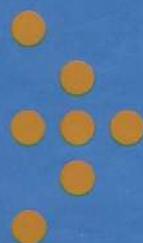


ЦЕНТР ГРОМАДСЬКОГО  
ЗДОРОВ'Я МОЗ УКРАЇНИ



## КЕРІВНИЦТВО

з оцінювання методів  
лабораторних досліджень  
в медичних лабораторіях  
у відповідності до вимог стандарту  
ДСТУ EN ISO 15189:2015  
та ISO 15189:2022  
«Медичні лабораторії.  
Вимоги до якості та компетентності»



ЦЕНТР  
ГРОМАДСЬКОГО  
ЗДОРОВ'Я

# КЕРІВНИЦТВО

з оцінювання методів  
лабораторних досліджень  
в медичних лабораторіях  
у відповідності до вимог стандарту  
ДСТУ EN ISO 15189:2015 та ISO 15189:2022  
«Медичні лабораторії.  
Вимоги до якості та компетентності»

Чуваков О.М., Вінарська Н.В., Андріанова І. В.,  
Черненко Л.М., Танасійчук І.С.

**КЕРІВНИЦТВО**  
з оцінювання методів лабораторних досліджень  
в медичних лабораторіях  
у відповідності до вимог стандарту  
ДСТУ EN ISO 15189:2015 та ISO 15189:2022  
«Медичні лабораторії.  
Вимоги до якості та компетентності»

ISBN 978-617-8085-92-6

УДК 006.4:61]:005.1

ISBN 978-617-8085-92-6

© Авторський колектив, 2023

## ЗМІСТ

1. Позначення та скорочення .....	5
2. Вступ .....	8
3. Основи оцінювання методів .....	9
4. Вибір припустимих специфікацій аналітичних характеристик .....	28
5. Верифікація кількісних методів .....	36
5.1. Верифікація прецизійності та правильності .....	36
5.2. Експеримент з порівняння методів .....	43
5.3. Оцінка аналітичної чутливості методу .....	48
5.4. Оцінка лінійності .....	49
5.5. Оцінка переносу .....	51
5.6. Верифікація калібрування .....	52
5.7. Оцінка невизначеності .....	54
5.8. Верифікація нового лоту реагентів .....	66
5.9. Верифікація референтних інтервалів .....	69
6. Верифікація напівкількісних та якісних методів .....	72
7. Особливості оцінювання методів в окремих галузях лабораторної медицини .....	83
7.1. Оцінка гематологічних методів .....	83
7.2. Оцінка молекулярно-діагностичних методів .....	85
7.3. Оцінка мікробіологічних методів .....	88

8. Додатки .....	95
8.1. Значення біологічної варіації та припустимі аналітичні специфікації щодо зсуву $v(\%)$ , неprecизійності $i(\%)$ , максимально припустимої невизначеності MAU та загальної аналітичної помилки TEA деяких тестів .....	95
8.2. Зразки звітів з верифікації та валідації методики дослідження .....	105
9. Бібліографія .....	109



#### Електронний додаток містить:

- електронну версію підручника;
- робочі інструменти:
  - інструмент верифікації прецизійності та правильності кількісної методики досліджень;
  - інструмент верифікації при зміні лоту кількісної методики досліджень;
  - інструмент оцінювання невизначеності кількісної методики згідно з ISO TS 20914;
  - інструмент верифікації референтних інтервалів кількісної методики;
  - інструмент верифікації лінійності кількісної методики;
  - інструмент верифікації прецизійності та узгодженості якісної методики дослідження;
  - інструмент верифікації діагностичних чутливості та специфічності якісних методик досліджень.

## 1. ПОЗНАЧЕННЯ ТА СКОРОЧЕННЯ

AMI	аналітичний діапазон вимірювання (Analytical measuring interval)
ANOVA	дисперсійний аналіз (analysis of variance)
ATE (TEa)	припустима загальна помилка (allowable total error)
ASTS	система визначення чутливості до антибіотиків (antimicrobial susceptibility testing system(s))
BV	біологічна варіація (biological variability)
CA	узгодженість категорій
CD	критична різниця (critical difference)
CLSI	Інститут клінічних та лабораторних стандартів (Clinical and Laboratory Standards Institute)
CV	коефіцієнт варіації (coefficient of variation)
CV <sub>w</sub> , або CV <sub>i</sub>	внутрішньоіндивідуальний коефіцієнт варіації
CV <sub>b</sub> , або CV <sub>g</sub>	міжіндивідуальний коефіцієнт варіації
%CV	коефіцієнт варіації
%CV <sub>B</sub>	коефіцієнт варіації між серіями (between run)
%CV <sub>R</sub>	внутрішньосерійний коефіцієнт варіації, повторюваність (repeatability)
%CV <sub>WL</sub>	внутрішньолабораторний (within-laboratory) коефіцієнт варіації
DF	ступені свободи (degrees of freedom)
df <sub>c</sub>	комбінований ступінь свободи (combined)
df <sub>R</sub>	ступені свободи для повторюваності (repeatability)
df <sub>RM</sub>	ступені свободи для референтного матеріалу (reference materials)
df <sub>WL</sub>	ступені свободи для внутрішньолабораторної прецизійності (within-laboratory imprecision)
EA	есенціальна узгодженість
EFLM	Європейська федерація клінічної хімії та лабораторної медицини (European Federation of Clinical Chemistry and Laboratory Medicine)
EMI	розширений інтервал вимірювання (extended measuring interval)
EQA	зовнішня оцінка якості (external quality assessment)
FN	хибнонегативні результати (False negative)
FP	хибнопозитивні результати (False positive)
IFCC	Міжнародна федерація клінічної хімії та лабораторної медицини (International Federation of Clinical Chemistry and Laboratory Medicine)

IQR	міжквартильний інтервал (interquartile range)
ISO	Міжнародна організація зі стандартизації (International Organization for Standardization)
VMD	дуже велика невідповідність (very major discrepancy)
VME	дуже велика помилка (very major error)
JCTLM	Об'єднаний комітет з простежуваності в лабораторній медицині (Joint Committee for Traceability in Laboratory Medicine)
<i>k</i>	фактор покриття (coverage factor)
K-S	критерій Колмогорова – Смирнова (Kolmogorov Smirnov)
LoB	ліміт бланку (limit of blank)
LoD	ліміт визначення (limit of detection)
LoQ, LLQ	ліміт кількісного визначення (limit of quantitation, Lower limit of quantitation)
MAU	максимально припустима розширена невизначеність (maximum allowable measurement uncertainty)
mD	незначна різниця (minor discrepancy)
ME	велика помилка (major error)
mE	незначна помилка (minor error)
MIC	мінімальна інгібуюча концентрація (minimal inhibitory concentration)
MIS	система ідентифікації мікроорганізмів (microbial identification system(s))
MRSA	метицилін-резистентний золотистий стафілокок (methicillin-resistant <i>Staphylococcus aureus</i> )
MS	середнє квадратів (mean squares)
MU	невизначеність результатів вимірювань (measurement uncertainty)
M-W U	критерій Манна-Уїтні U (Mann-Whitney U)
<i>n</i> Sam	кількість зразків (number of samples)
NPA	відсоток негативної узгодженості (negative percent agreement)
NPV	прогностична цінність негативного результату (negative predictive value)
$\sigma_R$	заяви виробника щодо повторюваності (manufacturer's claim for repeatability)
$\sigma_{WL}$	заяви виробника щодо внутрішньолaboratorної неprecizійності (manufacturer's claim for within-laboratory imprecision)
PPA	відсоток позитивної узгодженості (positive percent agreement)

PPV	прогностична цінність позитивного результату (positive predictive value)
QC	контроль якості (quality control)
<i>r</i>	коефіцієнт кореляції (correlation coefficient)
RCPA	Королівський коледж патологів Австралії (Royal College Pathologists of Australia)
ROC	крива похибок (receiver operating characteristic curve)
SD	стандартне відхилення (standard deviation)
$s_R$	отримана користувачем оцінка повторюваності (user estimate for repeatability)
SS	сума квадратів sum of squares
$s_{WL}$	отримана користувачем оцінка внутрішньолaboratorної неprecizійності (user estimate for within-laboratory imprecision)
TP	істинно позитивні результати (true positive)
TN	істинно негативні результати (true negative)
TV	цільове значення, опорне значення (target value)
ULQ	верхня межа кількісного визначення (upper limit of quantitation)
UVL	верхній ліміт верифікації (upper verification limit)
VIM	словник з метрології (Vocabulary in Metrology)
$V_B$	варіація між серіями (variance between runs)
$V_W$	варіація в аналітичній серії (variance within run)
$U(x)$	розширена невизначеність (expanded uncertainty)
$u_c(x)$	комбінована невизначеність (combined uncertainty)
$u(x)$	стандартна невизначеність (standard uncertainty on the quantity value)
$u_{briola}$	припустима різниця між лотами (between lot reagents, allowable)
АЛТ	аланінамінотрансфераза
АСТ	аспартатамінотрансфераза
ВЛКЯ	внутрішньолaboratorний контроль якості
ДІ	довірчий інтервал (CI, confidence interval)
ЗОЯ	зовнішня оцінка якості
ІФА	імуноферментний аналіз
МЛ	медична лабораторія
ПЗ	програмне забезпечення
СКВ	середньоквадратичне відхилення

## 2. ВСТУП

Завдяки запровадженню стандартів ISO 15189 і ISO 17025 клінічні лабораторії проводять велику роботу щодо покращення якості. Використання в клінічній практиці тільки валідованих щодо свого призначення методів та верифікація досягнення в лабораторії припустимих аналітичних характеристик тесту є однією з фундаментальних вимог стандартів, що містять вимоги до технічної компетентності вимірювальних, в тому числі медичних, лабораторій.

Разом з тим, методики валідації та верифікації лабораторного методу не зазначено безпосередньо в стандарті, тому вони часто значно відрізняються від лабораторії до лабораторії.

Крім того, на цей час вже затверджено нову редакцію – ISO 15189:2022, яка гармонізована з ISO 9001:2015 та ISO/IEC 17025:2017 та містить важливі доповнення щодо оцінювання методів досліджень.

Багато з того, що публікується на цю тему, містить складну статистику і її важко реалізувати в звичайних лабораторіях. Результатом є те, що цим пунктом часто нехтують або реалізують його неправильно, що, у свою чергу, може призвести до хибних висновків щодо ефективності методу, що потенційно може поставити під загрозу безпеку пацієнтів або спричинити неправильний діагноз.

Окрім того, сучасними для акредитованих лабораторій є вимоги оцінювати та документувати аналітичну ефективність усіх методів не лише до їх першого впровадження, але також і під час поточної експлуатації, отже, є потреба в чітких, стандартизованих і практичних рекомендаціях з цього питання.

Це знайшло місце і в новій редакції ISO 15189:2022, п.7.3.1: «Уповноважений персонал повинен **періодично** оцінювати методи дослідження, що використовуються лабораторією, щоб переконатися, що вони клінічно відповідають отриманим запитам».

Керівництво узагальнює поточну літературу провідних джерел (CLSI, ACB, EFLM та інших) з цієї теми, зосереджуючи увагу на вимогах до валідації методів або, залежно від випадку, верифікацій, та які статистичні тести підходять для кожного. Наведено правильну інтерпретацію результатів статистичних тестів і критерії прийнятності для кожної процедури. Надані інструменти, реалізовані у вигляді робочих листів EXCEL, дають можливість обробити результати верифікаційних експериментів у будь-якій лабораторії.

Отже, відповідно до ISO 15189, методики вимірювань повинні піддаватися незалежній валідації або верифікації медичною лабораторією перед введенням у використання. Аналітичні характеристики та критерії прийнятності, які є важливими для валідації або верифікації, залежно від передбачуваного використання, повинні бути встановлені у плані валідації або верифікації, щоб об'єктивно оцінити, чи методика вимірювання придатна для передбачуваного використання.

## 3. ОСНОВИ ОЦІНЮВАННЯ МЕТОДІВ

### Статистичні методи, що використовуються для оцінювання методу дослідження

Оцінка аналітичних методів передбачає надання об'єктивних доказів щодо аналітичних характеристик метода вимірювання. Оцінка методів є принциповою вимогою стандарту та може передбачати як оцінювання перед використанням методики, так і моніторинг незмінності аналітичних характеристик протягом усього періоду використання її в лабораторії.

Для оцінювання аналітичних характеристик методики дослідження використовується досить складний комплекс методів лабораторної біостатистики, який містить в собі описову статистику (розподіл, середнє, СКВ, гістограми), тестування гіпотез, ANOVA, регресійний аналіз, кореляційний аналіз.

У зв'язку зі складністю даних, що можуть бути результатами експериментів з оцінювання методів, використовуються як параметричні (припускається нормальний розподіл даних), так і непараметричні методи, які не вимагають нормального розподілу.

ТАБЛИЦЯ 3.1. ОСНОВНІ МЕТОДИ ЛАБОРАТОРНОЇ БІОСТАТИСТИКИ

Показники	Параметричні	Непараметричні
Центральної тенденції	Середнє	Медіана
Варіабельності, розкиду	Дисперсія ( $\sigma$ ) Середньоквадратичне відхилення (SD) Коефіцієнт варіації (CV)	Міжквартильний розмах (IQR)
Коефіцієнт кореляції	Пірсона ( $r$ )	Спірмена ( $\rho - \rho_0$ , або $r_s$ ), Кендала ( $\tau$ -tau)
Статистичні тести (критерії оцінювання)	Z-критерій (z-тест) t-критерій Стьюдента (t-тест) F-критерій Фішера (F-тест)	Критерій узгодження Пірсона $\chi^2$ (хі-квадрат) Критерій знакових рангів Вілкоксона Критерій Манна-Уїтні
Регресія	Проста лінійна (найменших квадратів) Регресія Демінга	Пассинга-Баблока

Використання цих методів значно полегшується при застосуванні програмного забезпечення, спеціально призначеного для проведення оцінювання методів (method evaluation).

Найбільш відомими ПЗ є спеціалізовані програми Analyze It, EP Evaluator, Abacus. Разом з тим, поширеним є застосування спеціально приготованих робочих листів електронних таблиць, що реалізують відповідні розрахунки.

Раціональний дизайн експериментів з оцінювання методів та обробки їх результатів забезпечується використанням стандартизованих протоколів оцінювання, найбільш відомими з яких є протоколи оцінювання (Evaluation Protocols, EP) Clinical Laboratory Standards Institute (CLSI, США). Інші лабораторні товариства також створюють свої протоколи оцінювання методів, які будуть висвітлені в цьому керівництві.

### Основні статистичні поняття

**Закон розподілу** – математична функція, яка описує розподіл величини.

**Закон нормального розподілу** – математична функція Гауса, що описує ділянку усіх можливих значень змінної випадкової величини та ймовірність отримання конкретного результату. Функція Гауса характеризує розподіл величини за допомогою лише двох параметрів: математичного сподівання  $\mu$  (середнє всіх можливих значень) та дисперсії  $\sigma$ . Основним постулатом закону нормального розподілу є доведене ствердження про те, що ймовірність невеликих відхилень результатів вимірювань від дійсного значення є набагато вищою за ймовірність великих відхилень. На основі нормального розподілу можна обчислити теоретичний розподіл результатів навколо середнього значення. Крива нормального розподілу, що ілюструє вірогідність для результату, який буде міститись у заданих межах ( $X_{\text{сер}}$  – середнє значення,  $S$  – стандартне відхилення). Діапазон навколо середнього значення, в межах якого, як очікується, лежить встановлений відсоток значень. Наприклад, для нормального розподілу, приблизно 95 % значень лежать у межах  $\pm 2 S$ .

**Стандартний нормальний розподіл** – найпростіший варіант нормального розподілу, особливий випадок, коли математичне сподівання  $\mu = 0$  та дисперсія  $\sigma = 1$ .

**Нульова гіпотеза ( $H_0$ )** – це загальне припущення при статистичному висновуванні про те, що не існує ніякого зв'язку між двома вимірюваними явищами або не існує ніякого зв'язку між групами. Поняття нульової гіпотези передбачає існування **альтернативної гіпотези ( $H_1$ )**, що її заперечує. Прийнято вважати, що нульова гіпотеза відповідає

стану «за умовчанням» (природному, найбільш очікуваному стану речей, відповідно до оцінювання методів – це гіпотеза, що функціональні характеристики в межах припустимих значень). Відповідно, альтернативна гіпотеза позначає протилежну ситуацію, яка зазвичай трактується як менш вірогідна, неординарна, така, що вимагає якої-небудь реакції (функціональні характеристики не відповідають припустимим).

**Помилки першого роду (type I errors,  $\alpha$  errors, false positives)** – істинна нульова гіпотеза помилково відкидається, це є хибнопозитивним результатом, помилковою тривоною (хибне припущення, що функціональні характеристики не відповідають вимогам).

**Помилки другого роду (type II errors,  $\beta$  errors, false negatives)** – помилково приймається хибна нульова гіпотеза – пропуск події (хибне припущення, що функціональні характеристики відповідають вимогам).

**Статистичний критерій** – це правило чи процедура, яка дозволяє прийняти чи відкинути статистичну гіпотезу. Це кількісна оцінка показників роботи процесу, що отримана за допомогою методів математичної статистики, яка дозволяє зробити обґрунтовані висновки щодо виникнення значущих змін в роботі процесу.

**Статистична значущість критерію** – це вірогідність того, що отриманий результат правильно представляє популяцію, вибірка з якої досліджувалася ( $p$ -рівень значущості), тобто виявлення змін у стабільності роботи процесу там, де вони є. Рівень значущості – це характеристика критерію, яка базується на ймовірності ухвалити рішення відхилити нульову гіпотезу, якщо насправді нульова гіпотеза правильна (рішення відоме як похибка першого роду). В лабораторній медицині зазвичай використовують критерії з 90%, 95% та 99% рівнем ймовірності.

**Статистична потужність критерію** – це ймовірність того, що тест виявляє різницю, правильно відкидає нульову гіпотезу ( $H_0$ ), коли альтернативна гіпотеза ( $H_1$ ) істинна (похибка другого роду). Це здатність не припуститися помилкових висновків про стабільність роботи процесу, тобто не спрацювати там, де зміни відсутні.

**Подія (event)** – під подією в теорії ймовірностей розуміють будь-який факт, який може відбутися або не відбутися в результаті експерименту при заданих умовах, причому, існує визначена ймовірність того, що вона відбудеться при певних умовах. Подія може вважатися випадковою, лише коли вона теоретично може повторитись довільну кількість разів.

**Ймовірність (probability)** – це числова характеристика можливості того, що деяка подія відбудеться в регламентованих умовах, які можуть

бути відтворені необмежену кількість разів. Ймовірність завжди позитивна (не може бути нижче нуля), сума усіх можливих ймовірностей для конкретної події (величини) завжди дорівнює одиниці або 100 %. Стосовно лабораторної аналітики ймовірність характеризує можливість отримання того чи іншого результату вимірювання в певних умовах.

**Постійна величина, константа (constant)** – це така величина, яка в умовах експерименту не змінюється.

**Випадкова змінна величина (variable)** – це характеристика або значення, яке збільшується, зменшується чи коливається в певних межах протягом часу або в залежності від умов вимірювання. Випадкові змінні величини утворюють розподіл.

**Середнє (арифметичне) значення (mean)** – оцінка математичного сподівання, тобто середнє значення обмеженої кількості даних. Обчислюється за формулою:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

де  $X_i$  – окреме значення набору даних,  $\Sigma$  – знак суми,  $n$  – кількість значень.

**Медіана** – величина, що розташована в середині ряду величин, розташованих у зростальному або спадному порядку.

**Міжквартильний розмах (interquartile range, IQR)** – це різниця між 75-м та 25-м процентилями.

### Основні метрологічні поняття

**Фізична величина (quantity)** – властивість, спільна в якісному відношенні у багатьох матеріальних об'єктів (вид величини) та індивідуальна у кількісному відношенні у кожного з них (розмір величини).

**Результат вимірювання (result of a measurement)** – значення фізичної величини, що отримане шляхом її вимірювання. Результат вимірювання є безперервною кількісною величиною і завжди містить похибку.

**Істинне (ідеальне) значення величини (true value)** – результат вимірювання, який можна було б отримати в ідеальних умовах. На практиці отримати істинне значення величини, що вимірюється, неможливо через об'єктивні причини (вплив різноманітних факторів, які неможливо повністю усунути), внаслідок чого результат вимірювання відхиляється в ту чи іншу сторону. Істинне значення величини є теоретичним поняттям і, як правило, не може бути точно відомим. Вважається, що ідеальне значення наближається до математичного сподівання за умови проведення безкінечної кількості вимірювань.

**Прийняте опорне (еталонне) значення (accepted reference value)** – це поняття застосовується в лабораторній медицині замість терміну «істинне значення». Воно має сенс узгодженого значення, яке застосовується для порівняння (зіставлення) результатів вимірювань, та використовується для обчислення систематичної помилки вимірювання. Синоніми: атестоване значення, консенсусне значення, цільове значення, опорне значення, середнє групи порівняння та ін.

**Помилка (похибка) вимірювання (measurement error)** – кількісна характеристика точності результату. Це різниця між отриманим результатом вимірюваної величини та її істинним (ідеальним) значенням. Може бути подана в одиницях вимірюваної величини (абсолютна похибка: різниця між отриманим результатом та істинним значенням) або у відсотках (відносна похибка: відношення абсолютної похибки вимірювання до істинного значення вимірюваної величини). Через недоступність інформації про істинне значення вимірюваної величини похибку окремого результату неможливо розрахувати точно, її можна тільки оцінити (тобто знайти її наближене значення). Традиційно в лабораторній медицині вживається термін «помилка», хоча метрологи вважають його невдалим. В даному документі ці два терміни використовуються як взаємозамінні.

**Загальна помилка (total error, TE)** – термін, подібний до метрологічного терміну «похибка», але, на відміну від нього, є характеристикою не окремого результату, а процесу вимірювання. Загальну похибку окремого результату лабораторного дослідження розрахувати неможливо, тому що його істинне значення завжди невідоме. Оскільки загальна помилка залежить від сумарної дії факторів, що спричиняють систематичну та випадкову помилки, її можна лише приписати окремому результату з певною долею ймовірності на підставі розрахованих кількісних мір систематичної та випадкової похибок. Концепція **Загальної помилки** запропонована Джеймсом Вестгардом.

**Загальна аналітична помилка (analytical total error, aTE)** – загальна помилка аналітичного етапу лабораторного дослідження.

**Допустима загальна помилка (allowable total error, TEa)** – вимоги до якості, сформульовані у вигляді загальної похибки. Можуть бути представлені в абсолютному значенні (в одиницях вимірювання величини) й у відносному.

**Систематична помилка (systematic error, SE)** – складова загальної помилки, яка залишається постійною (константна) або прогнозовано змінюється (пропорційна) у ряді вимірювань однієї й тієї ж величини. Зазвичай SE однакова за розміром та знаком, якщо умови вимірювання стабільні. Для оцінювання систематичної помилки усереднені результати серії вимірювань порівнюють з опорним, атесто-

ваним або іншим узгодженим значенням, прийнятим за істинне. З систематичною помилкою пов'язані такі терміни, як **зсув** методу та **дрейф** методу.

**Випадкова помилка (random error, RE)** – складова загальної помилки, що не прогнозовано змінюється в ряді вимірювань однієї й тієї ж величини. Випадкові похибки не однакові за розміром та знаком, вони обумовлені впливом сукупності випадкових факторів, які неможливо усунути. Для оцінювання випадкової помилки розраховують показники варіації, отримані на підставі певної кількості вимірювань. Випадкова помилка має нормальний розподіл (Гауса).

**Варіація результатів дослідження (test results variation)** – це статистичний термін, що характеризує числову різницю між результатами повторних вимірювань однієї й тієї ж величини. Варіація виникає через вплив на процес вимірювання випадкових причин. Кількісною оцінкою варіації є дисперсія  $\sigma$  (міра розсіювання даних).

**Репліката (replicate)** – повторні вимірювання, що виконуються в регламентованих умовах з метою оцінювання випадкової помилки.

**Надмірна похибка вимірювання (gross measurement error)** – похибка вимірювання, що суттєво перебільшує очікувану (в даних умовах) похибку.

**Промах, викид, аномальний результат, абсурдний результат (outlier)** – результат вимірювання, який містить надмірну похибку. Такий результат виникає через грубі порушення процедури лабораторного дослідження або роботи аналітичної системи.

**Аналітична надійність результатів (analytical reliability of the results)** характеризує ступінь достовірності лабораторних даних про аналіт, що досліджується. Цільовим параметром аналітичної надійності лабораторних досліджень є їх здатність достовірно розмежувати значення, які властиві станам здоров'я та патології. Аналітична надійність залежить від аналітичних характеристик та стабільності роботи метода.

**Цілі аналітичної якості (analytical quality goals)** – кількісні критерії, що встановлені як вимоги до аналітичної надійності методики з урахуванням її призначення до застосування. Як синонім в лабораторній медицині вживаються терміни «**припустимі функціональні характеристики методики**» та «**норми точності**».

- В залежності від виду методики – кількісна це методика чи якісна, визначено наступні аналітичні характеристики:
- селективність;
- межа виявлення (LOD) та межа кількісного визначення (LOQ);
- робочий діапазон;
- аналітична чутливість;

- правильність (зсув, ступінь вилучення);
- прецизійність (збіжність, проміжна прецизійність) та відтворюваність);
- непевність (невизначеність) вимірювань;
- стійкість (робастність);
- референтні значення;
- перенесення (carryover);
- лінійність та перевірка автоматичного розведення;
- діагностична чутливість та діагностична специфічність.

**Прецизійність вимірів (measurement precision)** – близькість між показами або виміряними значеннями величини, отриманими під час повторних вимірювань на одному або на подібних об'єктах за заданих умов (VIM 2.15). Аналог терміну, що раніше використовували – «відтворюваність».

В залежності від умов проведення вимірювань, розрізняють прецизійність день-у-день, прецизійність внутрішньосерійну, міжлотову, міжсерійну, внутрішньолабораторну, міжлабораторну, проміжну, довгострокову тощо.

На прецизійність впливає величина випадкової похибки: чим менше похибка, тим краще прецизійність.

Кількісною мірою прецизійності є **середньоквадратичне, або стандартне відхилення (Standard Deviation, SD, S)** та коефіцієнт варіації (**Coefficient of Variation, CV**).

**SD(S)** та **CV** є похідними дисперсії та розраховуються за формулами:

$$S = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{n-1}} \quad CV = \frac{S}{\bar{x}} * 100 \%$$

Прецизійність вимірів чисельно виражають через міри розсіювання, наприклад, стандартний відхил, розрахований за результатами, отриманими під час повторних вимірювань відповідного зразка за певних умов.

**VIM 3 визначає три види умов вимірювання:**

- умови збіжності (VIM 2.20);
- умови проміжної прецизійності (VIM 2.22);
- умови відтворюваності (VIM 2.24).

**Збіжність** – прецизійність в умовах збіжності.

**Умови збіжності** – умови, за яких вимірювання виконуються із застосуванням одного й того ж методу на ідентичних об'єктах випробування, в одній лабораторії, одним аналітиком, з викорис-

танням одного й того ж обладнання впродовж короткого проміжку часу. Збіжність вимірів часто використовують як оцінку розсіювання результатів всередині партії, збіжність вимірів відображає мінливість результатів протягом лише короткого періоду часу. Оцінювання прецизійності в таких умовах переважно не дозволяє врахувати всі фактори, що впливають на розсіювання результатів під час регулярного застосування відповідної методики.

**Проміжна прецизійність** – прецизійність в умовах проміжної прецизійності.

**Умови проміжної прецизійності** – умови, за яких вимірювання виконуються із застосуванням одного й того ж методу на ідентичних об'єктах випробовування, в одній лабораторії протягом тривалого проміжку часу.

