографія	0,61	0,47	0,59
Моделі А (бінарна)	0,59	0,84	0,67
Моделі Б (бінаризовано $F \geq 1$)	0,83	0,73	0,83

Таблиця 3

Порівняння сумарного розподілу стадій фіброзу: еластографія vs. модель Б

Стадія	Еластографія, n (%)	Моделі Б, n (%)
F0	72 (38,9%)	53 (28,6%)
F1	64 (34,6%)	51 (27,6%)
F2	27 (14,6%)	58 (31,4%)
F3	17 (9,2%)	17 (9,2%)
F4	5 (2,7%)	6 (3,2%)
Разом	185 (100%)	185 (100%)

На рівні агрегованих нозологічних груп ($n = 6$) спостерігалася тенденція до позитивного зв'язку між середніми показниками жорсткості та середнім прогнозом моделі. З огляду на малу кількість груп цей аналіз має описовий характер і потребує підтвердження на індивідуальному рівні та/або в більшій мультицентровій вибірці.

Ефективність класифікації «норма vs. патологія» різними методами. У додатковому аналізі було оцінено здатність кожного методу розрізняти дітей із ХДЗП (основна група) та здорових дітей (контрольна група). Для цього як еталон використовували належність до групи (контроль vs ХДЗП), а як тест – класифікацію за порогом $F \geq 1$ (F0 проти F1–F4). У табл. 2 наведені показники діагностичної ефективності (чутливість, специфічність, точність) для еластографії, моделі А та моделі Б у задачі «норма vs патологія» (патологія = стадія ≥ 1). У цьому аналізі еталонним критерієм була належність пацієнта до основної групи з ХДЗП або до контрольної групи, тоді як еластографія, модель А та модель Б розглядалися як індекс-тести (патологія = $F \geq 1$). Цей аналіз розглядав окрему клінічну задачу – виявлення будь-якого фіброзу ($F \geq 1$) як маркера наявності ХДЗП, а не точну стратифікацію стадій F0–F4. За таких умов еластографія продемонструвала чутливість 64 %, специфічність 47 % та прогностичну цінність позитивного результату 78 % щодо виявлення дітей з ХДЗП. Модель А мала близьку чутливість

(59 %), але суттєво вищу специфічність (84 %), формуючи більшу групу «норма». Натомість модель Б виявилася більш «агресивною»: її чутливість становила 83 % при специфічності 73 %, тобто вона частіше відносила пацієнтів до категорії «патологія».

Це свідчить про різні порогові властивості підходів: модель Б має більш «агресивний» профіль і схильна ширше виявляти потенційні патологічні зміни, тоді як модель А, ймовірно, розглядає частину випадків зі стадією F1 як клінічно малозначущі.

Розподіл стадій фіброзу за еластографією та моделлю Б. Порівняння розподілу стадій між еластографією та моделлю Б наведено в табл. 3. Отримані дані свідчать про тенденцію моделі Б відносити частину пацієнтів із ранніх стадій до категорії помірного фіброзу, що може відображати підвищену чутливість до ранніх текстурних змін, але потребує калібрування для зниження ризику хибнопозитивних висновків на межі ранніх стадій.

Узгодженість із доплерівськими показниками. Для оцінки клінічної узгодженості результатів було проаналізовано зв'язок між доплерівськими характеристиками кровотоку в печінкових венах (звичайного калібру у частках печінки та дрібного калібру при ВЧ скануванні) (табл. 4) і показниками фіброзу та класифікацією моделей.

У табл. 5 наведено кореляційний аналіз між часткою спектрів та показниками фіброзу.