У кожному конкретному випадку ці умови можуть відрізнитися (різні аналітики, різна тривалість вимірювання, різне обладнання тощо), але в будь-якому разі вони є більш мінливими, ніж умови збіжності. Оцінювання прецизійності в таких умовах відображає максимальну кількість факторів розсіювання результатів, що мають місце в одній лабораторії під час її регулярної роботи.

Проміжну прецизійність часто використовують як оцінку розсіювання результатів між партіями. Показник проміжної прецизійності є більш реальною оцінкою розсіювання результатів вимірювань у лабораторії для тривалого періоду часу і є основою для оцінювання невизначеності результатів вимірювань в медичних лабораторіях.

**Відтворюваність** – прецизійність в умовах відтворюваності.

**Умови відтворюваності** – умови, за яких вимірювання виконуються із застосуванням одного й того ж методу на ідентичних об'єктах випробовування, в різних лабораторіях, що забезпечує зазвичай найбільше розсіювання результатів.



РИС. 3.1. ВІДТВОРЮВАНІСТЬ ТА ПОВ'ЯЗАНІ ПОНЯТТЯ

**Аналітична селективність** – ступінь, до якого метод можна застосувати для визначення конкретних аналітів у сумішах або матрицях без завад з боку інших компонентів з аналогічними властивостями (у деяких галузях використовують терміни «специфічність» або «аналітична специфічність»).

**Інтервал вимірювання (measuring interval)** – множина значень величин одного роду, які можна виміряти даним засобом вимірювання або вимірювальною системою із зазначеною інструментальною невпевністю за заданих умов (VIM4.7).

**Межа виявлення (LOD, detection limit)** – мінімальний рівень аналіту, який може бути виявлений при проведенні вимірювання. Розрізняють межу виявлення приладу і межу виявлення методу. Межа виявлення приладу є характеристикою безпосередньо вимірювальної процедури, найчастіше її можна визначити за результатами аналізування холостої проби. Межа виявлення методу визначається за даними аналізування проб, які було піддано усім операціям методики вимірювання, які застосовуються для рутинних проб.

**Стійкість (робастність)** – здатність методики зберігати свої характеристики за умови впливу можливих небажаних зовнішніх чинників. Стійкість є показником надійності методу під час його нормального застосування.

**Межа кількісного вимірювання (LOQ, limit of quantitation)** – найнижчий рівень вмісту аналіту, який можна визначити з прийнятними показниками точності.

**Робочий діапазон, інтервал вимірювання (measuring interval)** – інтервал, у межах якого метод забезпечує результати з прийнятною невизначеністю. Робочий діапазон обмежений знизу межею кількісного визначення LOQ. Верхня межа робочого діапазону визначається концентрацією, за якої мають місце значні аномалії аналітичної чутливості.

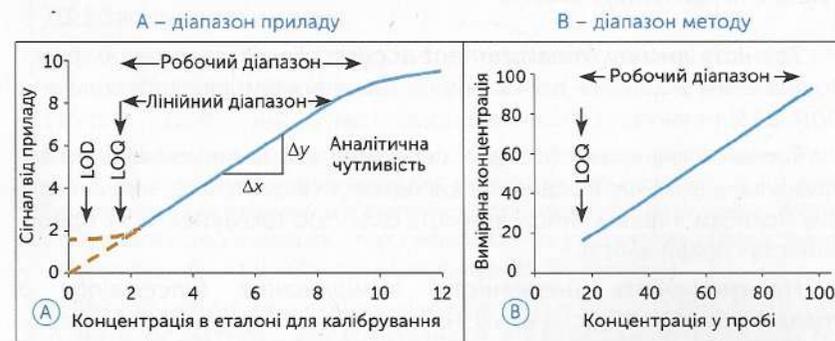


РИС. 3.2. РОБОЧИЙ ДІАПАЗОН

**Правильність вимірів (measurement trueness)** – близькість середнього арифметичного нескінченно великого числа повторно вимірних значень величини до опорного значення величини (VIM 2.14).

На правильність – характеристика якості вимірювання, що відображає близькість до нуля систематичної похибки: чим менше похибка, тим вище правильність. Кількісною мірою правильності є показник зсув ( $B$ ).

**Зсув (Bias,  $B$ )** – це різниця між середнім значенням серії результатів вимірювання та істинним (опорним) значенням. Найчастіше використовується у вигляді відносної величини та розраховується за формулою:

$$B = (X_{\text{сер}} - a) / a \times 100 \%$$

$$\text{або } B = ((X_{\text{сер}} - X_0) / X_0) \times 100 \%$$

де:  $X_{\text{сер}}$  – середнє значення серії результатів вимірювання;  
 $a$  (або  $X_0$ ) – істинне (опорне, атестоване) значення.

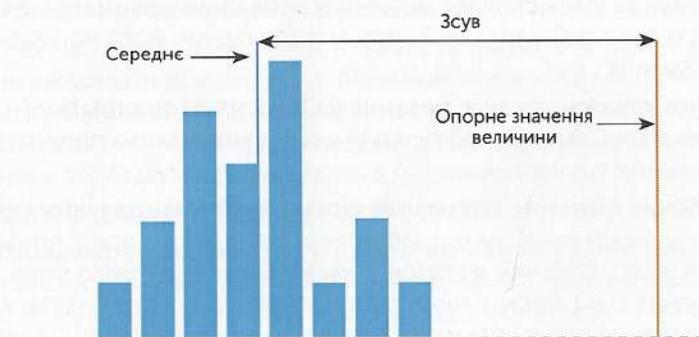


РИС. 3.3. ПРАВИЛЬНІСТЬ ВИМІРІВ

**Точність виміру (measurement accuracy)** – близькість виміряного значення величини до істинного значення вимірюваної величини (VIM 2.13).

Термін «точність» стосовно серії результатів вимірювань чи випробувань включає поєднання випадкової та систематичної складових похибки, і таким чином точність охоплює такі аспекти, як прецизійність і правильність.

**Невизначеність (непевність) вимірювання (uncertainty of measurement)** – сучасна кількісна характеристика точності результату, що описує діапазон, в якому лежить істинне значення вимірюваної величини.

В сучасній лабораторній медицині необхідність такої оцінювання результату лабораторного дослідження набула такого значення як новий параметр вимог щодо аналітичної якості методики.

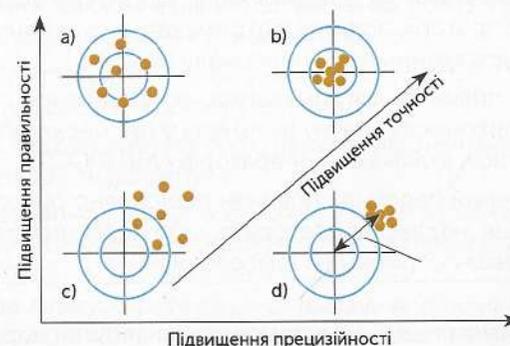


РИС. 3.4. ТОЧНІСТЬ ВИМІРІВ

### Процес оцінювання методів у медичній лабораторії

Відповідно до вимог ISO 15189, методики вимірювань повинні піддаватися незалежній валідації або верифікації медичною лабораторією перед введенням у використання. Аналітичні характеристики та критерії прийнятності, які є важливими для валідації або верифікації, залежно від передбачуваного використання, повинні бути встановлені у плані валідації або верифікації, щоб об'єктивно оцінити, чи методика вимірювання придатна для передбачуваного використання.

#### ISO 15189:2022

### 7.3 Процеси дослідження (Examination processes)

#### 7.3.1 Загальні положення

a) Лабораторія повинна вибрати та використовувати методи обстеження, валідовані для їхнього використання за призначенням, щоб забезпечити клінічну точність обстеження для тестування пацієнтів.

**Примітка.** Рекомендованими методами є ті, що вказані в інструкціях із використання медичних пристроїв для діагностики *in vitro*, або ті, які були опубліковані в усталених/авторитетних підручниках, рецензованих текстах чи журналах, або в міжнародних і національних узгоджених стандартах чи настановах, або національних чи регіональних положеннях.

b) Функціональні характеристики кожного методу обстеження повинні стосуватися передбачуваного використання цього обстеження та його впливу на допомогу пацієнту.

## ISO 15189:2022

с) Усі процедури та допоміжна документація, така як інструкції, стандарти, посібники та довідкові дані, що мають відношення до діяльності лабораторії, повинні підтримуватися в актуальному стані та бути легкодоступними для персоналу (див. 8.3).

д) Персонал повинен дотримуватись встановлених процедур, а дані осіб, які виконують значну діяльність у процесах іспиту, повинні реєструватися, включаючи операторів АМЛ (РОСТ).

е) Уповноважений персонал повинен періодично оцінювати методи дослідження, надані лабораторією, щоб переконатися, що вони клінічно відповідні отриманим замовленням.

Призначення валідації – **встановити** аналітичні характеристики методики та підтвердити їх відповідність призначенню – вимогам споживача – відповідним **цілям аналітичної якості / діагностичним характеристикам**.

## ISO 15189:2022

**3.31 Валідація (validation)**

підтвердження придатності для конкретного передбачуваного використання або застосування шляхом надання об'єктивних доказів того, що визначені вимоги були виконані

**Примітка 1 до запису.** Валідація може бути застосована до тверджень для підтвердження інформації, заявленої у заяві, щодо передбачуваного майбутнього використання.

**Примітка 2 до запису.** Об'єктивні докази можна отримати шляхом спостереження, вимірювання, дослідження або іншим способом.

**Примітка 3 до запису.** Слово «валідовано» використовується для позначення відповідного статусу.

**Примітка 4 до запису.** Визначені вимоги до методу дослідження можуть включати наступне: специфікації ефективності; правильність вимірювання, прецизійність вимірювання, включаючи повторюваність вимірювання, і інтерферуючі речовини, межу виявлення та межу кількісного визначення, інтервал вимірювання, клінічну релевантність, діагностичну специфічність і діагностичну чутливість.

Призначення верифікації – **підтвердити** отримані під час валідації (та надані виробником наприклад) **аналітичні характеристики** в умовах конкретної лабораторії.

## ISO 15189:2022

**Верифікація (verification)**

підтвердження шляхом надання об'єктивних доказів достовірності тверджень щодо виконання зазначених вимог

**Приклад 1.** Підтвердження того, що характеристики ефективності вимірювальної системи досягнуто.

**Приклад 2.** Підтвердження того, що цільова невизначеність вимірювання може бути досягнута.

**Примітка 1 до запису.** Верифікація може бути застосована до тверджень, щоб підтвердити інформацію, заявлену у твердженні щодо подій, які вже відбулися, або результатів, які вже були отримані.

**Примітка 2 до запису.** Верифікація – це процес, за допомогою якого лабораторія підтверджує, що встановлені вимоги щодо ефективності вимірювальної системи (наприклад, істинність, точність, звітний діапазон) можуть бути відтворені в лабораторії перед проведенням дослідження зразка людини.

**Примітка 3 до запису.** Об'єктивними доказами, необхідними для верифікації, можуть бути результати перевірки або інші форми визначення, такі як виконання альтернативних розрахунків або перегляд документів.

**Примітка 4 до запису.** Верифікація може бути достатньою для впровадження нового пристрою IVD за обставин, коли дослідження (3.8) виконується та використовується відповідно до вказівок у інструкції виробника.

**Примітка 5 до запису.** Слово «верифіковано» використовується для позначення відповідного статусу.

Таким чином при верифікації треба підтвердити, що:

1. Характеристики методики не відрізняються від наданих виробником (з певною ймовірністю);
2. Цілі аналітичної якості, обрані лабораторією, виконуються.

В стандарті не визначені методики валідації/верифікації, вибір аналітичних характеристик та методики оцінювання здійснюється спеціалістами лабораторії на основі аналізу ризиків, щоб забезпечити виконання вимог за обставин даної лабораторії.

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<b>5.5.1.3 Валідація методик досліджень</b> Лабораторія повинна провести валідацію методик досліджень, що мають наступне походження:	<b>7.3.3 Валідація методів дослідження</b> а) Лабораторія повинна валідувати методи дослідження, отримані з таких джерел:

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<p>a) нестандартизовані методики;</p> <p>b) методики, розроблені та відпрацьовані лабораторією;</p> <p>c) стандартизовані методики, які використовуються поза межами їх цільового застосування;</p> <p>d) валідовані методи, що в подальшому зазнали модифікації.</p> <p>Валідація має бути настільки повною, наскільки це необхідно, і підтверджувати, за допомогою надання об'єктивних доказів (у вигляді функціональних характеристик), виконання в повному обсязі спеціальних вимог щодо використання дослідження за призначенням.</p> <p><i>Примітка.</i> Функціональні характеристики методики дослідження мають враховувати правильність, точність і прецизійність вимірювань, включаючи повторюваність вимірювань і проміжну прецизійність вимірювань; невизначеність вимірювань, аналітичну специфічність, включаючи речовини, що заважають, аналітичну чутливість, межі виявлення і межі кількісного визначення, діапазон вимірювань, діагностичну специфічність та діагностичну чутливість.</p> <p>Лабораторія повинна задокументувати процедуру, використану для валідації, і отримані результати. Уповноважений персонал повинен проаналізувати результати валідації та задокументувати висновки.</p>	<p>1) методи, запропоновані або розроблені лабораторією;</p> <p>2) методи, що використовуються поза їх первісно призначенням (тобто поза інструкціями виробника щодо використання або оригінальним валідованим діапазоном вимірювань, сторонніми реагентами, окрім призначених приладів, і якщо дані валідації відсутні);</p> <p>3) валідовані методи, які згодом були модифіковані.</p> <p>b) Валідація має бути настільки обширною, наскільки це необхідно, і підтверджувати, шляхом надання об'єктивних доказів у формі функціональних характеристик, що конкретні вимоги до належного використання дослідження були виконані. Лабораторія повинна гарантувати, що ступінь валідації методу обстеження є достатнім для забезпечення постійної валідності результатів, що стосуються прийняття клінічних рішень.</p> <p>c) Персонал з відповідними повноваженнями та компетенцією має переглянути результати валідації та зафіксувати, що результати відповідають встановленим вимогам.</p> <p>d) Коли пропонуються зміни до валідованого методу обстеження, слід перевірити клінічний вплив і прийняти рішення щодо застосування модифікованого методу.</p> <p>e) Необхідно зберігати такі записи щодо валідації:</p> <p>1) використана процедура валідації;</p>

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<p>Якщо до валідованої методики дослідження вносяться зміни, вплив таких змін має бути задокументований і, за необхідністю, має бути проведена нова валідація.</p>	<p>2) особливі вимоги до використання за призначенням,</p> <p>3) визначення функціональних характеристик методу,</p> <p>4) отримані результати, і</p> <p>5) заява про валідність методу з детальним описом його придатності для використання за призначенням.</p>

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<p><b>5.5.1.2 Верифікація методик досліджень</b></p> <p>Валідовані методики досліджень, які використовуються без змін, підлягають незалежній верифікації лабораторією перед введенням у рутинне використання.</p> <p>Лабораторія повинна отримати інформацію від виробника/розробника методу для підтвердження функціональних характеристик методики.</p> <p>Незалежна верифікація лабораторією повинна підтвердити, за допомогою отримання об'єктивних доказів (у вигляді функціональних характеристик), що встановлені вимоги до функціональних характеристик методики дослідження виконані. Встановлені вимоги до функціональних характеристик методики дослідження, що підтверджуються в процесі верифікації, повинні визначатися можливістю використання результатів цього дослідження за призначенням.</p>	<p><b>7.3.2 Верифікація методів дослідження (Verification of examination methods)</b></p> <p>a) Лабораторія повинна мати процедуру верифікації того, що вона може належним чином виконувати методи дослідження перед введенням у використання, забезпечуючи досягнення необхідних функціональних характеристик, як зазначено виробником або методом.</p> <p>b) Функціональні характеристики методу дослідження, підтверджені під час процесу верифікації, повинні відповідати призначеному використанню результатів дослідження.</p> <p>c) Лабораторія повинна забезпечити достатній обсяг верифікації методів дослідження для забезпечення постійної достовірності результатів, що мають відношення до прийняття клінічних рішень.</p> <p>d) Персонал з відповідними повноваженнями та компетентністю повинен переглянути результати верифікації та зафіксувати, що ре-</p>

ДСТУ EN ISO 15189:2015	ISO 15189:2022
Лабораторія повинна задокументувати процедуру, використану для верифікації, і отримані результати. Уповноважений персонал повинен проаналізувати результати верифікації та задокументувати висновки.	<p>результати відповідають встановленим вимогам.</p> <p>е) Якщо метод переглядається розробником, лабораторія повторює верифікацію в необхідному обсязі.</p> <p>ф) Зберігаються такі записи щодо верифікації:</p> <ol style="list-style-type: none"> <li>1) Функціональні характеристики, яких необхідно досягти,</li> <li>2) отримані результати, і</li> <li>3) заява про те, чи були досягнуті визначені функціональні характеристики і, якщо ні, вжиті заходи.</li> </ol>

Валідація та верифікація методик вимірювань означає, що фахівці лабораторій використовують свою професійну оцінку, щоб судити, які аналітичні характеристики потрібно оцінити, щоб переконатися, що передбачувана якість може бути забезпечена.

Можна виділити такі різновиди верифікації:

**В залежності від стадії життєвого циклу методу:**

- при впровадженні методу;
- поточна;
- верифікація нового лоту реагентів.

**В залежності від вибору формату експериментів:**

- за результатами спланованих верифікаційних експериментів або
- з використанням наявних даних (внутрішньолабораторного контролю якості, міжлабораторних порівнянь, зовнішньої оцінки якості).

EFLM розроблено рекомендації з процедур валідації та верифікації в медичних лабораторіях (*Validation and verification of examination procedures in medical laboratories: opinion of the EFLM Working Group Accreditation and ISO/CEN standards (WG-A/ISO) on dealing with ISO 15189:2012 demands for method verification and validation (Roseri Roelofsen-de Beer et al., 2020)*). Процес оцінювання методу є стандартизованим та містить в собі етапи планування, виконання верифікаційних експериментів та прийняття висновку щодо придатності даного методу (рис. 3.5).

**Рекомендація I**

Лабораторія повинна використовувати лише процедури дослідження, які були валідовані або верифіковані лабораторією для використання за призначенням.



РИС. 3.5. ПРОЦЕС ВАЛІДАЦІЇ/ВЕРИФІКАЦІЇ МЕТОДИК ДОСЛІДЖЕНЬ  
(Validation and verification of examination procedures in medical laboratories: opinion of the EFLM WG-A/ISO)

**Рекомендація II**

Лабораторія повинна верифікувати процедури дослідження, отримані з таких джерел:

- нестандартні процедури дослідження;
- розроблені в лабораторії процедури дослідження;
- валідовані процедури експертизи, що використовуються поза призначенням;
- валідовані процедури експертизи, згодом модифіковані.

**Рекомендація III**

Якщо процедура дослідження змінена, повинно бути прийнято до уваги та задокументовано, чи є потенційні/відповідні наслідки для експлуатаційних характеристик і, якщо так, то які ці наслідки.

**Рекомендація IV**

Коли процедура дослідження змінюється згідно з задокументованими причинами, має бути виконана (додаткова) валідація або верифікація.

**Рекомендація V**

Процедури дослідження, які, на основі наявної документації, не обов'язково повинні бути валідовані, повинні бути принаймні верифіковані щодо відповідних характеристик.

**Рекомендація VI**

Якщо докази валідації, проведеної в іншому місці, є неповними, верифікації недостатньо, і необхідна додаткова валідація у власній лабораторії.

**Рекомендація VII**

При використанні більш ніж одного аналізатора для однієї і тієї ж вимірюваної величини, відповідна верифікація кожного окремого пристрою повинна бути виконана із застосуванням відповідних критеріїв прийнятності.

**Рекомендація VIII**

Валідація/верифікація має відбуватися відповідно до попередньо встановленого плану валідації/верифікації, який затверджений персоналом з відповідними повноваженнями, що містить принаймні такі елементи:

- цільове використання процедури дослідження;
- опис вимірюваної величини;
- вибір відповідних експлуатаційних характеристик;
- критерії прийнятності, дійсні для цільового використання;
- метод дослідження;
- особи, які проводитимуть дослідження та які будуть затверджувати результати.

**Рекомендація IX**

Під час верифікації співробітник, який має відповідні повноваження, повинен розглянути наступні аналітичні характеристики: прецизійність вимірювання, правильність вимірювань, межі виявлення, стабільність, референтний інтервал, порівняння з попереднім методом, якщо є, пороги медичних рішень та інтерференцію.

Такий розгляд має ґрунтуватися на професійній оцінці у світлі передбачуваного використання.

Прецизійність та порівняння з попереднім методом повинні бути оцінені локальними експериментами.

**Рекомендація X**

Під час валідації співробітник, який має відповідні повноваження, повинен враховувати всі експлуатаційні характеристики, наведені в ISO 15189:2012 (п.5.5.1.3).

**Рекомендація XI**

Якщо конкретні функціональні характеристики не є ні застосовними, ні можливими з огляду на характер процедури дослідження або поширеність патології, це має бути задокументовано та враховано.

**Рекомендація XII**

Отримані результати повинні бути задокументовані та збережені у звіті про валідацію/верифікацію принаймні протягом періоду, протягом періоду використання процедури дослідження

**Рекомендація XIII**

Отримані результати мають бути наочно оцінені з використанням критеріїв прийнятності, встановлених у плані валідації/верифікації.

Висновок оцінювання та те, чи підходить він для передбачуваного використання, повинні бути встановлені та задокументовані у звіті про валідацію/верифікацію.

Звіт про валідацію/верифікацію оцінюється та затверджується співробітником, який має відповідні повноваження.

**Рекомендація XIV**

Дозвіл на використання процедури дослідження на підставі відповідності критеріям прийнятності здійснюється працівником з відповідними повноваженнями із зазначенням дати в документації.

**Рекомендація XV**

Процедура дослідження, яка оцінюється, повинна бути задокументована у СОП відповідно до вимог ISO 15189.

**Рекомендація XVI**

При впровадженні методу повинна бути встановлена відповідна процедура забезпечення якості з межами прийнятності для внутрішнього та зовнішнього контролю якості, беручи до уваги межі прийнятності, що мають медичне значення.

**Рекомендація XVII**

Процес впровадження нової або зміненої процедури дослідження викладається в інструкції або контрольному списку.

**Рекомендація XVIII**

Коли зміни в процедурах експертизи вважаються важливими для замовника, має бути задокументовано, як замовники мають бути інформовані.

**Рекомендація XIX**

Повинна існувати процедура, за допомогою якої можна періодично перевіряти, чи дотримуються початкові критерії прийнятності для точності та істинності вимірювань, визначених під час валідації/верифікації. Якщо, подумавши, більш гнучкі критерії виправдані, це має бути обґрунтовано.

**Необхідні умови для проведення оцінювання методів**

- Визначені вимоги до аналітичних характеристик методики (див. розділ 4);
- Завершений тренінг персоналу та проведене первинне оцінювання компетентності персоналу, що буде брати участь в верифікаційних експериментах;
- Лабораторія забезпечена необхідними референтними матеріалами та зразками пацієнтів відповідно до обраного протоколу;
- Для атестованих референтних матеріалів наявні дані щодо їх невизначеності;
- Лабораторія забезпечена спеціалізованим програмним забезпеченням або підготувала робочі листи для обробки отриманих даних.

## 4. ВИБІР ПРИПУСТИМИХ СПЕЦИФІКАЦІЙ АНАЛІТИЧНИХ ХАРАКТЕРИСТИК

Процес вибору та оцінювання методики дослідження перед її використанням в лабораторії за призначення не можливий без розуміння вимог до її аналітичної якості або, іншими словами, без встановлення цілей (стандартів) якості. Управління аналітичною якістю в медичній лабораторії практично має розпочинатися саме із пошуку відповіді на питання про те, який рівень якості методики дослідження має бути досягнутий в конкретній лабораторії в конкретний проміжок часу. Звісно, цей рівень має бути визначено у кількісному форматі.

Найбільш поширеним способом встановлення вимог до аналітичної якості є використання допустимої загальної помилки (allowable total error, TEa).

В багатьох країнах світу значення TEa встановлюється та регулюється на державному рівні. Як приклад, нормативні документи США, зокрема Поправки щодо удосконалення клінічних лабораторій (Clinical Laboratory Improvement Amendments, CLIA), встановлюють єдині для всіх медичних лабораторій США стандарти щодо контролю та забезпечення якості їх роботи, в тому числі специфічні вимоги до аналітичної якості досліджень великого переліку аналітів.

Однак дедалі більшу популярність, особливо в країнах Європи, набуває використання науково обґрунтованих вимог до аналітичної якості.

На Першій стратегічній конференції Європейської федерації клінічної хімії та лабораторної медицини (EFLM) у 2014 році було запропоновано три незалежних моделі, які надаються в ієрархічному порядку. Ієрархічний підхід до визначення цілей аналітичної якості покликаний забезпечити пріоритет наукових доказів щодо встановлення вимог та максимально можливу об'єктивність при виборі і передбачає їх послідовний пошук від вищих за ієрархією моделей до нижчих.

### Модель 1. На основі впливу аналітичної якості на клінічні результати

Ця модель використовується для встановлення вимог до якості вимірювання аналітів, що відіграють центральну роль в діагностиці, скринінгу або моніторингу захворювань, тобто мають безпосередній вплив на прийняття клінічних рішень. Такий вплив можливо оцінити за допомогою різних видів досліджень:

- **1a. Дослідження прямих результатів** – дослідження впливу аналітичної ефективності тесту на клінічні результати.

- **1b. Дослідження непрямих результатів** – дослідження впливу аналітичної ефективності тесту на клінічні класифікації або рішення і, отже, на ймовірність правильної оцінки результатів обстежень пацієнта, наприклад, шляхом моделювання або аналізу рішень.

Аналіти, вимоги до аналітичної якості досліджень яких рекомендовано встановлювати згідно з цією моделлю, включають нейтрофільні лейкоцити, тромбоцити, гемоглобін, глікований гемоглобін, глюкозу, холестерин та його фракції, альбумін, тропонін тощо. Отримати таку інформацію можна в спеціалізованих базах даних, таких як Національна освітня програма з холестерину США (National Cholesterol Education Program, NCEP), Програма по стандартизації вимірювання HbA1c (National Glycohemoglobin Standardization Program, NGSP), Американська асоціація діабету (American Diabetes Association, ADA) та інші.

Так, наприклад, на основі результатів дослідження холестерину та його фракцій визначаються серцево-судинні ризики, пов'язані з гіперхолестеринемією. Для показників рівня холестерину, а також холестерину ліпопротеїнів високої щільності, холестерину ліпопротеїнів низької щільності та тригліцеридів NCEP визначено чіткі рівні прийняття клінічних рішень та встановлено цілі аналітичної якості у вигляді TEa, яка становить 8,9 %, 13 %, 12 % та 15 % відповідно.

Перевага **Моделі 1** полягає в тому, що вона стосується впливу аналітичної якості на клінічні результати, які мають значення для пацієнтів і суспільства. Основним недоліком є те, що вона корисна лише для випадків, де зв'язок між тестом, прийняттям клінічних рішень і клінічними результатами є простими і міцними.

Крім того, на аналітичні характеристики, отримані в дослідженнях прямих або непрямих результатів, часто впливає поточна якість вимірювань, і результати можуть відрізнятися залежно від фактично використаного методу тестування, досліджуваної популяції та закладів охорони здоров'я.

Ця модель не враховує аналітичні характеристики методів, доступних на сьогоднішній медичним лабораторіям, тому вимоги до якості, прийняті на її основі, можуть бути не завжди досяжними з використанням існуючих аналітичних систем.

### Модель 2. На основі компонентів біологічної варіації вимірюваної величини

В основі цієї моделі лежить розуміння природної фундаментальної властивості складу біоматеріалів людини – біологічної варіації

(biological variability, BV), яка віддзеркалює результати гомеостатичних процесів в організмі. Так, у стабільному стані вміст більшості вимірюваних величин характеризується випадковими коливаннями навколо певної гомеостатичної точки.

BV складається з двох компонентів: внутрішньоіндивідуального та групового. Внутрішньоіндивідуальна варіація (within subject variation) визначається як коливання вимірюваної величини навколо її гомеостатичної заданої точки в межах одного індивіда, тоді як міжіндивідуальна варіація (between subject variation) характеризує різницю між індивідуальними гомеостатичними точками у різних людей, що складають популяцію (рис. 4.1). Згідно з термінологією, рекомендованою А.М. Simundic та ін. [25], обидва компоненти виражаються у вигляді коефіцієнтів варіації – коефіцієнту внутрішньоіндивідуальної варіації (CV individual, CVi) та коефіцієнту групової варіації (CV group, CVg) відповідно.

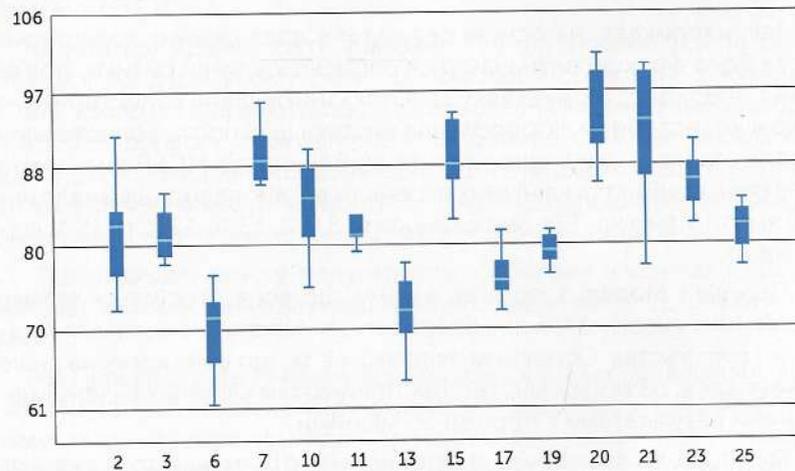


РИС. 4.1. ПРИКЛАД ДАНИХ ЩОДО БІОЛОГІЧНОЇ ВАРІАЦІЇ КОНЦЕНТРАЦІЇ КРЕАТИНІНУ (МКМОЛЬ/Л) В ПЛАЗМІ КРОВІ 25 ВОЛОНТЕРІВ (BAYSOY ET AL., 2022)

Величина BV залежить, перш за все, від фізіологічної функції, яку виконує в організмі речовина, яка аналізується. Найменша біологічна варіація виявлена у речовин з найкращою регуляцією їх рівня в організмі (натрій, хлориди, кальцій, магній, альбумін, загальний білок). Варіація середнього ступеня характерна для речовин, що беруть участь у процесах метаболізму: глюкоза, холестерин, фосфор. Найбільша біологічна варіація виявлена для компонентів сироватки, що є кінцевими продуктами катаболізму, або для ферментів: сечова кислота, сечовина, креатинін, АЛТ, АСТ.

BV має безліч застосувань у клінічній лабораторії, включаючи внутрішньолабораторний контроль якості аналітичних процесів, зовнішнє оцінювання якості, верифікацію та валідацію аналітичних методів тощо. Величини компонентів BV використовуються також і для встановлення вимог до аналітичної якості. Саме в цьому полягає суть Моделі 2, коли дані щодо BV використовуються для встановлення рівнів гранично допустимих випадкової аналітичної похибки (допустимого коефіцієнту аналітичної варіації, CVa) та систематичної аналітичної похибки (у вигляді показника зсуву, B), на основі яких розраховується показник TEa (табл. 4.1). Інформацію про CVi, CVg та, відповідно, TEa можна знайти в різних базах даних. З моменту, коли концепція біологічної варіації була вперше запропонована С. Fraser та Е. Harris [26], проведено велику кількість експериментальних досліджень різного рівня складності та отримано чисельні оцінки BV для різноманітних популяцій та умов. І протягом останніх десятиліть користувачам були доступні різні джерела даних BV.

ТАБЛИЦЯ 4.1. РОЗРАХУНОК ВИМОГ ДО ТОЧНОСТІ НА ОСНОВІ ДАНИХ БІОЛОГІЧНОЇ ВАРІАЦІЇ (FRASER, 2001)

Рівень прийнятності	Гранично допустимі похибки		
	випадкова, CVa	систематична, B	загальна, TEa
Мінімальний	$CVa < 0,75 * CVi$	$B < 0,375(CVi^2 + CVg^2)^{1/2}$	$TEa = 1.65 * CVa + B$
Базовий	$CVa < 0,50 * CVi$	$B < 0,25(CVi^2 + CVg^2)^{1/2}$	$TEa = 1.65 * CVa + B$
Оптимальний	$CVa < 0,25 * CVi$	$B < 0,125(CVi^2 + CVg^2)^{1/2}$	$TEa = 1.65 * CVa + B$

Примітка: В формулі для розрахунку TEa 1,65 – Z-квантиль нормального розподілу, який забезпечує 95 %-ий односторонній довірчий інтервал для випадкової похибки.

Найбільш відомою для українських фахівців галузі лабораторної медицини є база даних Кармен Рікос, яка й досі доступна за посиланням <https://www.westgard.com/biodatabase1.htm>, але з 2014 р. не підлягає оновленню.

На Першій стратегічній конференції EFLM у 2014 році було визнано, що велика частина наявних даних щодо BV є скомпрометованою через невизначеність вимірювань або інші фактори. За результатами конференції різні робочі групи EFLM започаткували ініціативи щодо поліпшення методології отримання, розрахунку та оцінювання даних BV і похідних параметрів. Результати багатьох із цих ініціатив були включені в базу даних біологічної варіації EFLM, створену в рамках проєкту з

вивчення біологічної варіації EuBIVAS (The European Biological Variation Study). Ця база даних доступна для користувачів у всьому світі за посиланням [www.biologicalvariation.eu](http://www.biologicalvariation.eu). База постійно оновлюється.

На додаток до даних BV, база даних EFLM також забезпечує автоматичний розрахунок критерію критичної різниці (Reference Change Value, RCV), а також аналітичних специфікацій продуктивності (Analytical Performance Specification, APS), зокрема показників загальної допустимої помилки (Total allowable error) та гранично допустимої невизначеності вимірювання (Maximum allowable measurement uncertainty). Для отримання інформації щодо APS або RCV необхідно натиснути на відповідний заголовок, як це показано на рис. 4.2 на прикладі креатиніну.

Matrix	BV Estimate	median CV estimate	lower CI limit	higher CI limit	Comments	Date Updated
Serum/plasma	4.5	14.1	4.2	5.7		13/4/2023
Serum/plasma			7.0	17.4		13/4/2023

РИС. 4.2. СКРІНШОТ СТОРІНКИ БАЗИ ДАНИХ EFLM З ІНФОРМАЦІЄЮ ПРО CV<sub>i</sub> (ЧЕРВОНА СТРІЛКА) ТА CV<sub>c</sub> (ЧОРНА СТРІЛКА) КРЕАТИНІНУ В СИРОВАТЦІ ТА ПЛАЗМІ КРОВІ

В 2022 році перелік аналітичних специфікацій був доповнений показниками **максимально припустимої розширеної невизначеності (Maximum allowable measurement uncertainty, MAU)**:

$$MAU = k \times F \times CV_i$$

де  $k$  – коефіцієнт покриття (2 або 3),  $F$  – рівень вимог до якості (0,5 – бажаний, 0,75 – мінімальний),  $CV_i$  – внутрішньоіндивідуальний коефіцієнт варіації.

На цей час рекомендується використовувати мінімальні (!) вимоги – (фактор 0,75).

Specification	Imprecision (CV <sub>a</sub> )	BIAS	MAU	Total Error
Minimum	3.4	5.6	6.8	11.1
Desirable	2.3	3.7	4.5	7.4
Optimal	1.1	1.9	2.3	3.7

РИС. 4.3. СКРІНШОТ СТОРІНКИ БАЗИ ДАНИХ EFLM З ІНФОРМАЦІЄЮ ПРО ПРИПУСТИМІ АНАЛІТИЧНІ СПЕЦИФІКАЦІЇ КРЕАТИНІНУ В СИРОВАТЦІ ТА ПЛАЗМІ КРОВІ

Отже, **Модель 2** передбачає встановлення вимог до аналітичної якості вимірювань на основі компонентів біологічної варіації величин, що вимірюються. Модель застосовується у випадках, коли лабораторний показник не відіграє провідної ролі у прийнятті клінічного рішення і для нього доступні дані BV. Останні реально встановити для більшості вимірюваних в медичних лабораторіях величин, що надає перевагу цій моделі. Крім того, відсутність чітких відмінностей між досліджуваними когортами з Туреччини, Норвегії, Нідерландів, Іспанії та Італії підтверджує, що отримані дані можна переносити між системами охорони здоров'я, що уможлиблює їх використання на міжнародному рівні [27, 28].

Можливим обмеженням Моделі 2 є умови отримання даних BV. Так, дані BV визначаються, зокрема в рамках проекту EuBIVAS, в суворо контрольованих умовах відповідно до найкращих практик (жорсткий контроль переданалітичних факторів, використання сучасних методів дослідження тощо), що може робити їх «занадто гарними», тобто такими, що не відображають «реальної» ситуації. Недоліком цієї моделі є також нестача відповідної інформації щодо різних вікових груп та стану здоров'я осіб в досліджуваних популяціях.

### Модель 3. На основі найсучаснішого рівня технологій (state-of-the-art)

Цю модель застосовують у випадках, коли перші дві моделі не можуть бути використані. Вона базується на наявному рівні аналітичної якості, який відображає реальний стан розвитку лабораторних технологій. Для встановлення такого рівня існують різні підходи, коли використовуються дані, які віддзеркалюють:

- найвищий рівень аналітичної ефективності, технічно досяжний на даному етапі розвитку лабораторних технологій;
- аналітичну якість певної частки найкращих лабораторій в національних програмах зовнішнього оцінювання якості;
- середній рівень аналітичної якості, що декларується для конкретної методики найбільш популярними виробниками.

Крім того, існує декілька визнаних у світі систем зовнішнього оцінювання якості, що визначають вимоги до аналітичної якості та можуть бути прийняті за основу при встановленні вимог до точності результатів лабораторних досліджень, зокрема:

- критерії програм перевірки кваліфікації CLIA, США;
- вимоги до аналітичної якості результатів лабораторних досліджень Федеральної медичної Ради ФРН (Richtlinien der Bundesärztekammer, RiliBak);
- вимоги до гранично допустимої загальної аналітичної помилки вимірювань Королівської колегії патологів Австралазії (Royal College of Pathologists of Australasia, RCPA).

Перевагою **Моделі 3** є те, що вона базується на аналітичних можливостях сучасних медичних лабораторій, отже, забезпечення відповідності вимогам до точності, встановленим на її основі, є абсолютно реальною задачею (на відміну, наприклад, від Моделі 2, коли вимоги до точності вимірювання аналітів із низькою BV є дуже жорсткими та не завжди досяжними на сучасному етапі розвитку лабораторних технологій). В той же час недоліком цієї моделі є те, що зв'язок між «технічно досяжним» та тим, що дійсно необхідно для забезпечення точності вимірювання та поліпшення клінічного результату, не завжди наявний.

Слід зазначити, що всі три зазначені вище моделі використовують різні принципи, а отже, мають свої переваги та недоліки. Деякі моделі краще підходять для одних вимірювань, ніж для інших. Тому рекомендується скласти список вимірюваних величин з їх розподілом для різних моделей. Перевагу слід віддавати моделям 1 і 2.

В більшості випадків кожний аналіт необхідно «закріпити» за однією моделлю. Хоча інколи може бути виправданим поєднання різних моделей. Так, наприклад, в залежності від клінічного застосування одна й та сама вимірювана величина може мати різні специфікації ефективності, що обумовлює необхідність застосування різних вимог до точності. Гарним прикладом є дослідження рівня глюкози в трьох різних клінічних ситуаціях: у реанімаційних пацієнтів, з метою самостійного моніторингу глюкози хворими на цукровий діабет 1 типу, та в рамках проведення діагностики різних захворювань. В даному випадку вимоги до точності вимірювання рівня глюкози можуть бути визначені шляхом моделювання впливу тесту на ймовірні результати пацієнта (модель 1b), за допомогою клінічних досліджень результатів (модель 1a) та на основі компонентів біологічної варіації (модель 2) відповідно.

Вимоги до точності результатів лабораторних досліджень мають бути задокументовані та супроводжуватися поясненнями та доказами ефективності обраних припустимих аналітичних специфікацій в конкретних умовах застосування. Порівняння рекомендацій щодо припустимих аналітичних специфікацій для деяких тестів наведені у додатку 8.1 цього Керівництва.

Якщо для вимірюваної величини наразі не встановлені вимоги до аналітичної якості у форматі TEa, можливі альтернативні підходи до їх визначення.

Зокрема, можливо використання припустимого коефіцієнту варіації, зазначеного виробником відповідної аналітичної системи, в якості гранично допустимої випадкової аналітичної похибки.

Також нещодавно запропоновано підхід до розрахунку допустимих меж для аналітичної похибки, заснований на емпіричних біологічних варіаціях, які, у свою чергу, виводяться з референтного інтер-

валу відповідного аналіту. Перевага цього підходу полягає в тому, що вимоги враховують нелінійний зв'язок між біологічними та аналітичними варіаціями. Прийнятна випадкова аналітична похибка ( $CV_A$ ) розраховується наступним чином:

$$CV_A = (CV^*_{\epsilon} - 0.25)^{0.5}$$

$CV^*_{\epsilon}$  виводиться з логарифмічних значень меж референтного інтервалу ( $RL_{lower}$ ,  $RL_{upper}$ ), оскільки більшість аналітів у клінічній хімії мають перекошений розподіл, і коли справжній розподіл невідомий, логарифмічний розподіл є найкращим припущенням:

$$CV^*_{\epsilon} = 100(\exp S_{E, \ln}^2 - 1)^{0.5}$$

$$S_{E, \ln} = \frac{(\ln RL_{upper} - \ln RL_{lower})}{3.92}$$

Якщо величина має кілька еталонних інтервалів, автори пропонують вибрати інтервал із найменшим діапазоном. Для аналітів із невідомою нижньою межею референтного інтервалу нижня межа повинна бути встановлена на 15 % від верхньої межі референтного інтервалу.

Припустима величина зсуву (pB) та припустима розширена невідомість (U) за цією моделлю розраховується за формулою:

$$pB = 0.7 * CV_A$$

$$U = 2.39 * CV_A$$

Деякі вимірювані величини можуть мати різні специфікації ефективності, в залежності, наприклад, від того, чи тест має кілька призначених клінічних застосувань. Наприклад, для рівня глюкози в крові характеристики можуть бути визначені в умовах реанімації шляхом моделювання впливу тесту на ймовірні результати пацієнта (модель 1b), для його самостійного моніторингу при цукровому діабеті 1 за допомогою клінічних досліджень результатів (модель 1a) або більш загальним підходом, заснованим на біологічній варіації (модель 2).

Враховуючи, що існують визнані у світі вимоги до точності вимірювань в медичних лабораторіях (МЛ), керівник МЛ може вибрати «дизайн» точності, наприклад, заснований на біологічній варіації (BV), на вимогах RILIBAEK (Федеральна Медична Рада ФРН – Richtlinien der Bundesärztekammer), CLIA (Clinical Laboratory Improvement Amendments) або інших обґрунтованих нормах точності, що задокументовані та погоджені із керівництвом закладу у встановленому порядку.

Порівняння рекомендацій щодо припустимих аналітичних специфікацій для деяких тестів наведені в додатку 8.1 до Керівництва.

## 5. ВЕРИФІКАЦІЯ КІЛЬКІСНИХ МЕТОДІВ ДОСЛІДЖЕНЬ

Верифікація методу дослідження передбачає, що в лабораторії встановлені вимоги до його функціональних характеристик. Верифікація проводиться з метою підтвердження того, що функціональні характеристики обраного методу відповідають цим попередньо встановленим вимогам.

На практиці верифікація методу переважно полягає у проведенні в лабораторії певних заходів, які дозволяють перевірити, чи підтверджуються в умовах конкретної лабораторії встановлені виробником функціональні характеристики методу. Для цього лабораторія має отримати від виробника/розробника методу відповідну інформацію.

Функціональні характеристики кількісних методів, що підлягають верифікації в медичних лабораторіях, включають:

- прецизійність (precision);
- правильність (trueness);
- референтні значення (reference intervals);
- аналітичну специфічність (analytical specificity);
- аналітичну чутливість (analytical sensitivity);
- перенесення (carryover);
- лінійність (linearity).

Верифікація всіх вищезазначених функціональних характеристик не завжди можлива та виправдана з практичної точки зору. Вибір характеристик, які мають пройти незалежну верифікацію лабораторією, залежить від методу, задачі лабораторного дослідження та, найголовніше, від рівня ризику заподіяння шкоди пацієнту внаслідок неправильної діагностики при отриманні недостовірного результату дослідження.

Мінімальні вимоги до верифікації кількісних методів дослідження передбачають оцінювання їх прецизійності та правильності.

### 5.1. ВЕРИФІКАЦІЯ ПРЕЦИЗІЙНОСТІ ТА ПРАВИЛЬНОСТІ

Існують багато варіантів дизайну експериментів для оцінювання прецизійності та правильності.

Можливо проведення окремих експериментів з оцінки прецизійності та правильності, при яких ці характеристики оцінюються окремо з використанням різних контрольних матеріалів та зразків пацієнтів. Саме такий дизайн застосовано у протоколі CLSI EP15-A2 «User Verification of Performance for Precision and Trueness».

Можливі також експерименти, які дозволяють перевірити обидва компоненти одночасно і потім інтерпретувати результати за допомогою добре збалансованої статистичної моделі. Такий підхід пропонується протоколом CLSI EP15-A3 «User Verification of Precision and Estimation of Bias» (далі – CLSI EP15-A3).

Відомим також є підхід, коли верифікація прецизійності проводиться з використанням даних довготривалого внутрішньолaboratorного контролю якості без організації та проведення додаткових спеціальних верифікаційних експериментів.

При виборі дизайну верифікаційного експерименту лабораторії мають керуватися власними технічними можливостями для отримання надійних статистично достовірних результатів.

#### 5.1.1. ВЕРИФІКАЦІЯ ПРЕЦИЗІЙНОСТІ ТА ПРАВИЛЬНОСТІ ЗА ПРОТОКОЛОМ CLSI EP15-A3

Відповідно до протоколу CLSI EP15-A3, повторне вимірювання принаймні двох зразків з різними вимірюваними концентраціями здійснюється протягом 5, 6 або 7 днів із загальною обробкою п'яти вимірювань на день за одну серію. Якщо використовуються зразки з відомими концентраціями (наприклад, зразки для контролю якості або сертифіковані еталонні матеріали), результати одного експерименту можна оцінити як щодо прецизійності, так і щодо правильності.

#### Верифікація прецизійності методу користувачем за протоколом CLSI EP15-A3

Передбачено тестування двох та більше рівнів контрольних або інших референтних матеріалів (якщо оцінювання прецизійності та зсуву планується проводити в одному експерименті) або зразків пацієнтів (якщо дизайн верифікаційного дослідження передбачає оцінювання прецизійності та зсуву в двох окремих експериментах).

#### Основний протокол:

5 днів тестування x 5 реплікацій на день для кожного зразка.

#### Повинні бути використані:

- один реагент;
- один калібратор;
- одне калібрування (оцінювання протягом 5 днів);
- одна інструментальна система.

#### Вимоги до зразків:

- мінімальна кількість зразків – 2 зразки;
- кілька рівнів прийняття клінічних рішень;
- нормальний та патологічний вміст аналіту;

- відповідає рівню прецизійності відповідно до вимог виробника;
- можливе використання зразків пацієнтів чи пулів зразків;
- можливе використання контрольних матеріалів відповідно до вимог виробника;
- дотримання вимог зберігання зразків із забезпеченням їх стабільності протягом всього дослідження (наприклад, приготувати та заморозити аліквоти, розморожування в день проведення тесту).

#### Процедура проведення:

- відкалібруйте процедуру вимірювання;
- кожного дня протягом 5 днів виконуйте по 5 вимірювань кожного зразка;
- паралельно виконуйте процедури внутрішньолабораторного контролю якості (ВЛКЯ);
- переглядайте результати щоденного ВЛКЯ, щоб впевнитись у прийнятності результатів вимірювань;
- проводьте своєчасне усунення невідповідностей та забезпечте належне документування;
- переглядайте результати верифікаційного експерименту щодня для своєчасного виявлення відсутності або неповних даних, або нетипових результатів;
- ідентифікуйте потенційні викиди за допомогою будь-якого відповідного тесту, наприклад, тесту Граббса для викидів.

Оцінювання прецизійності здійснюється за допомогою однофакторного дисперсійного аналізу (One-Way Analysis of Variance,

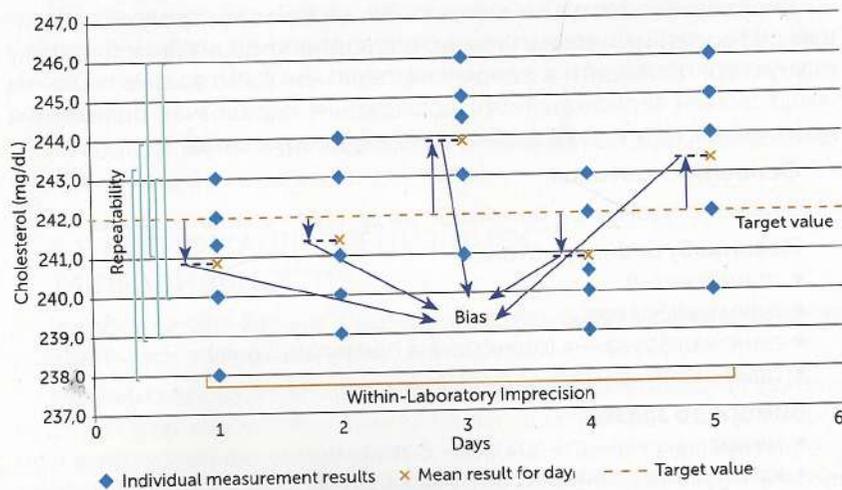


РИС. 5.1. ОГЛЯД РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТУ ЗА CLSI EP15-A3

ANOVA). Якщо обчислення не виконуються за допомогою спеціального програмного забезпечення, для проведення ANOVA можна використовувати майже будь-яке доступне програмне забезпечення для статистики, включаючи вбудований у Microsoft Excel «Analysis Tool Pak». Результатом оцінки є підсумкова таблиця, що відображає загальну мінливість набору даних, поділену на внутрішньогрупову та міжгрупову фракції.

Source of variation	Sum of squares (SS)	df	Mean square (MS)
Between-group	SS <sub>1</sub>	df <sub>1</sub>	MS <sub>1</sub>
Within-group	SS <sub>2</sub>	df <sub>2</sub>	MS <sub>2</sub>
Total	SS <sub>total</sub>	df <sub>total</sub>	

РИС 5.2. ВИГЛЯД ЗВІТУ ANOVA

Використовуючи ці результати, обчислюються три складові прецизійності: CV в межах циклу ( $CV(\%)_W$ ), CV між циклами ( $CV(\%)_B$ ) і внутрішньолабораторний CV ( $CV(\%)_{WL}$ ) за формулами:

$$CV(\%)_W = \frac{\sqrt{MS_2}}{\bar{x}} \times 100$$

$$CV(\%)_B = \frac{\left( \frac{MS_1 - MS_2}{n_0} \right)^{0.5}}{\bar{x}} \times 100$$

$$CV(\%)_{WL} = \frac{(V_W + V_B)^{0.5}}{\bar{x}} \times 100$$

Інтерпретація результатів оцінювання прецизійності проводиться шляхом порівняння розрахованих показників прецизійності методу із відповідними специфікаціями (заявами) виробника (manufacturer claims). Якщо отримані показники прецизійності методу не перевищують специфікації виробника, останні вважаються верифікованими.

Однак, оскільки експеримент складається з відносно невеликого розміру вибірки, ймовірність того, що цей експеримент випадково провалиться, досить висока. Насправді, якщо середня непрецизійність дорівнює заявленій виробником, непрецизійність користувача все одно перевищуватиме заявлену у 50 % випадків. З цієї причини CLSI рекомендує обчислювати верхню межу перевірки (UWL), яка являє собою верхній 95-й процентиль очікуваної непрецизійності, для експериментів, подібних за розміром і дизайном до тих, які виконують

виробник. Отже, ймовірність випадкового провалу експерименту зменшується до 5 %. Застосовані розрахунки є досить складними і реалізуються наведеним у додатку до даного Керівництва робочим інструментом або іншим спеціалізованим ПЗ згідно з протоколом CLSI EP 15-A3.

Проте, з практичних міркувань, UVL можна встановити на 30 % вище, ніж заявляє виробник (при дотриманні дизайну експериментів протоколу CLSI EP 15-A3). З огляду на це, спостережувана неprecизійність може бути вищою, ніж заявлена; однак, якщо вона все ще менша від UVL, специфікація виробника щодо прецизійності вважається верифікованою.

### Оцінювання зсуву

Оцінювання зсуву середнього значення результатів вимірювання аналізу в матеріалах з відомими концентраціями (таких як зразки, призначені для контролю якості, матеріали для перевірки кваліфікації або сертифіковані референтні матеріали) відносно відповідного, заздалегідь визначеного цільового значення. Якщо припустити, що такий матеріал використовувався для експерименту з оцінювання прецизійності, описаного вище, результати вимірювань, отримані для цього експерименту, також можна використовувати для розрахунку зсуву.

Оцінити зсув користувач може за різними сценаріями EP15-A3, в залежності від типу використаного матеріалу для експерименту:

- Сценарій А Матеріал з приписаною невизначеністю, CRM;
- Сценарій В Матеріал зовнішньої оцінки якості (ЗОЯ);
- Сценарій С Матеріал групи порівняння;
- Сценарій Е Матеріал внутрішнього контролю якості (ВЛКЯ) з невідомою невизначеністю.

### Основний протокол:

5 днів тестування  $\times$  5 реплікацій на день для кожного зразка.

### Повинні бути використані:

- один реагент;
- один калібратор;
- одне калібрування (оцінювання протягом 5 днів);
- одна інструментальна система.

### Вимоги до зразків:

- мінімум – один зразок (бажано більше) із визначеним вмістом аналізу – з цільовим значенням (TV) та невизначеністю цільового значення TV ( $se_{RM}$ ), функцію якого можуть виконувати:
  - сертифіковані референтні матеріали (Certified Reference Material, CRM) з приписаною невизначеністю (сценарій А);
  - матеріали, отримані з програм ЗОЯ, де цільове значення оцінювалося за допомогою референтної процедури вимірювання (сценарій В);

- матеріали, отримані з програм міжлабораторного контролю якості або екстерналізованого внутрішньолабораторного контролю якості (interlaboratory QC program), бажано з кількістю учасників понад 20 (сценарій С);
- матеріали, що використовуються для ВЛКЯ з невідомою невизначеністю (сценарій Е);
- також для окремих застосувань можна використовувати пули зразків пацієнтів, добре статистично охарактеризовані (для порівняння між центральною та допоміжними лабораторіями наприклад) або досліджені за допомогою референтних процедур;
- дотримання вимог зберігання зразків із забезпеченням їх стабільності протягом всього дослідження (наприклад, приготувати та заморозити аліквоти, розморожування в день проведення тесту).