Таблиця 4

Розподіл доплерографічних характеристик за нозологією

Нозологія	Стандарт			ВЧ УЗД		
	Монофазний спектр	Біфазний спектр	Трифазний спектр	Монофазний спектр	Біфазний спектр	Трифазний спектр
Контрольна група (n = 45)	0 (0,0 %)	5 (11,1 %)	40 (88,9 %)	0 (0,0 %)	6 (13,3 %)	39 (86,7 %)
MASLD (n = 65)	0 (0,0 %)	14 (21,5 %)	51 (78,5 %)	3 (4,6 %)	20 (30,8 %)	42 (64,6 %)
АІГ (n=23)	3 (13,0 %)	7 (30,4 %)	13 (56,5 %)	4 (17,4 %)	10 (43,5 %)	9 (39,1 %)
ХГС (n = 11)	0 (0,0 %)	1 (9,1 %)	10 (90,9 %)	0 (0,0 %)	4 (36,4 %)	7 (63,6 %)
ХГВ (n=16)	3 (18,8 %)	6 (37,5 %)	7 (43,8 %)	3 (18,8 %)	8 (50,0 %)	5 (31,3 %)
Хв. Вільсона (n = 25)	2 (8,0 %)	8 (32,0 %)	15 (60,0 %)	2 (8,0 %)	9 (36,0 %)	7 (28,0 %)

Таблиця 5

Аналіз кореляції між типами спектрів у печінкових венах та показниками фіброзу у печінці

Показник доплерографії	Кореляція (ρ) з середнім фіброзом (еластографія)	Кореляція (ρ) з % «патології» (модель А)	Кореляція (ρ) з середнім фіброзом (модель Б)
Монофазний спектр (стандарт)	0,70	0,88	0,88
Біфазний спектр (стандарт)	0,83	0,77	0,94
Трифазний спектр (стандарт)	-0,77	-0,89	-0,89
Монофазний спектр (ВЧ УЗД)	0,75	0,93	0,93
Біфазний спектр (ВЧ УЗД)	0,31	0,77	0,71
Трифазний спектр (ВЧ УЗД)	-0,77	-0,71	-0,89

Примітка. Використано коефіцієнт кореляції Спірмена (ρ).

Збільшення частки монофазних та біфазних спектрів на рівні нозологічних груп асоціювалося з вищим середнім ступенем фіброзу за еластографією та моделлю Б, а також з більшою часткою зрізів, класифікованих моделлю А як «патологія». Натомість вищий відсоток трифазних спектрів був пов'язаний із нижчими значеннями цих показників (табл. 5). Враховуючи невелику кількість груп (n = 6), отримані кореляції слід розглядати як гіпотезоутворюючі та такі, що потребують підтвердження в більшій когорті.

ДИСКУСІЯ

Проведене дослідження продемонструвало, що характеристика кровотоку в печінкових венах за доплерографією тісно пов'язана з вираженістю фіброзних змін у печінці у дітей з хронічними дифузними захворюваннями печінки. Зі збільшенням частки монофазних та біфазних спектрів у нозологічних групах зростали середні значення фіброзу за еластографією та за моделлю Б, а також частка зрізів, які модель А класифікувала як «патологія». Натомість вищий відсоток трифазних спектрів був асоційований з низьким середнім ступенем фіброзу та меншою часткою патологічних змін за результатами автоматизованого аналізу.

Зв'язок між спектральними характеристиками кровотоку та фіброзом має патофізіологічне підґрунтя. Прогресування фіброзу супроводжується перебудовою архітекτονіки синусоїдів та підвищенням внутрішньопечінкового судинного опору, що призводить до сплюснення доплерівського спектра (втрата трифазності). Реєстрація монофазних

спектрів при ВЧ-УЗД у пацієнтів з MASLD, які мали нормальний патерн при конвексному скануванні, свідчить про вищу роздільну здатність ВЧ-датчиків у детекції мікроциркуляторних порушень.

У нашій когорті саме групи з більш тяжкими формами ХДЗП (автоімунний гепатит, хвороба Вільсона, хронічний гепатит В) демонстрували вищу частку бі- та монофазних спектрів і вищий середній ступінь фіброзу за еластографією та моделлю Б, тоді як контрольна група та пацієнти з MASLD частіше мали трифазний спектр та низькі значення фіброзу. Водночас через невелику кількість агрегованих груп (n = 6) ці зв'язки слід розглядати як гіпотезоутворюючі, що потребують підтвердження в більших вибірках.

Використання високочастотного датчика дозволило підвищити чутливість доплерографії до ранніх змін гемодинаміки. Зокрема, у групі MASLD при стандартному датчику монофазні спектри не реєструвалися, тоді як при ВЧ-УЗД їх було виявлено у невеликої, але помітної частки пацієнтів (≈5 %). Це може свідчити про те, що високочастотне сканування здатне виявляти початкові порушення венозного відтоку та мікроциркуляції, які ще не призводять до вираженого підвищення жорсткості паренхіми та грубих змін спектра при стандартному дослідженні. Практично це підтримує доцільність використання ВЧ-датчика як уточнювального методу у дітей з груп ризику, особливо коли стандартне УЗД не виявляє відхилень, але клінічні чи лабораторні дані викликають підозру на фіброз.

Поєднання доплерографії печінкових вен та еластографії продемонструвало комплементарний

характер інформації. У випадках, коли монофазний або виражено біфазний спектр поєднувався з підвищеними значеннями жорсткості печінки, ймовірність фіброзного ураження була високою і клінічно правдоподібною. Натомість, за наявності трифазного спектра на фоні прикордонних значень еластографії результати слід інтерпретувати обережно, з урахуванням можливого впливу запалення, стеатозу або технічних чинників. Це узгоджується з сучасними рекомендаціями, де еластографія розглядається як частина мультипараметричного підходу, а не як ізольований «золотий стандарт».

Автоматизований аналіз ультразвукових зображень за допомогою моделей А та Б в нашому дослідженні продемонстрував різні профілі роботи, що важливо для їх потенційного клінічного застосування. Модель А, яка базується на текстурних ознаках та класифікації «норма / патологія», показала відносно високу специфічність (0,84) при помірній чутливості (0,59). Це вказує на її більш «консервативний» характер: модель схильна інтерпретувати частину випадків зі слабо вираженими змінами (зокрема F1) як варіант норми або мінімальних відхилень. Такий профіль може бути корисним для підтвердження відсутності значущої патології у пацієнтів з низьким ризиком, коли важливіше уникнути хибнопозитивних висновків.

Натомість модель Б, що реалізує ієрархічну мультикласову класифікацію стадій фіброзу, продемонструвала вищу чутливість (0,83) і дещо нижчу специфічність (0,73) у задачі «норма vs. патологія». Це відповідає більш «агресивному» профілю: модель схильна частіше відносити пацієнтів до категорії з наявністю фіброзу, що підтверджується перерозподілом частини випадків, які за еластографією належали до F0–F1, у стадію F2. З одного боку, це може відображати підвищену чутливість до тонких текстурних змін, пов'язаних з раннім фіброзом; з іншого – створює ризик гіпердіагностики помірного фіброзу і потребує обережного трактування в клінічному контексті.

Отримані результати загалом узгоджуються з даними сучасної літератури про те, що радіоміка УЗ-зображень та моделі машинного навчання здатні виявляти зміни паренхіми печінки, які частково випереджають традиційні візуальні ознаки та доплерографічні критерії. У нашій роботі це проявилось через сильні кореляції між часткою патологічних класифікацій за моделлю А, середнім ступенем фіброзу за моделлю Б та змінами спектрів печінкових вен. Таким чином, поєднання мультипараметричної УЗД (В-режим, еластографія, доплерографія) з автоматизованим аналізом зображень може сформувати основу для більш об'єктивної та відтворюваної оцінки фіброзу печінки у педіатричній популяції.

ВИСНОВКИ

1. Еластографія, модель А та модель Б демонструють тісний взаємозв'язок у оцінці стану паренхіми печінки. На рівні нозологічних груп спостерігалася сильна позитивна кореляція між середнім ступенем фіброзу за еластографією та оцінками моделі Б ($p \approx 0,77$), а також між еластографією та часткою зрізів, класифікованих моделлю А як «патологія» ($p \approx 0,77$), що свідчить про узгодженість напрямку змін між методами.

2. Модель Б, яка реалізує ієрархічну класифікацію стадій фіброзу, має тенденцію класифікувати частину пацієнтів із F0–F1 (за еластографією) як F2. Це відображає її підвищену чутливість до ранніх текстурних змін, але супроводжується ризиком гіпердіагностики помірного фіброзу. У цьому контексті модель Б може розглядатися як інструмент підвищеної чутливості для первинної стратифікації ризику, однак її клінічне застосування для «виключення» клінічно значущого фіброзу потребує зовнішньої валідації та визначення робочих порогів на незалежних даних.