**Процедура проведення верифікаційного експерименту із оцінювання правильності методу повністю відповідає описаній вище процедурі оцінювання прецизійності.**

**Оцінювання зсуву** передбачає, перш за все, отримання двох статистичних характеристик:

- загального середнього значення результатів вимірювання ( $\bar{x}$ );
- стандартної помилки загального середнього значення результатів вимірювання ( $SE_{\bar{x}}$ ), яка розраховується наступним чином:

$$df_{\bar{x}} = nRun - 1 \quad S_{WL} = (V_W + V_B)^{0.5}, \quad V_W = MS_2$$

$$SE_{\bar{x}} = \frac{S_{WL}}{(df_{\bar{x}} + 1)^{0.5}} \quad V_B = \frac{(MS_1 - MS_2)}{n_0}$$

де  $V_B$  – міжсерійний компонент варіації,  $V_W$  – внутрішньосерійний компонент варіації,  $S_{WL}$  – внутрішньолабораторна варіація,  $df$  – ступені свободи для реплікаційного експерименту

Щоб оцінити прийнятність зсуву, виводиться верифікаційний інтервал перевірки, який вимагає попереднього обчислення третьої статистичної величини – стандартної помилки цільового значення ( $SE_{TV}$ ). Розрахунок  $SE_{TV}$  залежить від типу матеріалу, що використовувався у верифікаційному експерименті:

- для матеріалів з приписаною невизначеністю (сценарій А)  $SE_{TV}$  дорівнює стандартній невизначеності приписаного значення матеріалу, а ступені свободи розраховуються за формулою:

$$df_C = df_{\bar{x}} \cdot \left( se_C / se_{\bar{x}} \right)^4 \quad df = (Runs - 1) \times \left( \frac{SE_C}{SE_{\bar{x}}} \right)^4$$

- для матеріалів, отриманих з програм ЗОЯ (сценарій В), де цільове значення оцінювалося за допомогою референтної процедури вимірювання або на основі статистики результатів учасників групового порівняння (сценарій С),  $SE_{TV}$  розраховується таким чином:

$$SE_{TV} = \frac{SD_{PT}}{(n_{Lab})^{0.5}}$$

де  $SD_{PT}$  – SD, розраховане на основі результатів учасників програми ЗОЯ;  $n_{Lab}$  – кількість лабораторій, що брали участь у програмі ЗОЯ.

Ступені свободи в такому випадку розраховуються за формулою Саттертвейта.

$$df_C = \frac{\left(\frac{se_{\bar{x}}^2}{df_{\bar{x}}} + \frac{se_{RM}^2}{df_{RM}}\right)^2}{\frac{se_{\bar{x}}^4}{df_{\bar{x}}^3} + \frac{se_{RM}^4}{df_{RM}^3}}$$

Протокол CLSI EP15-A3 містить додаткові таблиці (15 A-C), які полегшують розрахунок ступенів свободи, оминаючи формулу Саттерсвейта. Також для рутинних лабораторних досліджень можна застосувати дещо більш консервативний підхід, обчислюючи ступені свободи як:  $(n_{Run}-1)$  або  $(n_{Lab}-1)$ .

- для матеріалів, що використовуються для ВЛКЯ з невідомою невизначеністю (сценарій Е),  $SE_{TV} = 0$ , а ступені свободи обчислюються наступним чином:

$$df_C = df_{\bar{x}}$$

або:  $df = n_{Run} - 1$ , де  $n_{Run}$  – число серій експерименту.

На основі результатів оцінювання  $SE_{\bar{x}}$  та  $SE_{TV}$  розраховується комбінована стандартна помилка ( $SE_C$ ) за формулою:

$$SE_C = \left(SE_{\bar{x}}^2 + SE_{TV}^2\right)^{0.5}$$

Нарешті, розраховується верифікаційний інтервал (VI):

$$VI = TV \pm (t_{0.975, df} \times SE_C)$$

де  $TV$  – цільове значення (Target Value) аналізу у використаному контрольному матеріалі;  $t_{0.975, df}$  – критичне значення  $t$  для відповідних рівня значущості та числа ступенів свободи.

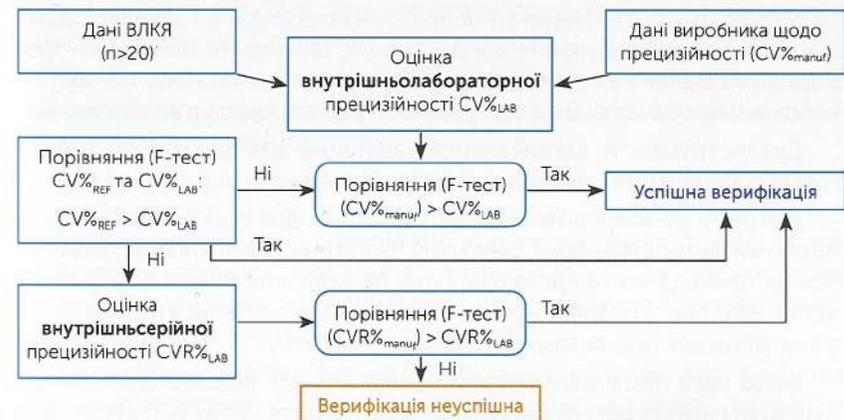
Інтерпретація результатів оцінювання зсуву проводиться шляхом співставлення загального середнього значення результатів вимірю-

вання із розрахованим верифікаційним інтервалом. Якщо VI включає  $\bar{x}$ , зсув не є статистично значущим, і правильність методу в такому випадку можна вважати верифікованою.

Якщо  $\bar{x}$  лежить поза межами розрахованого VI, зсув вважається таким, що статистично відрізняється від нуля. У цьому випадку необхідно визначити, чи є відхилення клінічно значущим. Для цього необхідно розрахувати показник аналітичного зсуву (Bias) у вигляді відносної величини та порівняти її із припустимою величиною аналітичного зсуву, встановленою у лабораторії. Якщо розрахований зсув не перевищує припустиму величину аналітичного зсуву, можна вважати, що верифікація правильності методу відбулася (навіть за умови, коли є статистично значущий зсув). У деяких випадках може знадобитися звернутися за допомогою до виробника.

### Поточна верифікація прецизійності з використанням даних довготривалого внутрішньолабораторного контролю якості:

Antonelli et al. (2017) були запропоновані кілька сценаріїв використання даних проведення ВЛКЯ та ЗОЯ для проведення поточної верифікації аналітичних характеристик методу. Один з них призначений для верифікації прецизійності методики, беручи до уваги довготривалий коефіцієнт варіації методики, який можна оцінити за даними ВЛКЯ  $CV\%_{LAB}$ , порівнюючи його з специфікаціями виробника  $CV\%_{manuf}$  та вимогами щодо прецизійності  $CV\%_{REF}$ :



## 5.2. ЕКСПЕРИМЕНТ З ПОРІВНЯННЯ МЕТОДІВ

Дослідження порівняння методів проводяться для оцінювання наявності зсуву (або у випадках, коли жоден із методів не є еталонним методом, різниці) між двома методами, які вимірюють один і той самий аналіт.

Як правило, це робиться перед заміною існуючої процедури вимірювання новою або під час додавання процедури вимірювання до існуючого тестового меню. Метою експерименту порівняння методів є визначення того, чи дає метод-кандидат аналогічні результати (у межах статистичної потужності експерименту), що й метод порівняння (існуючий). Це кількісно оцінюється в показниках зсуву (або різниці) і допоможе визначити, чи потрібно вносити зміни у звіти результатів, отриманих за допомогою процедури-кандидата, наприклад, зміна референтних діапазонів або меж для прийняття клінічних рішень.

Для проведення експерименту протокол CLSI EP 09-A3 рекомендує аналізувати щонайменше 40 зразків за допомогою обох процедур. Діапазон значень є більш важливим, ніж абсолютна кількість зразків, і в ідеалі він повинен охоплювати весь діапазон аналітичних вимірювань. Як правило, зразки аналізують як одиничні вимірювання, хоча середній результат кількох повторень зразка можна використовувати для мінімізації невизначеності оцінювання зсуву.

Бажаним є використання свіжих немодифікованих зразків, які аналізуються одночасно обома методами. Однак припустимим є використання збережених (з оглядом на стабільність) зразків та до 20 % модифікованих (розведених або з додатком), які аналізуються після певного періоду часу.

Після аналізу зразків і збору даних результати необхідно візуалізувати.

Метою цього є визначення аберантних результатів і викидів, оцінка діапазону випробовуваних концентрацій, виявлення можливих пропорційних (залежних від концентрації) і систематичних (постійних) помилок вимірювань і оцінка відповідності двох процедур вимірювання.

Два інструменти, які виявилися надійними для цієї мети, – це діаграма розсіювання (скатерограма) та діаграма різниці.

Діаграми розсіювання використовуються для опису кореляції між парними вимірюваннями у всьому діапазоні вимірюваних значень. Кожна точка даних представляє одну пару вимірювань. Зазвичай значення еталонного (або існуючого) методу відкладається вздовж осі x, а значення для порівняльного методу – по осі y.

Якщо були проведені повторні вимірювання, рекомендується побудувати лінію середніх значень двох результатів. У тих випадках, коли для кожного аналізу проводять три або більше вимірювань, замість середнього слід використовувати медіанне значення.

Якщо на діаграмі розсіювання виявлені очевидні відхилення або інші проблеми, їх слід усунути, перш ніж продовжити експеримент. Аналогічно, якщо діаграма розсіювання показує, що результати недостатньо охоплюють аналітичний діапазон, слід провести додаткові

вимірювання. Статистичний аналіз даних здійснюється за допомогою кореляційного та регресійного аналізу.

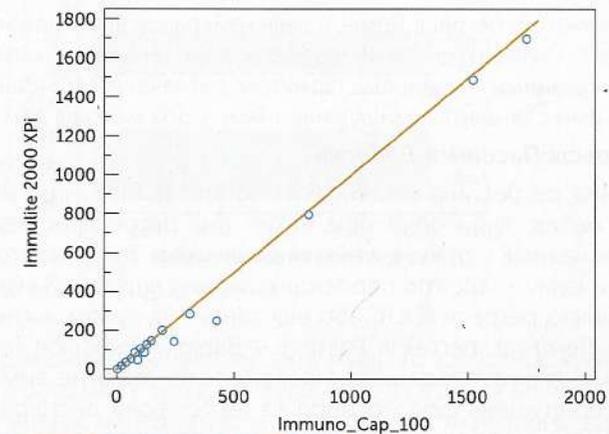


РИС. 5.3. СКАТЕРОГРАМА (D. VULJANIC, 2021)

**Кореляція** – це статистичний метод, який використовується для демонстрації зв'язку між парами змінних і для кількісної оцінки сили цього зв'язку.

Серед низки різних доступних методів кореляції найпоширенішими є кореляція Пірсона, або кореляція продукт-момент, кореляція рангового порядку Спірмена та тау Кендалла.

Числові значення коефіцієнтів кореляції завжди коливаються від -1,0 до 1,0, причому значення, ближчі до -1,0 або 1,0, вказують на сильну лінійну залежність.

Як зазначили Бланд і Альтман, кореляція не обов'язково означає згоду. Ідеальна згода досягається лише в тому випадку, якщо точки на діаграмі розсіювання розподілені вздовж лінії рівності, тоді як для досягнення ідеальної кореляції достатньо, щоб точки лежали тільки вздовж прямої. Тому кореляційний аналіз використовується тільки разом з регресійним та іншими техніками.

**Регресійний аналіз** – це методика, за допомогою якої лінія, яка найкраще передбачає залежність однієї змінної від іншої, підлаштовується до даних.

Доступні численні регресійні моделі, а ті, які найчастіше використовуються в умовах клінічної лабораторії, описані нижче.

### 1. Звичайна регресія найменших квадратів (OLR)

Звичайна регресія найменших квадратів оцінює лінію найкращого підходу шляхом мінімізації суми квадратів різниць між прогнозовани-

ми значеннями (на прямій) і спостережуваними значеннями у вертикальному напрямку.

## 2. Регресія Демінга

За допомогою регресії Демінга лінія найкращої відповідності оцінюється шляхом мінімізації суми квадратів відхилень між фактичними та спостережуваними значеннями під кутом, визначеним відношенням між аналітичними стандартними відхиленнями обох методів вимірювання.

## 3. Регресія Пассинга-Баблока

Описана регресійна модель Passing and Bablok – це непараметричний метод, який обчислює нахил лінії регресії як медіану всіх можливих нахилів і, отже, є стійким до викидів. Замість того, щоб мінімізувати суму квадратів перпендикулярних відстаней між точками даних і лінією регресії (OLR) або відстаней під кутом, визначеним  $\lambda$  (регресія Демінга), регресія Passing – Bablok обчислює нахили між будь-якими двома точками даних. Крім того, вона не вимагає спеціальних припущень щодо розподілів (включаючи дисперсії) X і Y. Є оптимальною моделлю для оцінювання результатів експерименту порівняння методів.

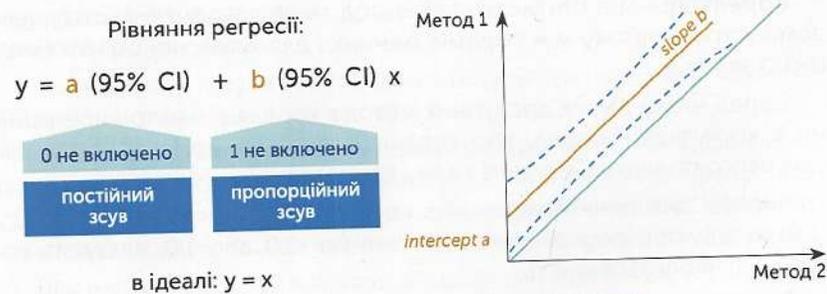


РИС. 5.4. ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ РЕГРЕСІЙНОГО АНАЛІЗУ

## Діаграми різниці (Difference plot)

Оскільки звичайні клінічні лабораторії рідко, якщо взагалі коли-небудь, використовують референтні процедури, зазвичай використовується підхід, який запропоновано Бленд і Альтман.

Діаграми різниці (діаграми Бленд-Альтмана) відображають результати експерименту порівняння методів із концентраціями існуючої процедури вимірювання на осі x і різницею між існуючими та порівняльними процедурами вимірювання на осі y.

Програмне забезпечення або робочі інструменти електронних таблиць обчислюють середню різницю та її стандартне відхилення, а також довірчий інтервал для середньої різниці та межі узгодження (limits of agreement, LOA).

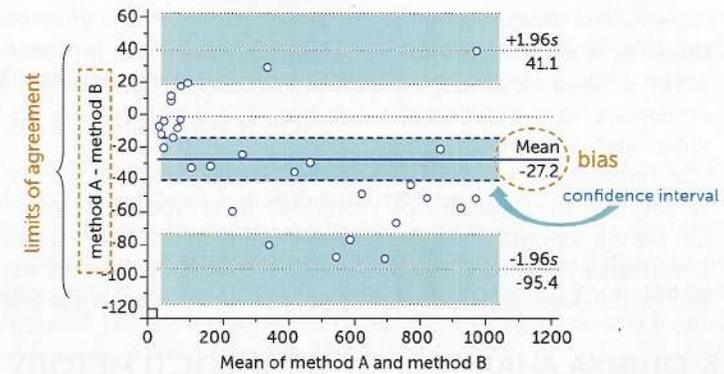


РИС. 5.5. ДІАГРАМА БЛЕНД-АЛЬТМАНА

Якщо довірчий інтервал містить нуль, то спостережуваний зсув має 95 % вірогідність того, що буде спостерігатися виключно випадково, що вказує на відсутність статистично значущого зсуву. Межі узгодження – інтервал, у якому очікується 95 % відмінностей вимірювань. За умови, що відмінності в межах цього інтервалу не є клінічно важливими (тобто меншими за межі прийнятного зсуву), обидва методи вимірювання можна використовувати як взаємозамінні.

## Інтерпретація результатів досліджень порівняння кількісних методів

Якщо проводилися регресійні дослідження, вони використовуються для оцінки наявності систематичних (постійних) або пропорційних (залежних від концентрації) помилок і для отримання враження про розподіл результатів вимірювань.

Ділянки різниці та процедура Бленда-Альтмана використовуються, щоб вирішити, чи відповідає зсув процедури вимірювання-кандидата критеріям прийнятності зсуву для методу порівняння.

Можливі різні результати, які необхідно оцінити самостійно:

- 95 % ДІ середнього зміщення включає нуль, а межі узгодження (LOA) більш жорсткі, ніж прийнятний зсув: систематичної помилки не спостерігалось, і два методи є взаємозамінними, оскільки принаймні 95 % різниць вимірювань потрапляють у заздалегідь визначені межі прийнятного зміщення;
- 95 % ДІ середнього зміщення не включає нуль, але межі узгодження (LOA) більш жорсткі, ніж прийнятний зсув: хоча спостерігалася статистично значуща систематична помилка, обидва методи є взаємозамінними, оскільки щонайменше 95 % відмінностей вимірювань потрапляють у заздалегідь визначені межі прийнятного зміщення. Систематична похибка має бути

оцінена на предмет клінічної значущості (в ідеалі, в точках прийняття медичних рішень). Якщо 95 % довірчий інтервал для зсуву в точці медичного рішення включає нуль, можна з 95 % впевненістю стверджувати, що похибка вимірювання не має клінічного значення;

- iii) LOA перевищує допустимий критерій зсуву: методи не є взаємозамінними. Лабораторії потрібно буде зробити висновок, що метод-кандидат неприйнятний, оцінити клінічну важливість спостережуваного зсуву (ідеально, в точках прийняття медичних рішень) або відкоригувати референтні інтервали для аналізу.

### 5.3. ОЦІНКА АНАЛІТИЧНОЇ ЧУТЛИВОСТІ МЕТОДУ

Термін «аналітична чутливість» (або здатність виявлення) описує групу характеристик продуктивності, які характеризують точність вимірювання при низьких концентраціях аналізу.

Такими атрибутами є «Межа бланку (LoB)», «Межі виявлення (LoD)» і «Межа кількісного визначення (LoQ)», і всі вони описують найнижчу концентрацію вимірюваної величини, яку можна надійно виміряти за допомогою аналітичної процедури.

Документ *CLSI EP17-A2 (Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures)* надає стандартизований підхід до визначення LoB, LoD і LoQ.

#### Межа холостої проби (бланку) LoB

LoB являє собою концентрацію, нижче якої лежать 95 % сигналів вимірювання при багаторазовому вимірюванні холостого зразка. Решта 5 % представляють частоту хибнопозитивних результатів, тобто кількість позитивних вимірювальних сигналів, генерованих пустим зразком. Статистично це відноситься до помилки типу I ( $\alpha$ ). З іншого боку, невелика частка зразків з дуже низькою концентрацією аналізу може давати вимірювальний сигнал, який розташований нижче LoB, таким чином генеруючи хибнонегативні результати. Це називається помилкою типу II ( $\beta$ ). Щоб визначити LoB, контрольні проби та проби низького рівня (зразки з концентрацією аналізу близько очікуваної LoD) вимірюють повторно протягом кількох днів. Приблизна оцінка концентрації для зразка низького рівня становить від одного до п'яти разів більше максимального значення, отриманого в результаті 20 повторних вимірювань одного холостого зразка.

#### Верифікація LoB

Заявлений LoB, зазначений виробником, може бути перевірений шляхом проведення загалом 20 вимірювань на двох холостих зразках протягом 3 днів, а потім розрахунку частки результатів холостого ви-

мірювання, які менші або дорівнюють заявленому LoB. Якщо спостережуваний відсоток становить щонайменше 85 % (17/20), заявлений LoD верифіковано.

#### Межа виявлення (LoD)

LoD визначається як найменша кількість або концентрація компонента, яку можна надійно відрізнити від LoB. Отже, він відображає рівень аналізу, при якому ймовірність того, що зразок низького рівня дасть хибнонегативний результат, дорівнює визначеному ризику помилки типу II. Ризик помилки типу II, як правило, встановлюється на значенні 0,05, що зазвичай те саме, що використовується для ризику помилки типу I (тобто  $\alpha = \beta = 0,05$ ).

#### Верифікація LoD

Для верифікації LoD, заявленого виробником, проводяться 20 вимірювань на двох зразках низького рівня з концентраціями аналізу при заявленому LoD протягом 3 днів і обчисленням частки результатів вимірювань, менших або рівних LoD. Якщо спостережуваний відсоток становить щонайменше 85 % (17/20), заявлений LoB верифіковано.

#### Межа кількісного визначення (LoQ)

LoQ визначається як найнижча концентрація аналізу в зразку, яка може бути визначена з прийнятною точністю і точністю за відповідних умов. Тому вона завжди має повідомлятися разом із відповідними цілями якості.

Якщо зсув неможливо оцінити, внутрішньолабораторний LoQ визначається лише в умовах внутрішньолабораторної (проміжної) прецизійності, що є еквівалентом функціональної чутливості методу.

#### Верифікація LoQ

Для верифікації заявленого LoQ проводиться 20 вимірювань на двох зразках із концентраціями на заявленому LoQ протягом 3 днів і розраховується частка результатів вимірювань, що відповідають допустимій межі похибки. Якщо спостережуваний відсоток становить щонайменше 85 % (17/20), заявлений LoQ верифіковано.

### 5.4. ОЦІНКА ЛІНІЙНОСТІ

Лінійний діапазон методу визначається як діапазон концентрації, в якому існує прямо пропорційний зв'язок між сигналом вимірювання та концентрацією аналізу. Хоча описані численні експериментальні методи для оцінювання лінійного діапазону, найбільш широко прийнятою процедурою є поліноміальний метод, вперше описаний Кроллом та ін. і згодом прийнятий документом CLSI EP06-A.

Процедура складається з двох кроків: оцінювання лінійності та, у разі нелінійності, обчислення ступеня нелінійності та порівняння з допустимими критеріями.

### Оцінка лінійності

П'ять-дев'ять зразків з різними відомими концентраціями або обчисленими згідно з коефіцієнтами розведень аналізуються в двох репліках, а результати наносяться на осі У проти відповідних концентрацій або коефіцієнтів розведень на осі Х. Незважаючи на те, що не вимагається і щоб передбачувані концентрації були отримані шляхом розведення, і концентрації зразків не повинні бути рівновіддаленими, практичний підхід, описаний CLSI, такий:

Отримують два пули зразків, один з високими і один з низькими значеннями концентрації аналіту. Пул з низькими значеннями повинен бути поблизу LoD.

**Три пули проміжної концентрації готуються наступним чином:**

- Три частини низького пулу з низькою концентрацією і одна частина пулу з високою концентрацією;
- Дві частини пулу з низькою концентрацією і дві частини пулу з високою концентрацією;
- Одна частина пулу з низькою концентрацією і три частини пулу з високою концентрацією.

Концентрація кожного пулу ( $C_i$ ) розраховується з об'ємів ( $V_L$ ,  $V_H$ ) і концентрацій ( $C_L$ ,  $C_H$ ) низьких і високих концентрацій в пулах:

$$C_i = \frac{(C_L \times V_L + C_H \times V_H)}{(V_L + V_H)}$$

Якщо валідується лінійний діапазон, слід проаналізувати 9 –11 рівнів з 2-4 реплікатами кожен. При верифікації лінійного діапазону слід виміряти 5 –7 рівнів з 2 реплікатами на кожному рівні.

Поліноміальний регресійний аналіз виконується для поліномів першого, другого та третього порядку за допомогою відповідного програмного засобу:

перший:  $y = a + bx$ ;

другий:  $y = a + bx + cx^2$ ;

третій:  $y = a + bx + cx^2 + dx^3$ .

Вихід програмного забезпечення повинен включати t-тест нелінійних коефіцієнтів ( $c$  і  $d$ ). Якщо t-критерій не є значущим для обох коефіцієнтів, пряма лінія демонструє найкращу відповідність даним. У цьому випадку лінійність підтверджується, і на цьому експеримент можна завершити. Однак, якщо один або обидва коефіцієнти є значу-

щими, це свідчить про нелінійність, клінічну значущість якої необхідно оцінити шляхом розрахунку ступеня нелінійності.

### Розрахунок ступеня нелінійності

Ступінь нелінійності розраховується шляхом першого вибору полінома другого або третього порядку з найменшою стандартною помилкою регресії ( $S_{yx}$ ). Це найкращий поліном і є основою для оцінювання.

Потім розраховується відхилення від лінійності ( $DL_i$ ) для кожної концентрації:

$$DL_i = p(x_i) - (a + bx_i),$$

де:  $p(x_i)$  – оцінка У з найкращим поліномом для кожної концентрації  $X_i$ .

Нарешті обчислений  $DL_i$  перевіряється на кожному рівні та порівнюється з допустимими критеріями зміщення для методу вимірювання.

Якщо прийнятний зсув не перевищено на жодному з рівнів концентрації, то ступінь нелінійності є неприйнятним. Якщо  $DL_i$  вище прийнятного зміщення на одному або кількох рівнях, це вказує на те, що на цих рівнях існує проблема лінійності, яку необхідно дослідити та усунути. Крім використання допустимого критерія зсуву %В для оцінювання лінійності, деякі автори рекомендують використовувати одну чверть загальної похибки для аналізованої речовини ( $TEa/4$ ).

## 5.5. ОЦІНКА ПЕРЕНОСУ

Перенесення (Carryover) визначається як «перенесення компонентів зразка з одного зразка на інший або випадкове перенесення реагенту в зразок». Перенесення зазвичай є проблемою лише для аналітів із широким діапазоном концентрації аналіту, таких як  $\beta$ -hCG, креатинкіназа або певні лікарські засоби, що викликають зловживання.

Перенесення слід оцінювати для всіх аналізаторів із системами автотіпування, які не використовують наконечники для одноразового використання. Якщо виявлено перенесення, необхідно визначити концентрацію аналіту, вище якої можуть постраждати наступні зразки. Згодом результати для кожного аналітичного запуску мають бути переглянуті, і, якщо будуть виявлені результати, що перевищують цю концентрацію, необхідно вжити відповідних заходів (наприклад, повторний аналіз). Зазвичай для оцінювання перенесення використовується одна з наступних двох процедур: стандартний протокол або короткий протокол.

Стандартний протокол включає вимірювання п'яти серій з двома зразками високої концентрації ( $H_1, H_2$ ), а потім трьома зразками низької концентрації ( $L_1, L_2, L_3$ ) у кожному циклі (разом – 25 визначень, 10 з високою концентрацією та 15 з низькою). Висока концентрація повинна бути поблизу (наприклад, 20 %) від верхньої межі кількісного визначення (ULOQ), а низька концентрація повинна бути в ділянці (наприклад, 20 %) нижньої межі кількісного визначення (LLOQ). Перенесення ( $h$ ) потім розраховується наступним чином:

$$h = \bar{L}_1 - \bar{L}_3.$$

Тобто принцип визначення полягає в оцінці різниці показників в зразку з низькою концентрацією, який іде одразу за зразком з високою концентрацією, та останнього в циклі зразка з низькою концентрацією, який іде відповідно за зразком з низькою концентрацією.

Його також можна виразити у вигляді частки або відсотка від  $L_3$ :

$$h = \frac{\bar{L}_1 - \bar{L}_3}{\bar{L}_3}$$

$$h(\%) = \frac{\bar{L}_1 - \bar{L}_3}{\bar{L}_3} \times 100.$$

**Короткий протокол** рекомендовано Європейським агентством з лікарських засобів (EMA). За цією процедурою холосту пробу аналізують після калібрувального стандарту найвищої концентрації, а перенесення визначається як відповідь холостого зразка, отриманого після вимірювання стандарту найвищої концентрації.

### Інтерпретація результатів перенесення

В якості критеріїв значущої різниці використовуються:

- **Нижня межа кількісного визначення (LLOQ)**. Перенесення не є значущим, якщо воно становить 20 % від LLOQ;
- **Стандартне відхилення (SD)**. Перенесення не є значущим, якщо воно становить 2 SD процедури вимірювання, розрахованих із вибірок з низькою концентрацією  $L_3, L_6, L_9, L_{12}, L_{15}$ ;
- **Парний t-тест**. Середнє значення зразків  $L_1, L_4, L_7, L_{10}, L_{13}$  порівнюють із середнім значенням результатів вимірів зразків  $L_3, L_6, L_9, L_{12}, L_{15}$  за допомогою парного t-критерію на рівні значущості  $\alpha=0,01$ . Значення  $P \geq 0,01$  вказує на те, що не було виявлено значного ефекту перенесення.

### 5.6. ВЕРИФІКАЦІЯ КАЛІБРУВАННЯ

Верифікація калібрування визначається як «аналіз матеріалів з відомою концентрацією таким же чином, як і зразків пацієнта, щоб під-

твердити придатність калібрування приладу або тест-системи в межах діапазону вимірювання для результатів тестів пацієнта».

Правила CLIA вимагають, щоб лабораторії перевіряли статус калібрування системи «згідно з графіком та інструкціями виробника, але принаймні раз на 6 місяців», та щоразу, коли трапляється будь-яке з наступного:

- усі реагенти, які використовуються для процедури тестування, змінюються на нові номери партій, якщо лабораторія не може продемонструвати, що зміна номерів партій реагентів не впливає на діапазон, який використовується для звітування про результати тесту пацієнта, а також на контрольні значення не впливають зміни номера партії реагентів;
- проводиться серйозне профілактичне обслуговування або заміна критичних частин, які можуть вплинути на результати тесту. Це стосується випадків, коли лабораторія надсилає обладнання виробнику для ремонту. Лабораторія повинна перевірити калібрування тест-системи на відремонтованому приладі, перш ніж відновити тестування пацієнта та повідомити результати;
- контрольні матеріали відображають незвичну тенденцію або зрушення, або виходять за допустимі для лабораторії межі, а інші засоби оцінювання та виправлення неприйнятних контрольних значень не дозволяють виявити та усунути проблему.
- Для деяких досліджень верифікація калібрування є підтвердженням того, що поточні налаштування калібрування для кожного аналізу залишаються дійсними для методу і дозволяє не виконувати повне калібрування або повторне калібрування випробувальної системи.

#### Винятками є:

- інструменти, що калібруються в умовах виробника;
- для автоматизованих лічильників клітин вимоги до перевірки калібрування вважаються виконаними, якщо лабораторія дотримується інструкцій виробника щодо роботи приладу та випробовує щодня два рівні контрольних матеріалів для тестування за умови, що результати контролю відповідають критеріям лабораторії для прийнятності;
- для автоматизованих біохімічних аналізаторів вимоги до перевірки калібрування вважаються виконаними, якщо лабораторія дотримується інструкцій виробника щодо роботи приладу та регулярно перевіряє три рівні контрольних матеріалів (найнижчий доступний, середній і найвищий рівень) більше ніж один раз на день тестування. Під час тестування результати контрольного матеріалу відповідають критеріям лабораторії щодо прий-

нятності, а контрольні матеріали простежуються до стандартних матеріалів Національного інституту стандартів і технологій (NIST);

- якщо процедура калібрування випробувальної системи включає три або більше рівнів калібрувального матеріалу та включає низьке, середнє та високе значення та виконується принаймні раз на шість місяців, тоді вимога щодо перевірки калібрування також виконується.

### Виконання верифікації калібрування

Статус калібрування системи можна перевірити шляхом тестування принаймні трьох матеріалів з відомою концентрацією (низького, середнього та високого рівнів), які охоплюють діапазон вимірювання аналізу, у такий самий спосіб, як і зразки пацієнтів. Зразки повинні максимально повно охоплювати діапазон вимірювання. Дослідження зразків проводять однократно, а результати порівнюють із призначеними значеннями. В деяких країнах випускаються спеціальні набори зразків для верифікації калібрування.

Верифікація калібрування є успішною, якщо результати вимірювань є прийнятно точними (в межах допустимих меж похибки).

В якості допустимих меж використовуються консенсусні значення, визначені в лабораторії (наприклад, 10-15 %), проте бажано використовувати припустимі специфікації щодо зсуву (В%) або половину значення загальної припустимої помилки (TEa).

## 5.7. ОЦІНКА НЕВИЗНАЧЕНОСТІ

Лабораторна інформація відіграє все більшу роль у забезпеченні діагностики та лікування, але, як і для будь-якої іншої клінічної інформації, всі обмеження в діагностичних тестах впливають на діагностичні та клінічні рішення. Щодо результатів лабораторних досліджень, є можливість оцінити надійність лабораторних даних за допомогою такого показника, як невизначеність вимірювання.

Необхідність проведення оцінювання невизначеності вимірювання підкреслюється і є особливою вимогою ISO 15189 (п.5.5.1.4), отже, для спеціалістів сучасних медичних лабораторій є актуальним знання методів розрахунку з урахуванням особливостей саме медичних лабораторій, встановлення максимально припустимої невизначеності вимірювання та подальшого використання даних щодо невизначеності результатів окремих тестів у різних клінічних ситуаціях.

В новій редакції ISO 15189:2022 є ряд уточнень щодо місця оцінювання невизначеності в сучасній лабораторній медицині:

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<p><b>5.5.1.4 Невизначеність вимірювання значень вимірюваної величини</b></p> <p>Лабораторія повинна встановити невизначеність вимірювань для кожної методики вимірювання на усіх етапах дослідження, яка використовується для кількісного характеризування значень вимірюваних величин у пробах пацієнтів. Лабораторія повинна визначити вимоги до функціональних характеристик стосовно невизначеності вимірювань для кожної методики вимірювань і регулярно аналізувати результати оцінювання невизначеності вимірювань.</p> <p><i>Примітка 1.</i> Відповідними компонентами невизначеності є ті, що пов'язані з фактичним процесом вимірювання, який починається з отримання проби для виконання вимірювання і завершується отриманням значення вимірюваної величини.</p> <p><i>Примітка 2.</i> Невизначеності вимірювань можуть обчислюватися з використанням кількісних значень величин, отриманих під час кількісних досліджень матеріалів, призначених для контролю якості, в умовах проміжної прецизійності, які включають у себе стільки рутинних змін, скільки є обґрунтовано можливим при звичайному виконанні процедури вимірювання, наприклад, заміни партій реагентів та калібраторів, заміни операторів, заплановане технічне обслуговування обладнання.</p> <p><i>Примітка 3.</i> Прикладами практичної корисності оцінювання невизначеності вимірювань може бути підтвердження того, що результати досліджень пацієнта відповідають цілям щодо якості, встановленим лабораторією, а також цілеспрямоване порівняння значення певного показника пацієнта із попередніми значеннями того ж показника або з межею прийняття клінічних рішень.</p>	<p><b>7.3.4 Оцінка невизначеності вимірювання (НВ) (Evaluation of measurement uncertainty (MU))</b></p> <p>а) Невизначеність вимірювання (НВ) значень вимірюваної величини необхідно оцінювати та підтримувати для використання за призначенням, де це доречно. MU необхідно порівняти з функціональними специфікаціями та задокументувати.</p> <p><i>Примітка ISO/TS 20914 надає детальну інформацію про ці дії разом із прикладами.</i></p> <p>б) Оцінки невизначеності вимірювань повинні регулярно переглядатися.</p> <p>с) Для процедур дослідження, де оцінка НВ неможлива або не застосовна, необхідно задокументувати обґрунтування виключення з оцінки НВ.</p> <p>д) Інформація про НВ (MU) повинна бути доступна користувачам лабораторії за запитом.</p> <p>е) Коли користувачі мають запити щодо НВ, відповідь лабораторії повинна брати до уваги інші джерела невизначеності, такі як біологічні варіації, але не обмежуватися ними.</p> <p>ф) Якщо якісний результат обстеження базується на тесті, який дає кількісні вихідні дані та визначається як позитивний або негативний на основі</p>

ДСТУ EN ISO 15189:2015	ISO 15189:2022
<p>Лабораторія повинна враховувати невизначеність вимірювань під час інтерпретації значень вимірюваних показників. На вимогу лабораторія повинна надавати користувачам лабораторних послуг свої оцінки невизначеності вимірювань.</p> <p>Якщо дослідження включають у себе етап вимірювання, але не передбачають повідомлення значення вимірюваного показника, лабораторія повинна обчислити невизначеність етапу вимірювання, якщо ця невизначеність вживатиметься для оцінювання надійності методики дослідження або впливатиме на результати дослідження.</p>	<p>порогового значення, НВ у результаті має бути оцінено за допомогою репрезентативних позитивних і негативних зразків.</p> <p>g) Для аналізів з якісними результатами, НВ на проміжних етапах вимірювання або результати ВЛКЯ(IQC), які дають кількісні дані, також слід розглядати для ключових (з високим ризиком) частин процесу.</p> <p>h) НВ слід брати до уваги під час виконання верифікації або валідації методу, якщо це доречно.</p>

У представленій таблиці додано посилання на стандарт (технічну специфікацію) ISO/TS 20914:2019 Medical laboratories – Practical guidance for the estimation of measurement uncertainty (Медичні лабораторії – Практична настанова щодо оцінювання невизначеності вимірювань).

У Міжнародному словнику з метрології (VIM) невизначеність вимірювання (НВ) визначається як «параметр, що характеризує дисперсію значень величини, які приписують вимірюваній величині на підставі наявної інформації». Він описує інтервал, в межах якого значення вимірюваної величини перебуває із заданою ймовірністю.

Знання НВ та визначення її допустимих меж для клінічного застосування вимірювань є одним із основних факторів, необхідних для отримання стандартизованих лабораторних результатів, придатних для клінічного використання. Оцінка та перевірка НВ в медичних лабораторіях необхідна для розуміння її впливу на результати вимірювань та, зрештою, їх клінічної придатності у визначених клінічних ситуаціях.

### Розрахунок невизначеності вимірювань в медичних лабораторіях

Для оцінювання невизначеності вимірювань запропоновано велику кількість документів та настанов, найбільш відомими серед яких є:

- GUM «Guide to the Expression of Uncertainty in Measurement»,
- NORDTEST TR 537 «Handbook for calculation of measurement uncertainty in environmental laboratories»,

- CLSI EP29A «Expression of Measurement Uncertainty in Laboratory Medicine»,
- ISO 21748:2017 «Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation».

Наприкінці 2019 року був прийнятий стандарт (технічна специфікація) ISO/TS 20914:2019 Medical laboratories – Practical guidance for the estimation of measurement uncertainty (Медичні лабораторії – Практична настанова щодо оцінювання невизначеності вимірювань), в якому надано практичні вказівки щодо оцінювання та вираження невизначеності вимірювань (НВ) кількісних значень вимірюваних величин, що виконуються медичними лабораторіями. Цей документ також застосовується для оцінювання НВ для результатів кількісних вимірювань, отриманих на межі прийняття медичного рішення системами тестування на місці лікування (POCT), а також отриманих якісними (номінальними) методами, які включають етап вимірювання.

Історично було запропоновано два підходи для оцінювання НВ: так званий підхід «знизу вгору» та «зверху вниз». Підхід «знизу вгору» – це модель, спочатку запропонована в «Настанові з оцінювання невизначеності вимірювань» (GUM).

Підхід «зверху вниз» простіший і є хорошою альтернативою попередньому підходу. Він оцінює невизначеність лабораторних результатів за допомогою даних внутрішнього лабораторного контролю якості (ВЛКЯ) для отримання випадкових компонентів невизначеності та інформації комерційного калібратора. Зараз такий підхід офіційно затверджений Технічною специфікацією ISO/TS 20914:2019, яка застосовується в медичних лабораторіях для оцінювання НВ значень, отриманих за допомогою процедур вимірювання, призначених для вимірювання біологічних вимірюваних величин.

Зокрема, невизначеність, пов'язана з неprecizійністю, може бути оцінена за допомогою даних ВЛКЯ, тоді як невизначеність калібратора повинна включати всі невизначеності, введені вибраною ієрархією калібрування для вимірюваної величини, починаючи з найвищого доступного референсного матеріалу або методики до присвоєного значення калібратора для комерційного медичного аналізатора, включаючи невизначеність корекції зсуву, оціненого виробником.

На теперішній час запропоновані об'єктивні критерії для визначення аналітичних специфікацій продуктивності (APS). Ці критерії базуються на трьох моделях: модель 1, заснована на впливі аналітичних показників на клінічні результати; модель 2, заснована на складових біологічних варіацій вимірюваної величини; і модель 3, заснована на сучасному рівні вимірювань (визначається, як найвищий рівень аналітичних показників, технічно досяжний).

Певні моделі краще підходять для одних вимірюваних величин, ніж для інших, і тому увага повинна, в першу чергу, бути спрямована на вимірювану величину та її біологічні та клінічні характеристики.

Класифікація різних рівнів якості (тобто мінімального, бажаного та оптимального) для APS також дуже важлива.

Визначення досліджуваної величини (y)		
Бюджет невизначеності вимірювання (ТВ <sub>u</sub> )	Невизначеність референтних матеріалів/методу	U <sub>ref</sub>
	Невизначеність калібратора	U <sub>cal</sub>
	Непрецизійність системи вимірювання	u(y)
	Індивідуальна продуктивність лабораторії	
Результат тесту зразка пацієнта		

РИС. 5.6. КОМПОНЕНТИ ЗАГАЛЬНОГО БЮДЖЕТА НЕВИЗНАЧЕНОСТІ ЗА Ф. БРАГА ТА ІН. (2020)

Наприклад, якщо вимірювана величина має суворий метаболічний контроль, найбільш відповідною моделлю для отримання APS є та, що базується на її біологічних варіаціях. Використовуючи опубліковані дані про середню внутрішньоіндивідуальну біологічну варіацію CVI та класичну парадигму Фрейзера для отримання APS, специфікації для невизначеності можна отримати як  $\leq 0,75$  CVI, мінімальна якість,  $0,50$  CVI, бажана якість та  $\leq 0,25$  CVI, оптимальна якість, відповідно.

Стандартна невизначеність результату вимірювання  $u(y)$  не має перевищувати отриману аналітичну специфікацію, тобто специфікація покриває загальний бюджет НВ (ТВ<sub>u</sub>), який слід виконувати на рівні результатів пацієнта. Досягнення ТВ<sub>u</sub> залежить від внесків MU кожного етапу метрологічного ланцюга простежуваності. Відповідними компонентами невизначеності є ті, що пов'язані з фактичним процесом вимірювання, який починається з отримання проби для виконання вимірювання і завершується отриманням значення вимірюваної величини.

Невизначеності вимірювань можна обчислювати, використовуючи кількісні значення величин, отримані під час кількісних досліджень матеріалів, призначених для контролювання якості, в умовах проміжної прецизійності, які охоплюють стільки рутинних змін, скільки є обґрунтовано можливим за звичайного виконання процедури вимірювання, наприклад, заміни партій реагентів та калібраторів, заміни операторів, заплановане технічне обслуговування обладнання.

## Оцінка невизначеності результатів вимірювань за допомогою протоколів ISO 20914:2019

Відповідно до рекомендацій ISO / TS 20914, стандартна НВ  $u(y)$  вимірюваної величини Y може бути розрахована на основі результатів вимірювань одного або кількох зразків контролю якості. А саме: невизначеність в умовах проміжної прецизійності, виражена як стандартне відхилення, визначається ретроспективно на основі даних, отриманих з лабораторної інформаційної системи. Ця стандартна невизначеність  $u_{RW}$  поєднується з невизначеністю цільових значень калібратора  $u_{cal}$ , як показано у рівнянні:

$$u(y) = \sqrt{u_{RW}^2 + u_{cal}^2 + u_{bias}^2}$$

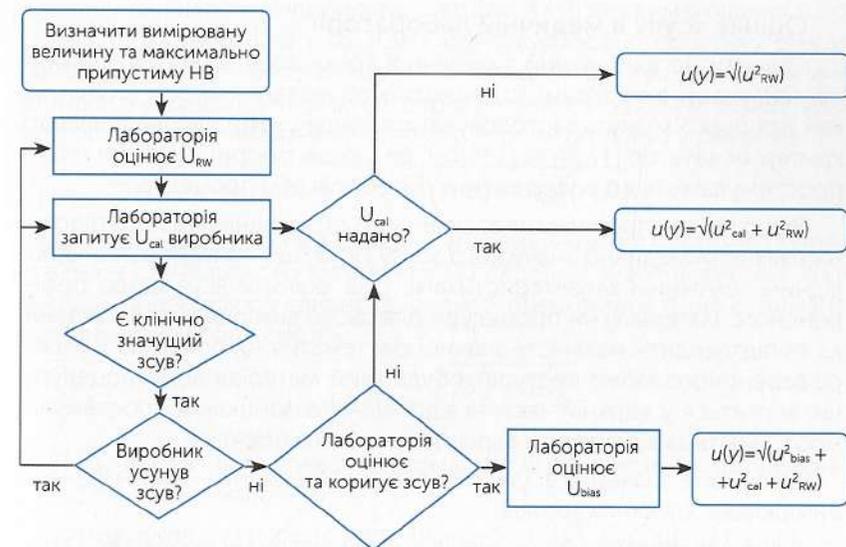


РИС. 5.7. АЛГОРИТМ ОЦІНЮВАННЯ НЕВИЗНАЧЕНОСТІ ВИМІРЮВАНЬ ЗА ISO / TS 20914

Виробник калібратора несе відповідальність за те, щоб надати користувачу значення  $u_{cal}$ . Якщо була проведена корекція зсуву, слід також додати її невизначеність  $u_{bias}$ .

Умови для отримання  $u_{RW}$  повинні відповідати внутрішньолабораторній прецизійності протягом визначеного періоду, достатнього для охоплення більшості змін умов вимірювання та систематичних джерел невизначеності (наприклад, за 6 місяців), таких як спричинені різними лотами реагентів, різними калібруваннями чи різними умовами навколишнього середовища. Характеристики контрольного ма-

теріалу для оцінювання  $u_{RW}$  повинні бути ретельно продумані, тобто матеріал має бути комутабельним і з концентрацією (концентраціями), що відповідає точкам прийняття клінічних рішень.

Якщо використовується більше одного калібратора (наприклад при імуноаналізі), згідно з ISO / TS 20914 для розрахунку використовується  $u_{cal}$ , найбільш відповідний до кожного рівня концентрації ВЛКЯ. Якщо в даному проміжку часу змінюється лот матеріалу контролю якості (КМ), показник  $u_{RW}$  розраховується як зведене (pooled) середнє значення для різних лотів, як описано у рівнянні, де  $u_n$  являє собою стандартне відхилення відповідного лоту КМ:

$$u_{RW}(lot1, lot2, \dots, lotn) = \sqrt{\frac{u_1^2 + u_2^2 + \dots + u_n^2}{n}}$$

### Оцінка зсуву в медичній лабораторії

Оцінити неприйнятний з медичної точки зору зсув є можливість під час участі в програмі зовнішнього оцінювання якості (EQA), однак для цього можуть застосовуватися лише схеми, що відповідають критеріям категорії I / IIA за G.Milller, де цільові опорні значення мають простежуваність до референтних матеріалів або процедур.

Якщо під час проведення зовнішнього оцінювання якості є підозра на наявність медично значущого зсуву (зсув, що не відповідає відповідним технічним характеристикам), слід оцінити зсув щодо референтного матеріалу чи процедури для цього вимірюваного показника та підтвердити наявність значної систематичної помилки. В якості референтного може виступати будь-який матеріал або процедура, що міститься у верхній частині відповідного ланцюжка простежуваності, навіть за відсутності варіантів високого порядку.

Отримане значення зсуву слід включити в оцінку невизначеності вимірювань клінічних зразків.

Якщо вирахована невизначеність вимірювань не відповідає заздалегідь визначеним технічним характеристикам, виробник несе відповідальність за негайне розслідування та, врешті-решт, вирішення проблеми коригуючими діями.

Якщо виробник не вирішить проблему, це може зробити лабораторія та ввести поправочний коефіцієнт для виявленого зсуву. Якщо так, то невизначеність поправочного коефіцієнта потрібно оцінити та включити до розрахунку результату.

В стандарті ISO 21748:2010 «Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation» наведено приклади різних варіантів розрахунку невизначеності вимірювань з урахуванням, в тому числі, значущого зсуву.

Для кожної аналітичної системи, аналіту та рівню контролю якості повинні бути визначені відповідні статистичні значення та визначена відносна розширена комбінована невизначеність  $U$ . Вона розраховується з коефіцієнтом охоплення  $k = 2$  (рівень довіри 95%):

$$U(y) = 2 \cdot u(y)$$

$$U(y)_{rel} = \frac{U(y)}{mean} \cdot 100,$$

де  $U(y)_{rel}$  – відносна розширена невизначеність у %.

Невизначеність вимірювань результату пацієнта, що повідомляється, повинна охоплювати наявні розбіжності вимірювальних систем, якщо в лабораторії для досліджень застосовується більше однієї аналітичної системи (наприклад, системи A і B) для вимірювання того самого показника. ISO / TS 20914 описує спосіб розрахунку об'єднаної середньої невизначеності  $u$  (pooled) шляхом комбінування стандартного відхилення  $u(A, B)$  середніх значень A і B з їх середньою довгостроковою невизначеністю  $u_{RW}(A, B)$ , як описано у рівнянні:

$$u(pooled) = \sqrt{u^2(A, B) + u_{RW}^2(A, B)},$$

де  $u^2(A, B)$  – дисперсія від середніх значень  $x$  кожної системи, розрахована із стандартного відхилення  $SD$  середніх значень систем, тоді як  $u_{RW}^2$  – дисперсія об'єднаної середньої прецизійності, що отримана зі зважених  $SD$  обох систем:

$$u_{RW}(A, B) = \sqrt{\frac{SD_A^2 + SD_B^2}{2}}.$$

Як альтернативу, можна повідомляти НВ тієї системи, що демонструє більшу невизначеність.

Визначення розширеної невизначеності  $U(A, B)$  досягається поєднанням з  $u_{cal}$  та коефіцієнтом покриття  $k = 2$  (95 % ймовірності покриття):

$$U(A, B) = k \cdot \sqrt{u(pooled)^2 + u_{cal}^2}.$$

Відносна розширена невизначеність  $U_{rel}(A, B)$  визначається за формулою:

$$U_{rel}(A, B) = \frac{U(A, B)}{x(A, B)}.$$

Невизначеність вимірювань у зразках пацієнтів може відрізнитися від визначеної при використанні комерційних матеріалів. Деякі компоненти невизначеності, які виникають при відновленні ліофілізованих контрольних матеріалів тощо, не існують для зразків пацієнтів або

їх пулів і, отже, можуть бути однією з причин різниці. Крім того, треба врахувати, що комерційний контроль не обов'язково є повною мірою репрезентативним для зразків пацієнтів. Зразок оцінювання невизначеності результатів дослідження (ISO 20914):

Компонент (аналіт)	Натрій		
Вимірювана величина	Концентрація натрію в сироватці, плазмі, сечі		
Одиниці вимірювання	ммоль/л		
Метод вимірювання	Непрямої, іонселективний електрод		
Процедура вимірювання	Виробник А		
Простежуваність калібратора	Простежуваний до SI SRM919b		
Невизначеність калібратора $u_{cal}$	0,71 ммоль/л		
Зсув	За результатами ЗОЯ несуттєвий		
Проміжна прецизійність	Plasma IQC Level 1	Plasma IQC Level 2	Urine IQC Level 1
Період	09/02/13 – 13/05/13	09/02/13 – 13/05/13	09/02/13 – 13/05/13
Кількість вимірювань КМ	343	338	122
Середнє, (ммоль/л)	134,8	149,8	86,4
$u_{RW}$ (Na, ммоль/л)	0,85	0,87	0,99
$u_c(\text{Na}) = \sqrt{(u_{RW}^2 + u_{cal}^2)}$	1,11	1,12	1,22
$U(\text{Na}); k=2$	2,22	2,24	2,44
$\%U_{rel}(\text{Na})$	1,6%	1,5%	2,6%
При застосуванні до результатів пацієнта	$\pm 1,6\%$	$\pm 1,5\%$	$\pm 2,6\%$

Інші методи та протоколи проведення оцінювання невизначеності, форми сповіщення замовників щодо невизначеності результату вимірювання та методи використання в клінічній практиці для діагностики або моніторингу.

Підхід «знизу вгору» – це модель, спочатку запропонована в «Настанові з оцінювання невизначеності вимірювань» (GUM). Ця мо-

дель зазвичай використовується референтними лабораторіями для отримання акредитації відповідно до стандартів ISO 17025 та 15195 та базується на побудові моделі процесу вимірювання, в якій кожне потенційне джерело невизначеності ідентифікується, визначається кількісно, а потім комбінується для формування результату невизначеності результату за допомогою статистичних правил.

Однак застосування цього підходу в медичних лабораторіях є надто складним і стикається з багатьма практичними проблемами та запереченнями, і на цей час рекомендується для використання розробниками та виробниками діагностичних систем або, за необхідності, для оцінювання внеску у невизначеність вимірювання окремих етапів методики для її оптимізації.

Цей та інші методи оцінювання невизначеності в медичних лабораторіях викладені в керівництві CLSI EP29 Expression of Measurement Uncertainty in Laboratory Medicine, 1st Edition.

### Сповіщення результатів замовнику

Якщо запит клініциста вимагає повідомлення про невизначеність, результат повідомляється у форматі [результат вимірювання]  $\pm U$  ( $k = 2,95\%$  довірчого інтервалу).

За необхідності, при використанні декількох однакових вимірювальних систем для отримання результатів вимірювання одного аналіту, кожна з яких має свою невизначеність вимірювання, можна розрахувати об'єднану НВ для всіх систем згідно з вищевикладеними принципами, щоб повідомити про одну загальну НВ на запит клініциста.

Оцінювання невизначеності вимірювань може бути підтвердженням того, що результати досліджень біологічних зразків пацієнта відповідають цілям щодо якості, встановленим лабораторією, а також для цілеспрямованого порівняння значення певного показника із попередніми значеннями того самого показника або з межею прийняття клінічних рішень.

Лабораторія повинна врахувати невизначеність вимірювань під час інтерпретації значень вимірюваних показників. На вимогу лабораторія повинна надавати користувачам лабораторних послуг інформацію щодо свого оцінювання невизначеності вимірювань.

Якщо дослідження охоплюють етап вимірювання, але не передбачають повідомлення про значення вимірюваного показника, лабораторія повинна обчислити невизначеність етапу вимірювання, якщо ця невизначеність використовується для оцінювання надійності методики дослідження або впливатиме на результати дослідження.

В залежності від клінічних завдань лабораторного дослідження, до уваги беруться відповідні компоненти невизначеності.

Призначення тесту	Приклад	Компоненти та спосіб оцінювання невизначеності вимірювання
Результати тесту потрібні, головним чином, для моніторингу стану пацієнта	Онкомаркери, імуносупресивні ліки	Тільки неprecизійність – дані ВЛКЯ в умовах проміжної прецизійності
Результат тесту використовується шляхом порівняння з референсним інтервалом, встановленим або верифікованим в лабораторії	Гормони	Тільки неprecизійність – дані ВЛКЯ в умовах проміжної прецизійності
Результат тесту звичайно порівнюється з точкою прийняття клінічного рішення	Глюкоза, холестерин	Неprecизійність – дані ВЛКЯ в умовах проміжної прецизійності, зсув – шляхом оцінювання невизначеності калібратора (дані виробника) або оцінка в лабораторії

Отже, основними завданнями оцінювання невизначеності вимірювань у медичних лабораторіях є:

- надання об'єктивної інформації про якість індивідуальних лабораторних результатів дослідження, підтримування, в свою чергу, належного процесу прийняття клінічних рішень;
- слугування інструментом управління для виробників тестів для клінічної лабораторії з примушенням їх дослідити та, врешті-решт, вирішити виявлену проблему;
- допомога тим виробникам, які випускають кращу продукцію та вимірювальні системи для демонстрації переваги цих продуктів;
- визначення вимірюваних величин, які потребують аналітичного вдосконалення для їх клінічного використання, та зацікавленості виробників працювати над підвищенням якості продукції;
- слугування підставою для відмови від аналізів з продемонстрованою їх недостатньою якістю.