3. Модель А, орієнтована на бінарну класифікацію («норма vs. патологія»), продемонструвала чутливість 59 % та специфічність 84 % у задачі відокремлення дітей із ХДЗП від контрольної групи. Це свідчить про її потенціал як автоматизованого інструменту для виявлення патологічних змін у паренхімі печінки, особливо у поєднанні з клінічними та лабораторними даними. Водночас обмежена чутливість вимагає обережного використання моделі А як самостійного скринінгового методу.

4. Обробка ультразвукових зображень за допомогою моделей А і Б забезпечує найкращу узгодженість з еластографією при діагностиці відсутності фіброзу (F0) та вираженого фіброзу/цирозу (F3–F4). Розмежування ранніх стадій є складним як для кількісних ультразвукових методів, так і для алгоритмічної класифікації. Це є очікуваним обмеженням підходу та ключовою мішенню подальшої валідації/калібрування.

5. Поєднання мультипараметричної ультразвукової діагностики (В-режим, еластографія, доплерографія) з автоматизованим аналізом зображень на основі моделей А та Б має потенціал стати основою більш об'єктивного підходу до неінвазивної стратифікації фіброзу печінки у дітей, однак потребує подальшої валідації в більших, мультицентрових когортних дослідженнях.

Перспективи подальших досліджень.

У майбутньому наші дослідження мають бути спрямовані на проспективну багатоцентрову валідацію запропонованого підходу на незалежних вибірках

із різних ультразвукових систем, стандартизацію протоколів отримання зображень та калібрування моделей для зменшення хибнопозитивних результатів на межі ранніх стадій фіброзу. Перспективним є поєднання кількісних показників еластографії, доплерографії та ознак текстури зображення з клініко-лабораторними даними для підвищення точності стратифікації, а також впровадження інтерпретованих моделей для прозорого обґрунтування діагностичних рішень.

ДОТРИМАННЯ ЕТИЧНИХ НОРМ

Дослідження виконано відповідно до Гельсінської декларації. Протокол схвалено комісією з медичної етики Державної установи «Інститут ядерної медицини та променевої діагностики Національної академії медичних наук України» (протокол № 2 від 09 квітня 2020 р.).

ФІНАНСУВАННЯ ТА КОНФЛІКТ ІНТЕРЕСІВ

Автори заявляють про відсутність конфлікту інтересів. Дослідження виконано без зовнішнього фінансування.

Стаття є фрагментом науково-дослідної роботи Державної установи «Інститут ядерної медицини та променевої діагностики Національної академії медичних наук України»: «Розробити технологію

цифрової обробки променево-діагностичних зображень на основі елементів штучного інтелекту при дифузній патології печінки у хворих різних вікових груп» (номер державної реєстрації: 0120U000437). Дослідження розпочато на базі ДУ «ІЯМПД НАМН України». З 01 серпня 2024 року Інститут увійшов до складу ДУ «Всеукраїнський центр материнства та дитинства НАМН України», де роботу продовжено. Стаття містить фрагменти дисертаційної роботи здобувача наукового ступеня доктора філософії кафедри радіології Національного університету охорони здоров'я України імені П. Л. Шупика Солодущенко Володимира В'ячеславовича «Ультразвукова діагностика дифузних захворювань печінки з використанням моделей штучного інтелекту» (номер державної реєстрації: 0124U003002, прикладна, термін виконання: 2024-2028 рр., науковий керівник Глоба М. В.).