### Звіт про оцінку невизначеності вимірювань значень вимірюваної величини

1. Методика досліджень.
2. Одиниці вимірювання.
3. Вимоги до аналітичних характеристик.

Концентрація показника	Неprecизійність CV% або SD, ( $u_c$ припуст)	Зсув В%	Загальна помилка (TEa)
Протягом усього діапазону вимірювання			

4. Оцінка невизначеності за даними проміжної прецизійності.

Дані ВЛКЯ для оцінювання проміжної прецизійності:

Контрольний матеріал	Період	Приписана концентрація TV	Середнє за період $\bar{x}_{сер}$	СКВ SD	КВ CV	Вимірювань КМ, N

Компоненти загальної невизначеності	Рівень 1	Рівень 2
Концентрація вимірюваного показника (log копій/мл)		
<b>Оцінка невизначеності, пов'язаної з неprecизійністю</b>	$u(R_w)$	
за даними проміжної прецизійності (log копій/мл)		
<b>Оцінка невизначеності, пов'язаної з корекцією зсуву</b>	$u(bias)$	
▶ за даними невизначеності калібратора		
▶ за даними ЗОЯ / міжлаб. порівнянь		
▶ за даними повторних вимірювань РМ		
Загальна комбінована невизначеність $u_c$		

Компоненти загальної невизначеності	Рівень 1	Рівень 2
$U_c < U_c$ припуст		
Розширена невизначеність	$U = 2 \times \sqrt{u(R_w)^2 + u(bias)^2}$	
Коефіцієнт покриття $k$		
Розширена невизначеність (U)		
Відносна розширена невизначеність ( $U_{rel}$ , %)		

Звіт складено «\_\_\_» \_\_\_\_\_ 202\_р. \_\_\_\_\_

Звіт затверджено «\_\_\_» \_\_\_\_\_ 202\_р. \_\_\_\_\_

## 5.8. ВЕРИФІКАЦІЯ НОВОГО ЛОТУ РЕАГЕНТІВ

Частою проблемою, яка обмежує здатність лабораторії отримувати послідовні результати з часом, є різниця між партіями (лотами), що впливає на калібратори та реагенти. Ця зміна може бути не позбавлена клінічних наслідків, і є кілька добре задокументованих прикладів несприятливих клінічних результатів.

Важливо, щоб у лабораторіях були встановлені процедури кількісного оцінювання цієї неточності та визначення того, чи є величина варіації прийнятною для оприлюднення результатів пацієнта.

До оцінювання нових партій використовуються різні підходи, включаючи протокол оцінювання EP26, опублікований Інститутом клінічних та лабораторних стандартів (CLSI). Матеріали внутрішнього контролю якості та зовнішнього оцінювання якості часто не є комутабельними, тому перевага віддається нативним зразкам пацієнтів.

Опубліковані протоколи оцінювання суттєво відрізняються простотою використання та статистичною точністю, і деякі можуть бути недостатніми для виявлення клінічно значущих змін між партіями. Крім того, поточні протоколи (включаючи протокол CLSI) не виявлятимуть кумулятивних зсувів між партіями реагентів.

Цей недолік можна принаймні частково усунути за допомогою процедур, які застосовують ковзні середні показники результатів досліджень пацієнтів або подібні процедури якості. Співпраця та обмін даними між лабораторіями та виробниками також відіграють важливу роль у виявленні відмінностей від партії до партії. Хоча лабораторія може вжити заходів для оцінювання та виявлення варіацій, ідеалом є зменшення розбіжностей між партіями на місці виробництва.

Використання відповідних критеріїв прийнятності, заснованих на медичних потребах або вимогах до біологічної варіації замість деякого довільного відсотка, є найбільш прийнятним.

Кроки процесу оцінювання варіації між лотами реагентів:

1. Критерії прийому визначаються заздалегідь;
2. Для оцінювання відбирається відповідна кількість комутабельних зразків;
3. Зразки перевіряються на старій та новій партії;
4. Результати між лотами порівнюються;
5. Якщо результати задовольняють критерії прийнятності, нова партія приймається.

Якщо результати не задовольняють критерії прийнятності, нова партія відхиляється і не повинна використовуватися для вимірювання показників у зразках пацієнтів.

Дії, які необхідно вжити, якщо альтернативні партії недоступні, включають надання клініцистам коментарів зі звітами про результати, зміну референтних інтервалів і впровадження факторів перерахунку, щоб нові результати були порівнянними з попередніми результатами.

### Протокол CLSI EP-26 для оцінювання нових партій реагентів:

1. Дані щодо прецизійності методу використовуються для визначення повторюваності ( $S_r$ ) і стандартного відхилення в межах партії реагентів ( $S_{wrl}$ );
2. Визначається бажана статистична потужність;
3. Відповідні таблиці протоколу використовуються для визначення:
  - а) кількості зразків, необхідних для кожної цільової концентрації;
  - б) межі відхилення (яка буде у відсотках від критичної різниці);
  - в) фактичної статистичної потужності, отриманої в результаті оцінювання;
4. Зразки відбираються та вимірюються на кожній партії реагентів;
5. Середня різниця між результатами для поточної та нової партії для кожної цільової концентрації порівнюється з межею відхилення:
  - а) якщо всі відмінності менші за межу відхилення, то нова партія приймається до використання;
  - б) якщо будь-яка з відмінностей перевищує ліміт відхилення, то нова партія не пройшла оцінку і не повинна використовуватися, поки не буде проведено подальше усунення невідповідності.

Протокол CLSI EP-26 є надійно теоретично обґрунтованим, але занадто складним для виконання в клінічних лабораторіях, тому дістав обмежене поширення в світі.

Запропоновані більш практичні протоколи, які на даний час можуть бути використані для оцінювання варіації між лотами реагентів:

### Протокол університету МакМастера, 2016:

1. Три зразки пацієнтів з низьким рівнем концентрації аналіту;
2. Три зразки з нормальним рівнем концентрації;
3. Три зразки з високим рівнем концентрації;
4. Додатковий зразок з екстремальним значенням концентрації;
5. Зразки ВЛКЯ всіх рівнів
  - в якості критеріїв прийнятності використовуються дані щодо цілей якості – максимально припустима аналітична помилка TEa (оптимально, якщо різниця між лотами не перевищує третини TEa, мінімально – не перевищує TEa)

Приклад робочого листа, що полегшує оцінку результатів, наведений в додаткових матеріалах до Керівництва.

### Протокол клініки Майо

Являє собою пряме порівняння між партіями з використанням 20 зразків пацієнтів у всьому діапазоні концентрації аналізу. Статистичний аналіз здійснюється із застосуванням регресії Пассінга-Баблока, критеріями прийнятності якого є:

- градієнт лінії регресії від 0,9 до 1,1;
- Y-перетин лінії регресії <50 % від найнижчої концентрації, про яку повідомляється;
- R<sup>2</sup> коефіцієнт детермінації >0,95;
- середня різниця між партіями реагентів <10 %.

В якості критеріїв прийнятності EFLM (Lantman et al., 2022) запропонували використовувати дані щодо припустимої стандартної невизначеності  $u_c$  за формулою

$$u_{\text{brlot, a}} \leq \frac{\sqrt{APC^2 - u_{\text{cal}}^2}}{\sqrt{n+1}},$$

де  $u_{\text{brlot, a}}$  – припустима різниця між лотами, APS – припустимі аналітичні специфікації для прецизійності,  $u_{\text{cal}}$  – невизначеність калібратора,  $n$  – середня кількість обстежень пацієнта в цій лабораторії із застосуванням одного лоту реагентів (від 1 до 6).

Отже, при використанні моделі біологічної варіації для встановлення припустимих аналітичних специфікацій та бажаного (desirable) рівня вимог формула матиме наступний вигляд:

$$u_{\text{brlot, a}} \leq \frac{\sqrt{(0.5 \times CV_1)^2 - u_{\text{cal}}^2}}{\sqrt{n+1}}.$$

Якщо невизначеність калібратора невідома, припускається прибрати її з розрахунку.

Варіація між лотами є потенційним джерелом неточності, і це питання найкраще вирішити шляхом розширення співпраці між лабораторіями та виробниками тест-систем. Однак, поки це не станеться, на лабораторії залишається поточна робота щодо оцінювання всіх нових партій реагентів і калібратора, перш ніж використовувати їх. Кожна лабораторія стикається зі своїм набором проблем з точки зору обслуговуваного населення, робочого навантаження та ресурсів. Це означає, що не може бути єдиного підходу до оцінювання варіабельності між лотами. Швидше за все, лабораторія повинна розробити підхід, прийнятний для її власних конкретних потреб і який забезпечує прийнятний рівень статистичної надійності.

## 5.9. ВЕРИФІКАЦІЯ РЕФЕРЕНТНИХ ІНТЕРВАЛІВ

Референтні інтервали є важливим інструментом підтримки прийняття рішень щодо інтерпретації результатів, оскільки лабораторні результати можуть бути інтерпретовані лише після порівняння зі значеннями, знайденими у здорових осіб.

Референтні інтервали можуть бути верифіковані за протоколами EFLM шляхом дослідження невеликої кількості референтних осіб ( $n=20$ ). Результати обстежень таких осіб не повинні містити патологічних відхилень. Референтні особи відбираються з референтної сукупності за специфічними, чітко визначеними критеріями. Коли ці критерії застосовуються до того, як зразки будуть зібрані та проаналізовані, вони називаються апіорними. Якщо ці самі критерії застосовуються після відбору проби, вони називаються апостеріорними.

Для перевірки можливих викидів у таких малих групах рекомендується тест Діксона.

Референтні особи для визначення референтного інтервалу, пов'язаного зі здоров'ям, не обов'язково повинні бути молодими дорослими; вони можуть більше нагадувати групу пацієнтів, які проходять медичне обстеження. Припускається, що вікові референтні інтервали в багатьох випадках можуть бути більш клінічно прийнятними. Однак деякі вікові зміни лабораторних показників можуть не свідчити про хороше здоров'я (наприклад, підвищення рівня лужної фосфатази у пацієнтів літнього віку).

Після випробування 20 зразків результати порівнюють із зазначеним референтним інтервалом.

Інтервал дійсний, якщо не більше двох результатів виходять за межі. Якщо більше чотирьох результатів виходять за межі, інтервал не можна використовувати для популяції лабораторії.

Якщо три або чотири результати лежать поза межами, необхідно отримати ще 20 зразків, знову ж таки без викидів. Якщо три або більше результатів знову виходять за контрольні межі, слід встановити нові контрольні інтервали. В іншому випадку референтний інтервал може бути прийнятий лабораторією. Використовуючи цей підхід, ймовірність помилкового відхилення (ймовірність того, що більше двох результатів виходять за межі, коли насправді 95 % населення перебуває в межах) становить 7,5 %. Після оцінювання іншої вибірки з 20 спостереженнями, коли три-чотири спостереження виходять за межі, рівень помилкового відхилення стає менше 1 %

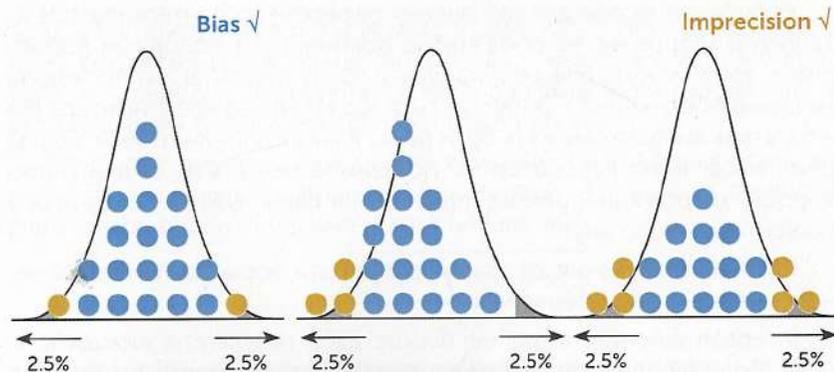
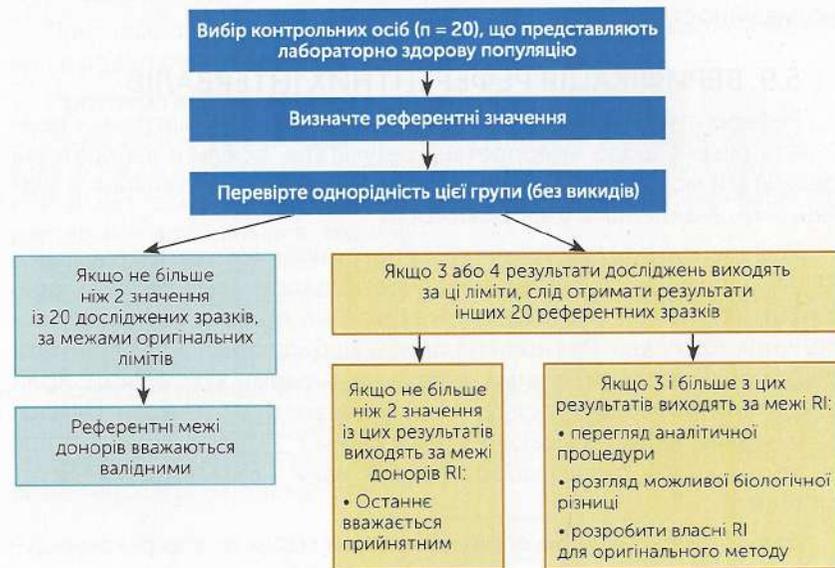


РИС 5.8. ПРОТОКОЛ EFLM ВЕРИФІКАЦІЇ РЕФЕРЕНТНИХ ІНТЕРВАЛІВ

Альтернативний підхід для перевірки запропонованих референтних діапазонів полягає в отриманні наявних лабораторних даних. Цей метод дійсний, лише якщо більшість непатологічних значень можна чітко ідентифікувати (наприклад, за допомогою фільтрації). Використовуючи ці дані, середню точку можна порівняти з відповідною середньою точкою даних, що використовуються для встановлення референтного інтервалу і, отже, оцінювання аналітичної або популяційної упередженості, що може призвести до неправильної класифікації пацієнтів. За деяких обставин (наприклад, у педіатрії) може бути особливо важко використовувати методи прямого забору проб. У цих випадках деякі дослідники виступають за використання непрямих методів, при яких особи не враховуються, але застосовуються певні статистичні методи для обробки бази даних результатів.

## 6. ВЕРИФІКАЦІЯ НАПІВКІЛЬКІСНИХ ТА ЯКІСНИХ МЕТОДІВ

Якісні діагностичні тести використовувалися з перших днів лабораторної медицини для скринінгу, діагностики та лікування різноманітних захворювань. Ці тести зустрічаються в багатьох спеціальностях клінічної лабораторії. До них належать:

- швидкі тести (імунохроматографія, суха хімія);
- морфологічні дослідження (цитологія, гістологія, цито-, гістохімія, цитогенетика тощо);
- класична бактеріологія та вірусологія;
- сучасна бактеріологія;
- серологічні тести (класичні, ІФА, ІХЛА);
- молекулярно-генетичні дослідження (ПЛР та інші молекулярні методи).

Іноді такі види тестів називають напівкількісними, а іноді якісними. Обидва терміни неоднозначні. Інститут клінічних та лабораторних стандартів (CLSI) використовує термін «якісні тести» для «... методів, які дають лише категоричні відповіді (тобто позитивні/ негативні або так/ні)». Інший спосіб класифікувати ці види досліджень – за результатами, які можуть бути номінальними (результат класифікації) або порядковими (оцінювання). Номінальні та порядкові значення можуть бути точно визначені і, отже, за визначенням не є напівкількісними в тому сенсі, що результати мають велику невизначеність. Тому використовувати термін «напівкількісний» треба, добре розуміючи природу отриманих результатів.

Типи якісних методів:

- якісні номінальні (групи крові та ін.);
- якісні бінарні (так/ні);
- кількісні з визначенням порогового значення та видачею бінарних відповідей (COI);
- напівкількісні (+, ++, +++, титри та ін.).

Номінальні та порядкові значення не обмежуються бінарною шкалою. Тест-смужка сечі на визначення альбуміну є прикладом «мультинарної» процедури дослідження з можливим результатом більш ніж двох альтернативних значень. Результати для порядкової величини є результатами вимірювання, хоча результати не мають одиниць вимірювання. Порядкові значення певним чином впорядковані за величиною. Це виключає, наприклад, алфавітне впорядкування значень.

Значення «так» і «ні», включені у CLSI, що використовується вище, не можуть бути впорядковані за величиною і тому не є порядковими значеннями. Результати «позитивні» та «негативні» зазвичай розгляда-

ються як упорядковані за величиною (або 0/1), у тому сенсі, що позитивні мають на увазі більше чогось, наприклад, концентрацію вище певної граничної межі.

Результат тесту на вагітність може бути негативним або позитивним, залежно від концентрації ХГЧ у зразку сечі. Результати «негативні» та «позитивні» можна інтерпретувати як порядкові значення, наприклад, <5 мМО/л і >5 мМО/л відповідно. Результати з бінарним результатом так/ні часто є номінальними. Однією з характеристик цих типів іменних результатів експертизи, наприклад, щодо появи гена або класифікації захворювання, є те, що значення не мають індивідуальних біологічних варіацій. Результати є дискретними, і значення між «так» і «ні» неможливі. Тест-смужки для сечі є класичним прикладом з процедури дослідження з можливим результатом більш ніж двох альтернативних значень, хоча назва кроків шкали (+ або 15 мг/л, наприклад) може імітувати одиниці вимірювання.

Поняття прецизійності, правильності, невизначеності та простежуваності для порядкових та номінальних результатів чітко не визначені. Відповідні визначення до метрологічної термінології та опису простежуваності вимірювань і невизначеності результатів для результатів шкали відношень, якими є результати кількісних досліджень, ще потрібно розробити для номінальних та порядкових результатів.

Невизначеність для порядкових і номінальних результатів може бути виражена як ймовірність правильних або хибних результатів, які є поняттями, які можна зрозуміти для набору результатів, наприклад, 1000 зразків, але їх важче зрозуміти щодо одного позитивного або негативного результату.

Також бракує чіткого визначення межі виявлення, за якої більшість результатів очікується, будуть негативними, і верхньої межі концентрації, вище якої очікується більшість результатів тесту позитивними. У проміжку між цими двома межами випадково можна очікувати негативні або позитивні результати. Цей інтервал можна назвати сірою зоною, ділянкою ненадійності або «інтервалом C5–C95», як запропоновано CLSI.

C5 описує концентрацію, нижче якої <5 % результатів є позитивними, C95 – це концентрація, вище якої >95 % результати є позитивними, коли вимірюється багато повторів. Значення C50, точка еквівалентності або граничний рівень, — це концентрація, при якій половина результатів тесту стають позитивними, а половина — негативними.

Значення C50 має перевагу порівняно зі значеннями C5 і C95, оскільки його можна оцінити з більшою точністю. Зі статистичних міркувань потрібна набагато більша кількість вимірювань, щоб точно оцінити C5 і C95. Коли інтервал C5–C95 вузький, порядковий тест

іноді називають більш точним. Для тестів із кількома кроками шкали інтервал C5–C95 можна описати для кожного з кроків.

Як зазначалося вище, простежуваність та інші метрологічні терміни для номінальних та порядкових результатів ще не були чітко визначені. Істинність порядкового вимірювання можна описати як близькість узгодження між точкою еквівалентності для бінарного порядкового тесту та значенням контрольної величини.

### Цілі аналітичної ефективності для якісних тестів

Цілі аналітичної ефективності для якісних процедур тестування повинні бути встановлені відповідно до передбачуваного використання процедури.

Коли тест використовується з метою скринінгу і позитивний тест буде підтверджено більш специфічним вторинним тестом, тоді цільовою ефективністю процедури скринінгу може бути дуже висока чутливість.

Якісні тести можуть використовуватися для широкого кола клінічних цілей. Їх використання можна описати як скринінг, діагностику, підтвердження або моніторинг.

Чутливість і специфічність тесту, прогностичні значення та поширеність захворювання або стану в тестованій популяції визначають клінічну корисність якісного тесту.

### Скринінгові тести

Клінічно скринінгові методи використовуються для перевірки підгруп населення на наявність аналіту або агента. Як правило, якісні тести, що використовуються для цілей скринінгу, повинні мати високу чутливість, щоб гарантувати виявлення істинно позитивних результатів. Як правило, скринінгові тести дають більше хибнопозитивних результатів, ніж діагностичні або підтверджувальні тести. Таку низьку специфічність можна допускати, якщо існує хороший підтверджувальний тест і якщо соціальні/економічні наслідки хибнопозитивних результатів не є надто серйозними. Необхідність супроводжувати позитивні результати скринінгу підтверджувальним тестом для виявлення хибнопозитивних результатів може бути кращою практикою, ніж отримання хибнонегативних результатів тесту, оскільки помилкові негативні результати можуть мати більш серйозні наслідки, такі як переливання інфікованої крові або невиконання лікування серйозного стану, що піддається лікуванню.

### Діагностичні тести

Якісні тести часто використовуються для діагностики певного захворювання або стану на підставі клінічної підозри на його наяв-

ність. Одним із прикладів діагностичного тесту є використання різних мікробіологічних культуральних методів для виявлення бактеріальної інфекції. Клінічно вимога до належного лікування вимагає, щоб діагностичні тести мали чудову чутливість і специфічність. Якщо підтверджувальний тест слідує за діагностичним тестом, вимоги до його специфічності можуть бути дещо нижчими.

### Підтверджувальні тести

Підтверджувальні тести використовуються для контролю результатів скринінгових або діагностичних тестів. Перевірка або підтвердження результату тесту дозволяє клініцисту встановити діагноз. Підтверджувальні тести розроблені так, щоб бути специфічними (за рахунок зниження чутливості, якщо необхідно) і мати високу позитивну прогностичну цінність (PPV).

Зазвичай неможливо заснувати ціль продуктивності для порядкового тесту на біологічній варіації. Сіра зона, або інтервал C 5–C95, ймовірно, більша за біологічну варіацію для багатьох вимірюваних величин. Цілі ефективності поки що, ймовірно, базувалися, в основному, на показниках state-of-art – добре досліджених та визнаних тестах деяких постачальників. З майбутнім технологічним розвитком можна очікувати покращення ефективності багатьох якісних тестів.

Коли дослідження виконується лише шляхом спостереження оператора, без об'єктивної реакції, яка зазвичай забезпечується приладами, мінливість, викликана інтерпретацією оператора, може бути великою щодо однорідності пристрою всередині та між виробничими партіями. У таких випадках варіація результатів великою мірою залежить від здатності операторів виявляти наявність кольору, на який впливають умови освітлення, колір фону (для контрасту), індивідуальне судження та інші параметри.

### Верифікація якісних методів

Процедури оцінювання методів для таких тестів різноманітні, причому, в різних галузях лабораторної медицини часто акцентують увагу на різних питаннях як при плануванні експерименту, так і в аналізі даних та інтерпретації таких досліджень. Було декілька спроб стандартизувати як експериментальні деталі, так і аналіз даних якісної інформації.

Верифікація методики дослідження здійснюється для визначення функціональних характеристик методики, що встановлені виробником, які підтверджуються в умовах конкретної лабораторії. Для цього лабораторія має отримати від виробника/розробника методу відповідну технічну документацію.

Перелік параметрів (діагностичних характеристик) залежить від типу тесту:

- прецизійність;
- точність;
- порівняння методів;
- верифікація Cut-off;
- верифікація ліміту детекції (LoD).

Оцінка правильності (точності) для якісних та напівкількісних методів може здійснюватися за результатами сигналу вимірювання. В цьому випадку розрахунки та критерії прийнятності ідентичні тим, що використовуються для верифікації правильності для кількісних методів.

Для результатів, представлених в біноміальній або порядковій шкалі, більш значну роль відіграють експеримент порівняння методів і перевірка чутливості та специфічності, як описано нижче. Напівкількісні аналізи використовуються для різноманітних застосувань у клінічній практиці лабораторії, починаючи від скринінгу і закінчуючи діагностикою.

Типовим прикладом таких процедур тестування є тест-смужки для сечі. Результати вимірювань класифікуються на декілька довільних кроків з категоріями результатів без одиниць вимірювання. Як і для кількісних аналізів, напівкількісні аналізи повинні пройти процес верифікації перед їх введенням у клінічну лабораторну практику.

Окрім верифікації прецизійності, також повинні бути проведені експерименти порівняння методів між порівняльними та кандидатами. Для цього необхідно проаналізувати щонайменше 10 зразків для кожної категорії з рівномірним розподілом по всіх категоріях. Через природу порядкових результатів статистичні інструменти, які використовуються для експериментів порівняння кількісних методів, тут не можна застосувати.

### Інтерпретація результатів досліджень порівняння напівкількісних методів

Каппа Коена (зважена,  $k$ ) – це статистичний показник надійності між оцінювачами, тобто ступеня згоди між оцінювачами. У клінічній лабораторії він використовується для оцінювання відповідності між якісними або напівкількісними аналізами.

Оскільки в напівкількісних аналізах присутні три або більше впорядкованих категорій, ступінь розбіжностей може змінюватися (наприклад, одна, дві або більше категорій), тому для цієї мети використовується зважена каппа. Верхня межа  $k$  дорівнює +1,00, що виникає лише тоді, коли обидва методи ідеально узгоджуються. Нижня

межа  $k$  лежить між 0 і 1,00. Значення  $k$ , рівне 0, означає, що угода склалася лише випадково. Аналогічно, від'ємні значення вказують на рівень згоди, який можна було б очікувати лише випадково.

У літературі можна знайти різні інтерпретації для  $k$ , наприклад, запропоновані Ландісом і Кохом: 0,8–1,0 вказує на майже ідеальну згоду, 0,6–0,8 істотна згода, 0,4–0,6 помірна згода, 0,2–0,4 справедлива згода, від нуля до 0,2 незначна згода та нульова або нижча – погана згода.

При інтерпретації результату для каппа Коена корисно перевірити 95 % ДІ. Якщо нижня межа більше 0,70, каппа перевищує 0,70 з 95 % впевненістю, що забезпечує високий рівень впевненості в узгодженні двох методів. Якщо каппа більше 0,70, але нижня межа  $k$  менша за 0,70, не можна стверджувати з 95 % впевненістю, що узгодження між кандидатом і порівняльним методом так само добре, як спочатку вказано каппа. Специфікації для зваженої каппи дедалі більш жорсткі, ніж для звичайної незваженої каппи. Отже, значення 0,70 – 0,90 є цільовим, що вказує на значну або майже повну узгодженість.

Експеримент порівняння методів проводять зі зразками пацієнтів, еталонними матеріалами або матеріалами для перевірки кваліфікації (PT).

CLSI рекомендує продовжувати тестування, поки не буде отримано принаймні 50 позитивних зразків як за допомогою тесту, так і за допомогою порівняльного методу при оцінці чутливості.

Специфічність методу-кандидата оцінюють шляхом вимірювання не менше 50 зразків з негативними результатами, отриманими порівняльним методом Крім того, рекомендується також включити зразки з концентраціями навколо граничного значення (наприклад, 20 %) для оцінювання ефективності методу-кандидата при цій критичній концентрації. Інші автори припускають, що аналізу 40 зразків (10 позитивних, 10 слабо позитивних і 20 негативних) достатньо для рутинного клінічного тестування.

Для статистичного оцінювання ступеня згоди (або між порівняльними методами і методами-кандидатами, або між методом-кандидатом і діагностичними критеріями) використовується каппа Коена.

Значення  $k$  Коена  $> 0,60$ – $0,80$  вказує на значну або майже повну узгодженість. У зв'язку з цим важливо перевірити 95 % ДІ  $k$ , оскільки це дасть цінну інформацію. Якщо нижня межа ДІ  $> 0,60$ ,  $k$  перевищує 0,60 з 95 % впевненістю можна стверджувати, що забезпечується високий рівень впевненості в узгодженні двох методів. Якщо  $k$  більше 0,60, але нижня межа ДІ менша за 0,60, ця узгодженість не може відповідати 95 % впевненості.

Узгодження між методами, коли діагноз для груп обстежуваних є невідомим.

ТАБЛИЦЯ 6.1. ТАБЛИЦЯ СПРЯЖЕНОСТІ 2x2, КОЛИ СПРАВЖНІЙ ДІАГНОЗ НЕВІДОМИЙ (УЗГОДЖЕННЯ МІЖ МЕТОДАМИ)

Досліджуваний метод	Метод порівняння		
	Позитивні	Негативні	Разом
Позитивні	a	b	a + b
Негативні	c	d	c + d
Разом	a + c	b + d	n

Результати підсумовуються в таблиці 2x2, що демонструє ступінь узгодження між порівняльним методом і методом-кандидатом:

$$100 \times \frac{(a + d)}{n}$$

де: a – кількість позитивних результатів для обох методів, d – кількість негативних результатів для обох методів і n – загальна кількість вимірювань.

Оскільки загальна згода не характеризує належним чином узгодження двох якісних методів, рекомендується також звітувати про «PPA» (Відсоток позитивної узгодженості) і «NPA» (Відсоток негативної узгодженості)

$$PPA = 100 \times \frac{a}{(a + c)}$$

$$NPA = 100 \times \frac{d}{(b + d)}$$

де: b – кількість негативних результатів, отриманих за допомогою порівняльного методу, але позитивних у методі-кандидаті, і c – кількість результатів, позитивних при порівняльному методі, але негативних при методі-кандидаті.

У вебдодатку наведені робочі інструменти для оцінки узгодженості.

При верифікації мають бути перевірені та підтверджені прецизійність та щонайменше дві функціональні характеристики методики: діагностична чутливість та діагностична специфічність.

Діагностична чутливість та діагностична специфічність є характеристиками узгодження між методом дослідження та відомим діагно-

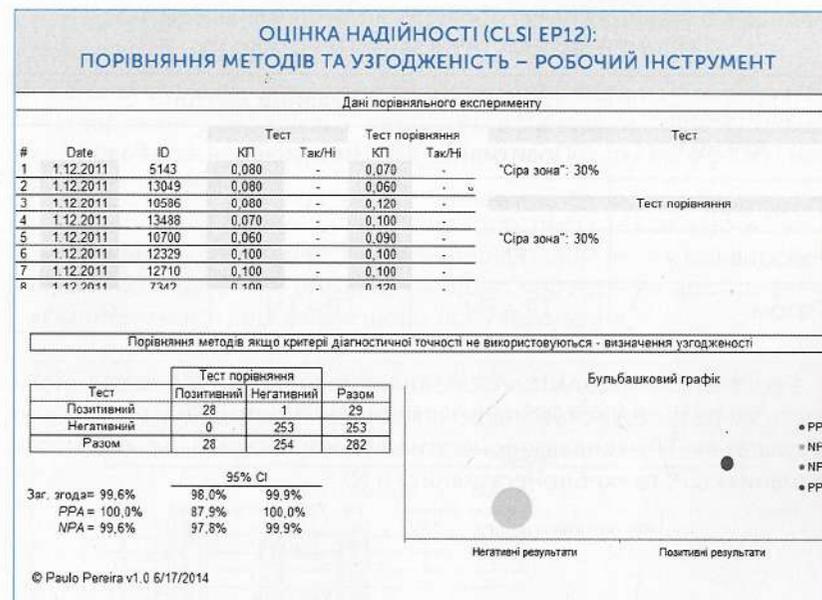


РИС 6.1. РОБОЧИЙ ІНСТРУМЕНТ ПОРІВНЯННЯ ЯКІСНИХ МЕТОДІВ



РИС 6.2. КАЛЬКУЛЯТОР ОЦІНЮВАННЯ НАДІЙНОСТІ ЗА EP12

зом. Результати підсумовуються в таблиці 2x2, яка порівнює результати з критеріями діагностичної точності.

ТАБЛИЦЯ 6.2. ТАБЛИЦЯ СПРЯЖЕНОСТІ 2x2, КОЛИ СПРАВЖНІЙ ДІАГНОЗ ВІДОМИЙ (ДІАГНОСТИЧНІ ХАРАКТЕРИСТИКИ ТЕСТУ)

Істинний діагноз	Досліджуваний метод		
	Позитивні	Негативні	Разом
Позитивні	TP	FP	TP + FP
Негативні	FN	TN	TN + FN
Разом	TP + FN	FP + TN	n

Ефективність підсумовується з точки зору чутливості та специфічності, які розраховуються на основі кількості «справжніх позитивних результатів» (TP), «справжніх негативних результатів» (TN), «хибно-позитивних» (FP) та «хибно-негативних» (FN):

$$Sensitivity = 100 \times \frac{TP}{(TP + FN)}$$

$$Specificity = 100 \times \frac{TN}{(FP + TN)}$$

Поширеність захворювання (в досліджуваній групі) оцінюється за:

$$Prevalence = 100 \times \frac{TP + FN}{n}$$

Прогностична цінність позитивного тесту (PPV) описує частку пацієнтів з позитивним результатом тесту, у яких присутнє захворювання, тоді як прогностична цінність негативного тесту (NPV) описує частку пацієнтів з негативним результатом тесту, у якого ця умова відсутня

$$PPV = 100 \times \frac{TP}{(TP + FP)}$$

$$NPV = 100 \times \frac{TN}{(FN + TN)}$$

Важливо зазначити, що ці оцінені показники ефективності можуть мати зсув, який не покращується простим збільшенням кількості вибірки. Скоріше за це відповідають такі питання, як вибір пацієнта, референтний тест, а також потік пацієнтів і час, і, лише враховуючи ці фактори, можна зменшити або усунути зсув.

Методи верифікації якісних тестів у дещо спрощеному варіанті використовуються в «Процедурі верифікації методики дослідження біологічного матеріалу з метою лабораторної діагностики коронаві-

русної хвороби (COVID-19)» (Додаток 15 до Стандартів медичної допомоги «Коронавірусна хвороба (COVID-19)»<sup>1</sup>.

Якщо функціональні характеристики методики не можуть бути підтверджені у верифікації, необхідно встановити причину невідповідності та, за необхідності, повторити процедуру верифікації методики дослідження.

При пошуку причини невідповідності, перш за все, необхідно переконатися у чіткому дотриманні персоналом інструкції виробника медичних виробів та обладнання щодо проведення дослідження та компетентності персоналу щодо його виконання.

#### ОЦІНКА ДІАГНОСТИЧНИХ ХАРАКТЕРИСТИК ТЕСТУ (CLSI EP12): ЧУТЛИВІСТЬ, СПЕЦИФІЧНІСТЬ ТА ІН. – РОБОЧИЙ ІНСТРУМЕНТ

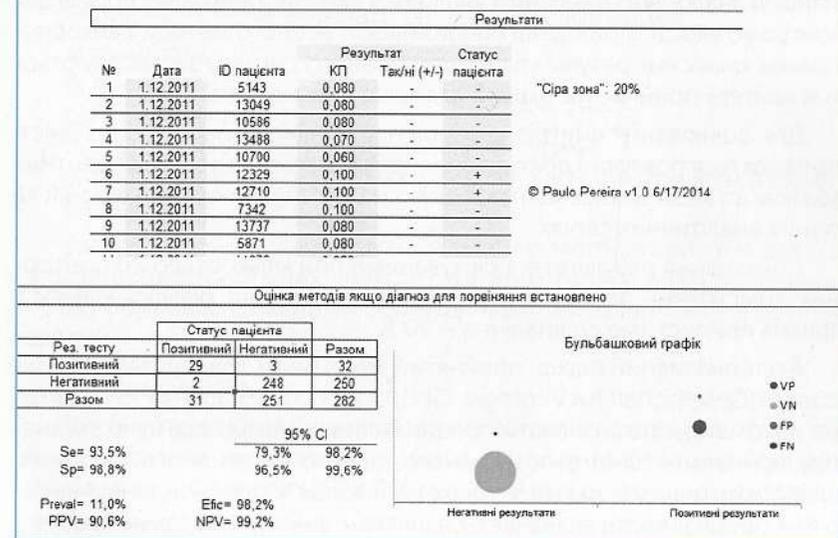


РИС. 6.3. ОЦІНКА ДІАГНОСТИЧНИХ ХАРАКТЕРИСТИК ЯКІСНИХ ТЕСТІВ

#### Прецизійність

Оцінити прецизійність якісних і напівкількісних даних важко, оскільки результати є або біноміальними (позитивні/негативні), або порядковими (можливо більше двох альтернативних значень). Визначення прецизійності необхідне для якісних і напівкількісних тестів, якщо результати отримані на основі кількісного значення, такого як оптична густина OD, або якщо в інструкції виробника описані аналітичні характеристики прецизійності для аналізу. Найпоширенішим підходом є оцінка преци-

<sup>1</sup> Джерело: <https://www.dec.gov.ua/?ZG93bmxyYWQ=d3AtY29udGVudC91cGxvYWRzLzlwMjlvMDlvMjAyMF83MjJfc3RhbmRhcncFY292aWRiMTkucGRm>

ОЦІНКА ДІАГНОСТИЧНИХ ХАРАКТЕРИСТИК ТЕСТУ ( ТАБЛИЦЯ 2X2 )						
Thomas Keller, ACOMED statistik, www.acomed-statistik.de.						
Тест	Захворювання (Disease)		Разом	Довірчий інтерв. CI (case control study y/n)	95%	Введіть преваленс.
	D+	D-				
T+	191	1	192			2%
T-	1	112	113		y	
Разом	192	113	305			
Чувливість	0.9946	0.9713	0.9999	5.750	99.48%	(95%-Довірчий інтерв. CI: 97.13% - 99.99%)
Специфічність	0.9912	0.9517	0.9996	4.719	99.12%	(95%-Довірчий інтерв. CI: 95.17% - 99.98%)
PPV	0.9964	0.2453	0.9417	0.036	69.64%	(95%-Довірчий інтерв. CI: 24.58% - 94.17%)
NPV	0.9999	0.9982	1.0000	5.140	99.99%	(95%-Довірчий інтерв. CI: 99.92% - 100.0%)
Preval	0.6296			3.00	Prevalenс	2.0%
DLR+	112.4115	15.9725	791.1322		DLR+	112.411 (95%-Довірчий інтерв. CI: 15.972 - 791.132)
DLR-	0.0053	0.0067	0.0371		DLR-	0.005 (95%-Довірчий інтерв. CI: 0.001 - 0.037)

РИС. 6.4. КАЛЬКУЛЯТОР ДІАГНОСТИЧНИХ ХАРАКТЕРИСТИК ЯКІСНИХ ТЕСТІВ

зійності аналогічно кількісним аналізам з використанням матеріалів для контролю якості. При оцінці прецизійності якісного методу з використанням кількісних результатів вимірювального сигналу застосовуються ті ж критерії прийнятності, що й для кількісних методів.

Для оцінювання внутрішньосерійної та міжсерійної прецизійності проводяться повторні дослідження зразка пацієнта або референтних зразків, а також зразків контрольного матеріалу виробника в одній та різних аналітичних серіях.

Співпадіння результатів з очікуваними при кількості до 20 повторних вимірювань повинно бути 100 %, при оцінці більшої кількості зразків припустиме співпадіння – 90 %.

Альтернативний підхід, прийнятий Німецьким товариством з вірусології (Gesellschaft für Virologie; GfV), описується Рабенау та ін. Автори рекомендують визначати прецизійність в межах серії шляхом аналізу принаймні одного позитивного, одного негативного та одного слабопозитивного зразка у трьох реплікатах у перший день. Міжсерійна прецизійність визначається шляхом аналізу тих самих зразків у одиничних зразках у два додаткові дні.

#### Верифікація прецизійності та точності якісних методів з бінарними результатами (Rabenau et al.)

##### Матеріали

Контролі виробника (K+ та K-)  
Позитивний зразок, Слабопозитивний зразок, Негативний зразок

##### Протокол

1 день (K+, K+, K+, Kсп+, Kсп+, Kсп+, K-, K-, K-)  
2 день (K+, Kсп+, K-)  
3 день (K+, Kсп+, K-)

Критерії верифікації – правильні результати (K+, Kсп+ позитивні, K- негативні) у всіх зразках

## 7. ОСОБЛИВОСТІ ОЦІНЮВАННЯ МЕТОДІВ В ОКРЕМИХ ГАЛУЗЯХ ЛАБОРАТОРНОЇ МЕДИЦИНИ

### 7.1. ОЦІНКА ГЕМАТОЛОГІЧНИХ МЕТОДІВ

З медичної точки зору цінність автоматизованого гематологічного аналізатора полягає в тому, щоб надати лікарям та іншим медичним працівникам надійні гематологічні дані для лікування пацієнтів. Надійні дані залежать від надійної конструкції системи, яка спочатку перевіряється виробником за допомогою офіційних протоколів/процедур дослідження, а потім перевіряється лабораторією – кінцевим користувачем.

Верифікація може включати оцінку таких показників :

- лінійність;
- прецизійність;
- перенесення;
- зсув в порівнянні з еталонними методами;
- порівнянність із іншим методом, який використовується в лабораторії.

Загальні принципи верифікації кількісних методів придатні для оцінювання гематологічних аналізаторів. Крім того, є деякі особливості, що обумовлені певними обмеженнями технології та стандартних зразків.

#### Прецизійність

Виконуйте відповідні вимірювання 10 разів поспіль на трьох або більше зразках, які включають низьку, високу та середню концентрації аналіту. Обчисліть SD і CV, порівняйте з ціллю щодо якості для тесту.

#### Лінійність

Лінійність демонструє ефекти розведення. Підготуйте зразок з високою концентрацією досліджуваної речовини і якомога точніше зробіть серію розведень у плазмі, щоб отримати 10 зразків з рівномірно розташованими рівнями концентрації від 10 % до 100 %. Кожен зразок виміряйте три рази та обчисліть середнє значення. Порівняйте з ціллю щодо якості для тесту. Перевірка графіка покаже, чи є лінійність у всьому діапазоні, чи вона обмежена частиною діапазону.

#### Перенос

Перенесення (Carryover) вказує, наскільки ймовірний вплив попереднього зразка на вимірювання аналіту в наступному зразку. Вимі-

рятьте зразок з високою концентрацією у трьох реплікатах, а потім так само зразок з низькою концентрацією аналіту та обчисліть процент переносу:

$$\text{Carryover (\%)} = \frac{l_1 - l_3}{h_3 - l_3} \times 100.$$

### Аналізатори ШОЕ

Для оцінювання аналізаторів ШОЕ використовуються як методозалежні штучні контрольні матеріали виробника (правильність та прецизійність), так і зразки свіжої крові пацієнтів для порівняння з референтним методом Вестергрена.

Зразки свіжої крові людини забираються безпосередньо в зазначені контейнери з антикоагулянтом або розчинником відповідно до інструкцій виробника. Зразки повинні охоплювати діапазон результатів седиментації від 15 до 105 мм, а також нормальні значення, приблизно з однаковою кількістю зразків у кожному квартилі. Кров для тесту ШОЕ слід зберігати при температурі навколишнього середовища (18 – 25 °C) до її проведення, а аналізи повинні розпочатися протягом не більше ніж чотирьох годин після забору. Якщо тест-система не включає пристрій автоматичного змішування, зразки слід змішати, як зазначено для стандартизованого методу (наприклад, принаймні 12 повних переворотів для контейнера з кров'ю розміром 10-13 мм x 75 мм і більше переворотів, якщо внутрішній діаметр контейнера менший).

Прецизійність слід вимірювати при повторних вимірюваннях (принаймні 10, якщо можливо) одного зразка з кожного квартиля. Прецизійність стандартизованого методу повинна бути визначена аналогічним чином для порівняльних цілей.

Порівнянність між тестовою системою та стандартизованим методом повинна бути перевірена паралельно на щонайменше в 60 зразках пацієнтів із широким спектром захворювань, а результати ШОЕ мають розподілятися рівномірно в діапазоні від 15 до 105 мм. Іноді зразки крові можуть не дати чіткого розділу плазми/еритроцитів після седиментації; якщо це відбувається в тестовій системі або стандартизованому методі, пару значень слід вилучити з набору даних.

Лінійну регресію слід обчислювати за допомогою методу, відповідного для таких даних, наприклад, Passing-Bablok, Bland Altman.

Метод випробування та стандартизований метод можна вважати еквівалентними, якщо: 1. 95 % довірчий інтервал для  $b$  містить значення 1 (гіпотезу  $b = 1$  можна вважати правильною); 2. 95 % довірчий інтервал для  $a$  містить значення 0 (гіпотезу  $a = 0$  можна вважати правильною).

## 7.2. ОЦІНКА МОЛЕКУЛЯРНО-ДІАГНОСТИЧНИХ МЕТОДІВ

Залежно від типу молекулярного методу, можуть бути застосовані загальні методи верифікації, що описані в розділах відповідно для кількісних та якісних методів дослідження.

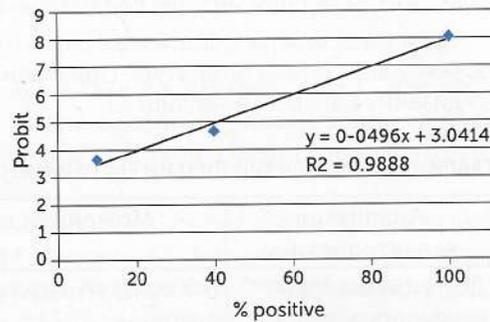
Крім того, можна обрати інші варіанти оцінювання, що рекомендовані у відповідній літературі. Приклади можливого альтернативного дизайну наведені в таблиці 7.1.

ТАБЛИЦЯ 7.1. МОЖЛИВИЙ ДИЗАЙН ЕКСПЕРИМЕНТІВ З ВЕРИФІКАЦІЄЮ (E/BURD, 2010)

Аналітичні характеристики	Можливий дизайн експериментів з верифікації
Діапазон вимірювання, дослідження лінійності (для кількісних аналізів)	5-7 концентрацій у заявленому лінійному діапазоні. 2 реплікатів кожної концентрації
Аналітична чутливість, дослідження ліміту виявлення	Не обов'язкове для лабораторії
Прецизійність (реплікаційний експеримент)	Для якісного тесту тестуйте 1 контрольний зразок/день протягом 20 днів або дублікати контрольних зразків протягом 10 днів: для кількісного тестування перевірте 2 зразки в кожній з 2 концентрацій (4 зразки) плюс один контроль протягом 20 днів або перевірте 2 концентрації в триплікаті протягом 5 днів
Аналітична специфічність, дослідження інтерференції	Не обов'язкове для лабораторії
Точність дослідження, порівняння методів	20 зразків пацієнтів у межах вимірювального інтервалу або референтних матеріалів в 2 концентраціях (низьких і високих) в дублікатах більше 2-5 аналітичних серій Зазначений виробником референтний інтервал може бути «перенесений», якщо заявлений референтний інтервал може бути застосовано до популяції, яку обслуговує клінічна лабораторія: при бажанні здійснити верифікацію, тестуйте 20 зразків, що репрезентують популяцію: якщо популяція відрізняється (див. Верифікація референтних матеріалів)

При необхідності визначення ліміта детекції (LOD) можна використовувати також метод пробіт-регресія. Досліджуються 10-кратні розведення в 10 реплікатах та за допомогою відповідного ПЗ розраховується ліміт детекції.

CFU/ml	% positive	Probit
1,000	100	8.09
100	100	8.09
10	40	4.75
1	10	3.72



Note: Probits should not use 100 % or 0 % in analysis.

%	0	1	2	3	4	5	6	7	8	9
0	-	2.67	2.95	3.12	3.25	3.36	3.45	3.52	3.59	3.66
10	3.72	3.77	3.82	3.87	3.92	3.96	4.01	4.05	4.08	4.12
20	4.16	4.19	4.23	4.26	4.29	4.33	4.36	4.39	4.42	4.45
30	4.48	4.50	4.53	4.56	4.59	4.61	4.64	4.67	4.69	4.72
40	4.75	4.77	4.80	4.82	4.85	4.87	4.90	4.92	4.95	4.97
50	5.00	5.03	5.05	5.08	5.10	5.13	5.15	5.18	5.20	5.23
60	5.25	5.28	5.31	5.33	5.36	5.39	5.41	5.44	5.47	5.50
70	5.52	5.55	5.58	5.61	5.64	5.67	5.71	5.74	5.77	5.81
80	5.84	5.88	5.92	5.95	5.99	6.04	6.08	6.13	6.18	6.23
90	6.28	6.34	6.41	6.48	6.55	6.64	6.75	6.88	7.05	7.33

	y (Probit)	x (LOD)
99 %	7.33	86.49194
95 %	6.64	72.58065
90 %	6.28	65.32258
85 %	5.84	56.45161
80 %	5.52	50
70 %	5.25	44.55645

Рис. 7.1. Визначення ліміта детекції (LOD) – пробіт – регресія (10-кратні розведення в 10 реплікатах)

Окремі методи молекулярної діагностики є багатостадійними та складними, і вимагають розробки окремої програми верифікації, що можуть мати специфічні для них показники і вимоги.

В якості прикладу можна навести варіант верифікації методики з визначення мутацій резистентності ВІЛ-1 до АРВ-препаратів методом секвенування геному ВІЛ.

Проводиться верифікація аналітичних характеристик методу у порівнянні із заявленими розробником та цілями аналітичної якості для наступних характеристик:

Функціональна характеристика	Визначення	Критерії прийнятності
Точність	Узгодження між результатом та очікуваним значенням (для аналізів генотипування – нуклеотидна послідовність)	$\geq 90\%$ порівнянь (базові виклики та виклики мутації стійкості до лікарських засобів) для кожного зразка (між результатом тесту та еталонним або стандартним результатом) повинні бути принаймні на $98\%$ однаковими
Чутливість ампліфікації	Відсоток успішних (точних та відтворюваних) тестів генотипування серед зразків із певним діапазоном вірусного навантаження	плазма: більше $90\%$ зразків або копій є позитивними щодо вірусного навантаження від 1000 до 5000 копій / мл та $> 95\%$ для вірусного навантаження, що перевищує 5000 копій / мл
Прецизійність внутрішньо-серійна та міжсерійна	Ступінь наближення один до одного результатів повторних вимірювань одного матеріалу в одній аналітичній серії або в різних аналітичних серіях	$\geq 90\%$ ( $9$ із $10$ ) парних порівнянь для кожного зразка повинні бути принаймні $98\%$ однаковими

### Прецизійність

Для оцінювання внутрішньосерійної та міжсерійної прецизійності проводяться повторні дослідження зразка пацієнта або референтних зразків, а також зразки контрольного матеріалу виробника в одній та різних аналітичних серіях.

Співпадіння результатів з очікуваними при кількості до 20 повторних вимірювань повинно бути 100 %, при оцінці більшої кількості зразків припустиме співпадіння – 90 %.

### Точність

Для оцінювання точності проводяться дослідження референтних зразків, що мають визначені послідовність нуклеотидів та види мутацій. Можуть використовуватися зразки ЗОЯ або іншого походження.

Задовільним є  $\geq 90\%$  порівнянь (базові виклики та виклики мутації стійкості до лікарських засобів) для кожного зразка (між результатом тесту та еталонним або стандартним результатом), що повинні бути принаймні у 98 % однаковими.

## 7.3. ОЦІНКА МІКРОБІОЛОГІЧНИХ МЕТОДІВ

Верифікація комерційних систем ідентифікації мікробів (далі – MIS) та систем тестування чутливості до антимікробних препаратів (далі – ASTS).

Верифікації підлягають як автоматичні системи, яких стає дедалі більше в клінічних лабораторіях, так і ручні методи ідентифікації мікробів (далі ID) та тести на чутливість до антимікробних препаратів (далі – AST), включаючи дискову дифузію та градієнтні дифузійні смужки.

Верифікація включає оцінку точності, прецизійності (відтворюваності) та діапазону звітних результатів випробувань, а також перевірку того, що референтні діапазони виробника (нормальні значення) відповідають популяції пацієнтів, які обстежуються в лабораторії.

Під час перевірки MIS або ASTS точність і прецизійність (відтворюваність) є найбільш застосовними характеристиками.

Лабораторії повинні перевірити, що їхня система може правильно ідентифікувати більш поширені види, і повинні мати політику для підтвердження ідентифікації незвичайних видів, які не були перевірені під час верифікації.

Верифікацію необхідно виконати для:

- тест-системи, введеної в лабораторії для вимірювання аналіту, який лабораторія раніше не вимірювала (наприклад, початкове введення комерційного MIS або ASTS в лабораторію);
- тест-системи введено в лабораторії для заміни тесту, який лабораторія в даний час виконує, на альтернативну тестову систему (наприклад, лабораторія переходить з однієї комерційної системи на іншу комерційну систему для MIS та/або ASTS);
- аналіт, щойно доданий до раніше використовуваної тестової системи, яка може вимірювати декілька аналітів (наприклад, новий антимікробний агент додається до комерційного ASTS).

ТАБЛИЦЯ 7.2. ВЕРИФІКАЦІЯ МІКРОБІОЛОГІЧНИХ МЕТОДІВ

Аналітичні характеристики	Рекомендації для методів ідентифікації (MIS)	Рекомендації для систем визначення антибіотикочутливості (ASTS)
Точність	Узгодження з методом порівняння ID; може бути визначено як відсоток узгодженості з певними довірчими інтервалами (критерії прийнятності >90 %)	Категоріальна узгодженість CA (S, I, R) Есенціальна узгодженість EA (MIC), якщо може бути застосовано (критерії прийнятності див. далі)
Прецизійність (Відтворюваність)	Відтворювана ідентифікація мікроорганізмів (деякі специфічні реакції можуть відрізнятися)	Відтворювана інтерпретація (S, I, чи R); Значення MIC відтворюється з прийнятною варіацією $\pm 1$ розведення для бактерій та $\pm 2$ розведення для дріжджів

Коли для виконання одного тесту використовується кілька інструментів (включаючи ту ж саму платформу та модель), лабораторія повинна перевірити характеристики продуктивності для кожного інструменту.

### Повне дослідження з верифікації

Повне дослідження з верифікації є виправданим, коли нова система вводиться в лабораторії. Перш ніж повідомляти результати пацієнта, отримані на новій системі, слід перевірити здатність лабораторії отримати результати, еквівалентні заявам виробника щодо MIS та/або ASTS. Тестування має бути зосереджено на передбачуваному використанні системи у випробувальній лабораторії.

#### Рекомендації щодо обсягу тестування:

**Точність:** мінімум 30 клінічних ізолятів, порівняння з існуючою системою чи референтним методом.

**Прецизійність:** дослідження 5 ізолятів (контрольних штамів та/або клінічних ізолятів) 3 рази протягом одного чи кількох днів та одним або різними операторами.

### Обмежене дослідження з верифікації

Обмежене дослідження з верифікації може виконуватися, коли нові функції та/або функціональні можливості додаються до попередньо перевіреної системи з програмою забезпечення якості. Об-

межене тестування особливо доречно, коли важливі етапи, такі як підготовка інокулята, залишаються незмінними.