ВНЕСОК АВТОРІВ

Бабенко В. О.^{B, C, D}

Солодущенко В. В.^{A, B, C, D}

Діба М. Б.^{B, E}

Глоба М. В.^{A, E, F}

Тарасюк Б. А.^{A, B, D, E, F}

REFERENCES

1. Neuberger, J., Patel, J., Caldwell, H., Davies, S., Hebditch, V., Hollywood, C., et al. (2020). Guidelines on the use of liver biopsy in clinical practice from the British Society of Gastroenterology, the Royal College of Radiologists and the Royal College of Pathology. *Gut*, 69(8), 1382-1403. <https://doi.org/10.1136/gutjnl-2020-321299>
2. Midia, M., Odedra, D., Shuster, A., Midia, R., & Muir, J. (2019). Predictors of bleeding complications following percutaneous image-guided liver biopsy: A scoping review. *Diagnostic and Interventional Radiology*, 25(1), 71-80. <https://doi.org/10.5152/dir.2018.17525>
3. Ferraioli, G., Barr, R. G., Berzigotti, A., et al. (2024). WFUMB guideline/guidance on liver multiparametric ultrasound: Part 1. Update to 2018 guidelines on liver ultrasound elastography. *Ultrasound in Medicine & Biology*, 50(8), 1071-1087. <https://doi.org/10.1016/j.ultrasmedbio.2024.03.013>
4. Ferraioli, G., Barr, R. G., Berzigotti, A., et al. (2024). WFUMB guideline/guidance on liver multiparametric ultrasound: Part 2. Quantification of liver fat with ultrasound. *Ultrasound in Medicine & Biology*, 50(8), 1088-1098. <https://doi.org/10.1016/j.ultrasmedbio.2024.03.014>
5. Liguori, A., Esposto, G., Ainora, M. E., et al. (2025). Liver elastography for liver fibrosis stratification: A comparison of three techniques in a biopsy-controlled MASLD cohort. *Biomedicine*, 13(1), 138. <https://doi.org/10.3390/biomedicine13010138>
6. Song, L., Zhao, L., Deng, J., Meng, F., Wu, X., Lu, Q., Xiang, H., Jing, X., & Luo, Y. (2023). Staging liver fibrosis in patients with chronic hepatitis B using two-dimensional shear wave elastography based on histopathological findings: A prospective multicenter study. *Quantitative Imaging in Medicine and Surgery*, 13(4), 2376-2387. <https://doi.org/10.21037/qims-22-831>
7. Xu, Z.-L., Wang, S.-R., Li, W.-X., Deng, Y.-Q., Kang, Y.-F., Zhang, Y.-R., Li, J., & Cui, X.-W. (2025). Diagnostic performance of point shear wave elastography and vibration-controlled transient elastography in paediatric metabolic dysfunction-associated steatotic liver disease: A systematic review

- and meta-analysis. *BMJ Open*, 15(6), e087101. <https://doi.org/10.1136/bmjopen-2024-087101>
8. Al-Hasani, M., Sultan, L. R., Sagreiya, H., Cary, T. W., Karmacharya, M. B., & Sehgal, C. M. (2022). Ultrasound radiomics for the detection of early-stage liver fibrosis. *Diagnostics*, 12(11), 2737. <https://doi.org/10.3390/diagnostics12112737>
 9. Popa, S. L., Ismaiel, A., Abenavoli, L., et al. (2023). Diagnosis of liver fibrosis using artificial intelligence: A systematic review. *Medicina*, 59(5), 992. <https://doi.org/10.3390/medicina59050992>
 10. Chen, B. R., & Pan, C. Q. (2022). Non-invasive assessment of fibrosis and steatosis in pediatric non-alcoholic fatty liver disease. *Clinics and Research in Hepatology and Gastroenterology*, 46(1), 101755. <https://doi.org/10.1016/j.clinre.2021.101755>
 11. Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., et al. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*, 351, h5527. <https://doi.org/10.1136/bmj.h5527>
 12. Collins, G. S., Dhiman, P., Riley, R. D., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, e078378. <https://doi.org/10.1136/bmj-2023-078378>
 13. Rinella, M. E., Lazarus, J. V., Ratziu, V., et al. (2023). A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology*, 78(6), 1966-1986. <https://doi.org/10.1097/HEP.0000000000000520>
 14. Bedossa, P., & Poynard, T. (1996). An algorithm for the grading of activity in chronic hepatitis C. *Hepatology*, 24(2), 289-293. <https://doi.org/10.1002/hep.510240201>
 15. Sasso, M., Beaugrand, M., de Lédinghen, V., et al. (2010). Controlled attenuation parameter (CAP): A novel VCTE-guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: Preliminary study and validation in a cohort of patients with chronic liver disease from various causes. *Ultrasound in Medicine & Biology*, 36(11), 1825-1835. <https://doi.org/10.1016/j.ultrasmedbio.2010.07.005>
 16. Vos, M. B., Abrams, S. H., Barlow, S. E., et al. (2017). NASPGHAN clinical practice guideline for the diagnosis and treatment of nonalcoholic fatty liver disease in children. *Journal of Pediatric Gastroenterology and Nutrition*, 64(2), 319-334. <https://doi.org/10.1097/MPG.0000000000001482>
 17. Babenko, V., Nastenka, I., Pavlov, V., et al. (2023). Classification of pathologies on medical images using the algorithm of random forest of optimal-complexity trees. *Cybernetics and Systems Analysis*, 59, 346-358. <https://doi.org/10.1007/s10559-023-00569-z>
 18. Babenko, V., Nastenka, I., Solodushenko, V., Tarasiuk, B., Pavlov, V., & Averianova, O. (2024). Technology of hierarchical classification applied to ensemble methods for liver fibrosis staging in ultrasound. In 2024 IEEE 19th International Conference on Computer Science and Information Technologies (CSIT) (pp. 1-4). IEEE. <https://doi.org/10.1109/CSIT65290.2024.10982652>

Abstract**DIAGNOSTIC PERFORMANCE OF MULTIPARAMETRIC ULTRASOUND WITH ARTIFICIAL INTELLIGENCE TOOLS IN PEDIATRIC DIFFUSE LIVER DISEASES**Vitalii O. Babenko¹, Volodymyr V. Soloduschenko^{1,2}, Maryna B. Dyba¹, Maryna V. Globa², Borys A. Tarasiuk¹¹SI «Ukrainian Center of Maternity and Childhood of NAMS of Ukraine», Kyiv, Ukraine²Shupyk National Healthcare University of Ukraine, Kyiv, Ukraine

Introduction. Chronic diffuse liver diseases in children require safe noninvasive assessment and repeated monitoring of fibrosis. Liver biopsy remains the reference method for morphological verification; however, in pediatric practice its use is limited by invasiveness, risk of complications, and ethical considerations. Multiparametric ultrasound, including shear wave elastography and Doppler assessment of blood flow, is suitable for longitudinal follow-up, yet interpretation of ultrasound features remains operator-dependent. Automated image analysis using artificial intelligence algorithms may increase objectivity and reproducibility of fibrosis stratification and support clinical decision-making.

Aim. To evaluate the diagnostic accuracy of multiparametric ultrasound, including two-dimensional shear wave elastography and machine learning algorithms, for staging liver fibrosis in children.

Materials and methods: This retrospective single-center study included 185 children (140 with chronic diffuse liver diseases and 45 controls). All participants underwent two-dimensional shear wave elastography, hepatic vein Doppler assessment, and grayscale ultrasound image analysis using two machine learning approaches: Model A (random forest algorithm; binary classification) and Model B (hierarchical ensemble; multiclass classification). The reference approach was clinical verification and risk stratification based on elastography findings.

Results. A strong positive association was observed between elastography measures and Model B predictions (Spearman rank correlation coefficient 0.77). Model A demonstrated 84% specificity in differentiating normal findings from pathological changes. Model B showed higher sensitivity (83%) but tended to overestimate fibrosis at the boundary between minimal and moderate stages. Doppler patterns, particularly a monophasic hepatic vein waveform, were significantly associated with higher fibrosis stages.

Conclusions. Integrating machine learning algorithms into the ultrasound protocol increases diagnostic objectivity; however, hierarchical approaches require calibration to reduce false-positive results in the borderline zone of early fibrosis stages.

Keywords: pediatric hepatology, chronic diffuse liver diseases, liver fibrosis, multiparametric ultrasound, shear wave elastography, hepatic vein Doppler, artificial intelligence, machine learning

Received: 17.11.2025

Accepted: 04.02.2026