Тестування має зосередитися на тих функціях, які були додані або змінені, щоб переконатися, що система продовжує відповідати очікуванням. Наприклад, якщо новий антимікробний агент додається до процедури дослідження, тільки цей агент потрібно перевірити.

Обмежена верифікація також доречна, коли інструмент замінюється новим приладом тієї ж платформи та моделі з використанням тих самих витратних матеріалів (наприклад, панелей або реагентів) або коли додаються додаткові інструменти тієї ж самої платформи та моделі. Користувач повинен переконатися, що характеристики нового інструменту еквівалентні поточному інструменту.

#### Рекомендації щодо обсягу тестування:

**Точність:** мінімум 10 клінічних ізолятів;

**Прецизійність:** дослідження контрольних штамів 3 рази протягом 1 дня.

#### Зміни, для яких не потрібні верифікаційні дослідження

Верифікаційні дослідження не потрібні після планового технічного обслуговування, ремонту або оновлення приладів чи програмного забезпечення, якщо контрольна панель не змінюється. Після ремонту інструментів рекомендується проводити щоденне тестування контролю якості протягом одного-п'яти днів.

Коли систему переміщують з одного місця в інше в лабораторії, рекомендовано тестування для підтвердження роботи приладу. Для забезпечення прийнятної роботи має бути достатньо одного дня тестування контролю якості безпосередньо після переїзду та до або одночасно з тестуванням пацієнта.

Внутрішні самоперевірки приладів також можуть бути застосовані.

#### Відбір мікроорганізмів для тестування

Лабораторія повинна вибрати кількість і тип організмів для тестування відповідно до обсягу тестування, яке буде виконано за допомогою системи в лабораторії.

Тестування штамів контролю якості є лише частиною верифікації і само по собі не-достатньо для верифікації MIS або ASTS. Необхідно також провести додаткове тестування клінічних або раніше охарактеризованих вихідних ізолятів.

Клінічні ізоляти, відібрані для верифікаційних досліджень, повинні бути свіжими і можуть включати вихідні ізоляти з відомими ідентифікаторами або моделями АСТ.

#### Верифікація систем визначення антибіотикочутливості

Відбір організмів для тестування має бути репрезентативним щодо тих, які клінічно показані для антимікробних агентів, які необхідно перевірити.

Рекомендуються для використання свіжі клінічні ізоляти (тобто ті, що виділені зі звичайних клінічних культур протягом семи днів після дослідження та не заморожені). На додаток до свіжих ізолятів, ізоляти з лабораторної колекції заморожених запасів можна використовувати для збільшення різноманітності роду і видів, а також для збільшення кількості стійких фенотипів. Можуть також використовуватися ізоляти, отримані в результаті досліджень ЗОЯ або інших джерел.

Усі заморожені ізоляти перед тестуванням необхідно двічі пересівати. Субкультури для тестування повинні бути не старше 24 годин.

Для ASTS необхідно перевірити мінімум 30 ізолятів на панель. Для більшості антимікробних засобів це призведе до 30 результатів АСТ. Для агентів з меншою кількістю результатів можуть бути перевірені додаткові ізоляти або результати контролю якості, щоб довести кількість результатів для кожного антимікробного засобу до 30. Не більше 50 % з 30 результатів повинні бути результатами контролю якості.

Таблиця 7.4. Кількість ізолятів та режими тестування при верифікації (CLSI M52)

Тип змін	Точність	Прецизійність (відтворюваність)
Повна верифікація		
<ul style="list-style-type: none"> <li>• New System</li> <li>• Модифікація інструменту</li> <li>• Зміна методу</li> <li>• Нова група організмів</li> </ul>	Мінімум 30 ізолятів (30 результатів / агентів)	Дослідження 5 ізолятів x 3 рази
Обмежена верифікація		
<ul style="list-style-type: none"> <li>• Новий антимікробний агент</li> <li>• Другий інструмент</li> </ul>	Мінімум 10 ізолятів	Рутинні контрольні зразки QC (20 днів або 3x5 днів)

Прецизійність (відтворюваність) визначається як близькість узгодження між результатами послідовних вимірювань одного і того ж аналізу.

Для ASTS точність (відтворюваність) повинна брати до уваги звичайну варіацію  $\pm 1$  розведення для антибактеріальних засобів ( $\pm 2$  подвійних розведень для протигрибкових засобів) для тестування.

Наприклад, показники MIC 1 мкг/мл або 2 мкг/мл будуть вважатися еквівалентними результатами.

Для перевірки точності (відтворюваності) ASTS може використовуватися репрезентативна підгрупа штамів QC та клінічних ізолятів.

Щонайменше п'ять ізолятів (QC або клінічні штами) повинні бути перевірені тричі кожен. Наприклад, два штами QC і три клінічні ізоляти можуть бути відібрані та запущені в новій системі. Ці штами слід перевірити три рази, використовуючи окремі препарати для посіву. Тестування можна проводити протягом одного або кількох днів.

Критерії прийнятності для систем тестування на чутливість до антимікробних препаратів:

Узгодженість категорій (category agreement, CA) – узгодженість чутливих, проміжних та резистентних результатів між досліджуваним методом та референсним методом.

Есенціальна узгодженість (essential agreement, EA) – мінімальна інгібуюча концентрація (MIC), отримана за допомогою системи тестування на чутливість до антимікробних препаратів, що перебуває у межах одного розведення для бактерій і двох розведень для дріжджів від значення MIC, встановленого за допомогою контрольного методу.

#### Помилки при визначенні антибіотикочутливості:

**«Невідповідність»** (MD або VMD) використовується в цьому документі для вказівки на різницю між ASTS, тоді як **«помилка»** (ME або VME) вказує на розбіжність між ASTS та еталонним методом;

**Велика невідповідність (major discrepancy, MD)** – розбіжність в результатах тестів, коли нова система тестування чутливості до антимікробних препаратів (ASTS) дає результат «стійкий», а результат методу порівняння є «чутливий»;

**Велика помилка (major error, ME)** – помилка, коли результат референтного методу є «чутливий», а система тестування чутливості до антимікробних препаратів (ASTS) отримала результат «стійкий»;

**Незначна невідповідність (minor discrepancy, mD)** – розбіжність у результаті тесту між новою системою тестування чутливості до антимікробних препаратів (ASTS) та результатом методу порівняння (з існуючого комерційного ASTS), де один результат є «проміжний», а інший – «чутливий» або «стійкий»;

**Незначна помилка (minor error, mE)** – різниця в результатах тесту між новою системою тестування чутливості до антимікробних препаратів (ASTS) та еталонним результатом AST, де один результат є «проміжний», а інший – «чутливий» або «стійкий»;

**Дуже велика невідповідність (very major discrepancy, VMD)** – розбіжність у результатах тесту, коли результат нової системи тестування на чутливість до антимікробних препаратів (ASTS) є «чутливий», а результат існуючої системи ASTS є «стійкий»;

**Дуже велика помилка (very major error, VME)** – помилка, коли результат контрольного методу є «стійкий», а результат системи тестування на чутливість до антимікробних препаратів (ASTS), що оцінюється, є «чутливий».

**ТАБЛИЦЯ 7.5. ПОМИЛКИ (ПРИ ПОРІВНЯННІ З РЕФЕРЕНТНОЮ МЕТОДИКОЮ) АБО НЕВІДПОВІДНОСТІ (ПРИ ПОРІВНЯННІ З ІНШОЮ МЕТОДИКОЮ) ПРИ ВЕРИФІКАЦІЇ СИСТЕМ АНТИБІОТИКОЧУТЛИВОСТІ**

Помилка або невідповідність	Результати антибіотикочутливості	
	Референтна методика/ методика порівняння	Досліджувана методика
Дуже велика VME/VMD	R	S
Велика ME/MD	S	R
Незначна mE/mD	S	I
	R	I
	I	S
	I	R

Усі ASTS, які отримали дозвіл FDA, продемонстрували значну еквівалентність еталонному методу (тобто із застосуванням мікробульйону або розведення агару), хоча не всі ASTS були розроблені за одним і тим же еталонним методом.

ASTS слід вважати перевіреними для конкретних перевірених комбінацій ліків/мікроорганізмів, коли:

- CA та EA (якщо повідомляють про MIC) становлять  $\geq 90\%$  порівняно з поточною системою;
- рівень VMD або VME становить  $< 3\%$  від загальної кількості резистентних ізолятів;
- Частота MD або ME становить  $< 3\%$  від загальної кількості чутливих ізолятів.

Якщо лабораторні критерії відповідали очікуванім, значить, система успішно верифікована. Якщо будь-який із цих критеріїв не виконано, необхідно вжити додаткових заходів. Якщо тестується мінімальна кількість ізолятів, що дає 30 результатів на протимікробний засіб,

лише три результати можуть бути поза CA або EA, щоб відповідати критеріям  $\geq 90\%$  на антимікробний засіб.

Можна прийняти лише один MD або ME і один VMD або VME.

Може знадобитися додаткове тестування, щоб продемонструвати прийнятну роботу, якщо початкове тестування не відповідає критеріям прийнятності. Додаткове тестування може включати повторне тестування невідповідних ізолятів, тестування додаткових ізолятів або тестування за еталонним методом, коли початкове тестування порівнювалося з іншим ASTS.

**ДОДАТОК 8.1. ЗНАЧЕННЯ БІОЛОГІЧНОЇ ВАРІАЦІЇ ТА ПРИПУСТИМИ АНАЛІТИЧНІ СПЕЦИФІКАЦІЇ ЩОДО ЗСУВУ V(%), НЕПРЕЦИЗИЙНОСТІ I(%), МАКСИМАЛЬНО ПРИПУСТИМОЇ НЕВИЗНАЧЕНОСТІ MAU ТА ЗАГАЛЬНОЇ АНАЛІТИЧНОЇ ПОМИЛКИ ТЕА ДЕЯКИХ ТЕСТІВ**

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM			Бажана мета за EFLM			Бажана мета за Ricos, 2014			MAU %	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCPA TEa(%)
		CW <sub>i</sub>	CW <sub>g</sub>	I(%)	V(%)	TEa(%)	I(%)	V(%)	TEa(%)	I(%)	V(%)	TEa(%)				
1	ALB (альбумін)	2.5	4.9	1.88	2.06	5.16	1.25	1.38	3.44	1.6	1.43	4.07	3.75	20	10	6
2	ALP (лужна фосфатаза)	5.3	24.7	3.98	9.47	16.03	2.65	6.32	10.69	3.23	6.72	12.04	7.95	21	30	12
3	ALT (аланінамінотрансфераза)	10.1	29.3	7.58	11.62	24.12	5.05	7.75	16.08	9.7	11.48	27.48	15.15	21	20	12
4	AMY (амілаза)	6.6	30.2	4.95	11.59	19.76	3.30	7.73	13.17	4.4	7.4	14.6	9.9		30	10
5	AST (аспаратаміно-трансфераза)	9.6	20.8	7.20	8.59	20.47	4.80	5.73	13.65	6.15	6.54	16.69	14.4	21	20	12
6	UREA (сечовина)	13.9	21	10.43	9.44	26.65	6.95	6.30	17.76	6.05	5.57	15.55	20.85	21	9	12
7	Ca (кальцій)	1.8	2.7	1.35	1.22	3.44	0.90	0.81	2.30	1.05	0.82	2.55	2.7	21		4
8	MG (магній)	2.9	5.8	2.18	2.43	6.02	1.45	1.62	4.01	1.8	1.8	4.8	4.35	10	25	8
9	IRON (залізо)	20.7	32.3	15.53	14.39	40.00	10.35	9.59	26.67	13.3	8.8	30.7	31.05	15	20	12
10	CRE (креатинін)	4.5	14.1	3.38	5.55	11.12	2.25	3.70	7.41	2.98	3.96	8.87	6.75	20	15	8
11	BD (білрубін прямий)	51.1	57	38.33	28.71	91.94	25.55	19.14	61.30	18.4	14.2	44.5	76.65			20

№	Аналіт	BioVariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCPA TEa(%)			
		CW	CWg	I(%)	V(%)	TEa(%)	I(%)	V(%)	TEa(%)					I(%)	V(%)	%
12	BT (білірубін загальний)	20	26,7	15,00	12,51	37,26	10,00	8,34	24,84	10,9	8,95	26,94	30	22	20	12
13	GLU (глюкоза)	5	8,1	3,75	3,57	9,76	2,50	2,38	6,50	2,8	2,34	6,96	7,5	15	10	8
14	LDH (лактатдегідрогеназа)	5,2	12,6	3,90	5,11	11,55	2,60	3,41	7,70	4,3	4,3	11,4	7,8	18	20	8
15	GGT (γ-глутаміл-трансфераза)	9,1	44,5	6,83	17,03	28,29	4,55	11,36	18,86	6,7	11,06	22,11	13,65	21		12
16	LIP (ліпопротеїн-ліпаза)	9,2	24,8	6,90	9,92	21,30	4,60	6,61	14,20	16,1	11,31	37,88	13,8	25		20
17	Phos (фосфор)	7,8	10,7	5,85	4,97	14,62	3,90	3,31	9,75	4,08	3,38	10,11	11,7	16		8
18	TP (загальний білок)	2,6	4,6	1,95	1,98	5,20	1,30	1,32	3,47	1,38	1,56	3,63	3,9	10	10	5
19	Chol (холестерин)	5,3	16,7	3,98	6,57	13,13	2,65	4,38	8,75	2,98	4,1	9,01	7,95	13	10	6
20	TG (тригліцериди)	19,9	37,4	14,93	15,89	40,51	9,95	10,59	27,01	5,95	9,57	25,99	29,85	16	25	12
21	HDL (ліпопротеїди високої густини)	5,7	24,3	4,28	9,36	16,41	2,85	6,24	10,94	3,65	5,61	11,63	8,55			12

№	Аналіт	BioVariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCPA TEa(%)			
		CW	CWg	I(%)	V(%)	TEa(%)	I(%)	V(%)	TEa(%)					I(%)	V(%)	%
22	LDL (ліпопротеїди високої густини)	8,3	26,6	6,23	10,45	20,72	4,15	6,97	13,81	3,9	5,46	11,9	12,45			
23	UA (сечова кислота)	8,3	22,4	6,23	8,96	19,23	4,15	5,97	12,82	4,3	4,87	11,97	12,45	13	17	
24	P-AMU (панкреатична амілаза)	6,7	25,7	5,03	9,96	18,25	3,35	6,64	12,17	5,9	8	17,7	10,05			
25	CK (креатинкіназа)	15	37,9	11,25	15,29	33,85	7,50	10,19	22,57	11,4	11,5	30,3	22,5			12
26	K (калій)	4,1	4,2	3,08	2,20	7,27	2,05	1,47	4,85	2,3	1,81	5,61	6,15	8		5
27	Na (натрій)	0,5	1	0,38	0,42	1,04	0,25	0,28	0,69	0,3	0,23	0,73	0,75	5		2
28	Cl (хлор)	1,1	1,3	0,83	0,64	2,00	0,55	0,43	1,33	0,6	0,5	1,5	1,65	8	5	3
29	Ca++(кальцій іонізований)	2,18	1,58	1,64	1,01	3,71	1,09	0,67	2,47	0,9	0,6	2	3,27	18\15		
30	Har (галтоглобін)	8,6	39	6,45	14,98	25,62	4,30	9,98	17,08	10,2	10,4	27,3	12,9			
31	S3 (компонент комплекнту C3)	4,6	15,2	3,45	5,96	11,65	2,30	3,97	7,77	2,6	4,1	8,4	6,9			
32	S4 (компонент комплекнту C4)	6,9	24,5	5,18	9,54	18,08	3,45	6,36	12,06	4,5	8,6	16	10,35			

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM			Бажана мета за EFLM			Бажана мета за Ricos, 2014			MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCRA TEa(%)
		CWi	CWg	I (%)	B (%)	TEa(%)	I (%)	B (%)	TEa(%)	I (%)	B (%)	TEa(%)				
33	IgA (імуноглобулін А)	5,7	19,5	4,28	7,62	14,67	2,85	5,08	9,78	2,7	9,1	13,5	8,55	20		
34	IgM (імуноглобулін М)	5,9	48,5	4,43	18,32	25,62	2,95	12,21	17,08	3	11,9	16,8	8,85	26		
35	IgG (імуноглобулін G)	3,7	17,1	2,78	6,56	11,14	1,85	4,37	7,43	2,3	4,3	8	5,55	18		
36	Apo A1 (аполіпопротеїн А1)	5,4	11,2	4,05	4,66	11,35	2,70	3,11	7,56	3,3	3,7	9,1	8,1			10
37	Apo B (аполіпопротеїн В)	7,4	20,2	5,55	8,07	17,22	3,70	5,38	11,48	3,5	6	11,6	11,1			10
38	HbA1c (глікований гемоглобін, NGSP)	1,2	4,8	0,90	1,86	3,34	0,60	1,24	2,23	0,9	1,5	3	1,8	18		
39	D-D (D-димер)	25,2	35,4	18,90	16,30	47,48	12,60	10,86	31,65	11,7	8,8	28,4	37,8			
40	HCY (гомоцистеїн)			0,00	0,00	0,00	0,00	0,00	0,00	4,15	8,63	15,48	0			
41	a2MG (a2-макро-глобулін)	3,1	16,6	2,33	6,33	10,17	1,55	4,22	6,78	1,7	4,8	7,56	4,65			

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM			Бажана мета за EFLM			Бажана мета за Ricos, 2014			MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCRA TEa(%)
		CWi	CWg	I (%)	B (%)	TEa(%)	I (%)	B (%)	TEa(%)	I (%)	B (%)	TEa(%)				
42	TSH (ТТГ)	17,7	35,9	13,28	15,01	36,91	8,85	10,01	24,61	9,7	7,8	23,7	26,55	24		15
43	T4 (загальний тироксин)	6,4	11,8	4,80	5,03	12,95	3,20	3,36	8,64	2,5	3	7	9,6	24	20	15
44	FT4 (вільний тироксин)	4,9	7,8	3,68	3,45	9,52	2,45	2,30	6,35	2,9	3,3	8	7,35	24		15
45	T3 (трийодтиронін)	9,4	12,2	7,05	5,78	17,41	4,70	3,85	11,61	3,5	3,5	9,2	14,1	24		15
46	FT3 (вільний трийодтиронін)	5	8,3	3,75	3,63	9,82	2,50	2,42	6,55	4	4,8	11,3	7,5	24		20
47	Tg (тиреоглобулін)	10,6	77,2	7,95	29,22	42,34	5,30	19,48	28,23	7	10,4	21,9	15,9			
48	PRL (пролактин)	29,5	43	22,13	19,55	56,06	14,75	13,04	37,37	11,5	10,5	29,4	44,25			20
49	FSH (фолікуло-стимулюючий гормон)	12,4	42,1	9,30	16,46	31,80	6,20	10,97	21,20	5,5	12,1	21,2	18,6			20
50	LH (лютеїнізуючий гормон)	22,8	30,8	17,10	14,37	42,59	11,40	9,58	28,39	11,5	8,9	27,9	34,2			20
51	TESTO (тестостерон)	12,5	21,4	9,38	9,29	24,76	6,25	6,20	16,51	4,6	6	13,6	18,75	35		20
52	DHEAS (дегідроепіандростерон)	6	21	4,50	8,19	15,62	3,00	5,46	10,41	3,2	7,8	13,1	9			10

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCRA TEa(%)	
		CW; CWg	I(%)	B(%)	TEa(%)	I(%)	B(%)	TEa(%)	I(%)					B(%)
53	SHBG (глобулін, що зв'язує статеві гормони)	9,7	36,2	7,28	14,05	26,06	4,85	9,37	17,37	6,5	9,7	20,4	14,55	10
54	E2 (Естрадіол)	15	13	11,25	7,44	26,01	7,50	4,96	17,34	11,3	8,3	26,9	22,5	20
55	PROG (прогестерон)	18,5	39,7	13,88	16,42	39,32	9,25	10,95	26,21				27,75	25
56	17PRG	28,3	38,6	21,23	17,95	52,97	14,15	11,97	35,31	9,8	13,5	29,7	42,45	20
57	CORT (кортизол)	16,3	48,7	12,23	19,26	39,43	8,15	12,84	26,29	7,6	10,3	22,8	24,45	15
58	ALD (альдостерон)	36,6	34,7	27,45	18,91	64,21	18,30	12,61	42,80	14,7	12,4	36,7	54,9	15
59	AFP (АФП)	4,6	55,2	3,45	20,77	26,46	2,30	13,85	17,64	6,1	11,8	21,9	6,9	20
60	CA125	8,7	25,5	6,53	10,10	20,87	4,35	6,74	13,91	12,4	15	35,4	13,05	20
61	CA15-3	4,4	36,8	3,30	13,90	19,34	2,20	9,27	12,90	3,1	15,8	20,8	6,6	15
62	CA19-9	4,3	57,4	3,23	21,59	26,91	2,15	14,39	17,94	8	32,9	46	6,45	15
63	SEA (раково-ембріональний антиген)	6,7	59,3	5,03	22,38	30,67	3,35	14,92	20,45	6,4	14,3	24,7	10,05	20

Продовження додатку 8.1.

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCRA TEa(%)	
		CW; CWg	I(%)	B(%)	TEa(%)	I(%)	B(%)	TEa(%)	I(%)					B(%)
64	PSA (простат-специфічний антиген загальний)	6,8	42	5,10	15,96	24,37	3,40	10,64	16,25	9,1	18,7	33,6	10,2	25
65	FPSA (простат-специфічний антиген вільний)	7,1	46,2	5,33	17,53	26,31	3,55	11,69	17,54				10,65	15
66	NSE (нейрон специфічна енолаза)	10,9	16,6	8,18	7,45	20,94	5,45	4,96	13,96				16,35	10
67	FOL (фолати)	11,8	37,9	8,85	14,89	29,49	5,90	9,92	19,66	12	19,2	39	17,7	30
68	FER (ферритин)	12,8	16,46	9,60	7,82	23,66	6,40	5,21	15,77	7,1	5,2	16,9	19,2	25
69	B12	7,2	36,1	5,40	13,80	22,71	3,60	9,20	15,14				10,8	20
70	INS (інсулін)	25,4	33,5	19,05	15,77	47,20	12,70	10,51	31,47	10,6	15,5	32,9	38,1	25
71	C-PEP (C-пептид)	23,8	33,1	17,85	15,29	44,74	11,90	10,19	29,83	8,3	7,1	20,8	35,7	10
72	IGF-1 (інсуліно-подібний фактор росту)	9,4	27	7,05	10,72	22,35	4,70	7,15	14,90	7,3	11,9	24	14,1	15
73	b2MG (β-2 мікроглобулін)	4,1	11,5	3,08	4,58	9,65	2,05	3,05	6,43	3	4,1	9	6,15	10

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCPA TEa(%)	
		CWi	CWg	I (%)	B(%)	TEa(%)	I(%)	B(%)	TEa(%)					I(%)
74	PTH (паратгормон)	15,7	23,5	11,78	10,60	30,03	7,85	7,07	20,02	13	8,8	30,2	23,55	25
75	OST (остеокалцин)	8,9	32,3	6,68	12,56	23,58	4,45	8,38	15,72	3,2	7,9	13,1	13,35	
76	25-OH VitD3 (холкальці-ферол)	6,9	26,1	5,18	10,12	18,66	3,45	6,75	12,44				10,35	15
77	RBC (еритроцити)	2,6	6,5	1,95	2,63	5,84	1,30	1,75	3,90	1,6	1,7	4,4	3,9	6
78	Hb (гемоглобін)	2,7	5,9	2,03	2,43	5,77	1,35	1,62	3,85	1,43	1,84	4,19	4,05	7
79	HCT (гематокрит)	2,8	5,5	2,10	2,31	5,78	1,40	1,54	3,85	1,35	1,74	3,97	4,2	6
80	MCV (середній об'єм еритроцитів)	0,8	3,7	0,60	1,42	2,41	0,40	0,95	1,61	0,7	1,26	2,42	1,2	
81	MCH (середній вміст гемоглобіну в еритроциті)	0,8	4,4	0,60	1,68	2,67	0,40	1,12	1,78	0,7	1,35	2,5	1,2	
82	MCHC (середня концентрація гемоглобіну в еритроцитах)	1	1,4	0,75	0,65	1,88	0,50	0,43	1,26	0,53	0,4	1,27	1,5	

№	Аналіт	Biovariation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak TEa(%)	Вимоги CLIA TEa(%)	Вимоги RCPA TEa(%)	
		CWi	CWg	I (%)	B(%)	TEa(%)	I(%)	B(%)	TEa(%)					I(%)
83	RDW CV (відносна ширина розподілу еритроцитів за об'ємом)	1,7	4,3	1,28	1,73	3,84	0,85	1,16	2,56	1,8	1,7	4,6	2,55	
84	RDW SD (відносна ширина розподілу еритроцитів за об'ємом)	1,7	4,3	1,28	1,73	3,84	0,85	1,16	2,56	1,8	1,7	4,6	2,55	
85	WBC (лейкоцити)	10,8	16,4	8,10	7,36	20,73	5,40	4,91	13,82	5,73	6,05	15,49	16,2	15
86	LYM (лімфоцити)	10,8	22,7	8,10	9,43	22,79	5,40	6,28	15,19	5,1	9,19	17,6	16,2	
87	MON (моноцити)	13,3	22,2	9,98	9,70	26,16	6,65	6,47	17,44	8,9	13,2	27,9	19,95	
88	NEUT (нейтрофіли)	14	23,6	10,50	10,29	27,62	7,00	6,86	18,41	8,55	9,25	23,35	21	
89	EOS (еозинофіли)	15	65,4	11,25	25,16	43,72	7,50	16,77	29,15	10,5	19,8	37,1	22,5	
90	BASO (базофіли)	12,4	26,3	9,30	10,90	26,25	6,20	7,27	17,50	14	15,4	38,5	18,6	

№	Аналіт	Bioevaluation EFLM database		Мінімальна мета EFLM		Бажана мета за EFLM		Бажана мета за Ricos, 2014		MAU	Вимоги Rilibak	Вимоги CLIA	Вимоги RCRA				
		CW <sub>i</sub>	CW <sub>g</sub>	I (%)	V (%)	TEa (%)	I (%)	V (%)	TEa (%)					I (%)	V (%)	TEa (%)	TEa (%)
91	PLT (тромбоцити)	7,6	18,7	5,70	7,57	16,97	3,80	5,05	11,32	4,6	5,9	13,4	11,4	13/15/18	25		
92	MPV (середній об'єм тромбоцитів)	2,3	7,1	1,73	2,80	5,64	1,15	1,87	3,76	2,15	2,29	5,84	3,45				
93	Reticul (ретикулоцити)	9,7	27,1	7,28	10,79	22,80	4,85	7,20	15,20	5	15,7	24	14,55				
94	APTT (активованій частковий тромбoplastиний час)	2,8	7,2	2,10	2,90	6,36	1,40	1,93	4,24	1,4	2,3	4,5	4,2	18	15		
95	PT (протромбінв час)	2,6	5,1	1,95	2,15	5,36	1,30	1,43	3,58	2	2	5,3	3,9	23	15		
96	FIB (фібриноген)	10,2	17,1	7,65	7,47	20,09	5,10	4,98	13,39	5,4	4,8	13,6	15,3		20		
97	Prot S (протеїн S)	7,3	20,3	5,48	8,09	17,12	3,65	5,39	11,42	2,9	15,9	20,7	10,95				

## ДОДАТОК 8.2. ЗРАЗКИ ЗВІТІВ З ВЕРИФІКАЦІЇ ТА ВАЛІДАЦІЇ МЕТОДИКИ ДОСЛІДЖЕННЯ

	Лабораторія	Індекс форми в СУЯ
1	Вимірювана величина	
2	Біоматеріал	
3	Методика вимірювання	
4	Обладнання/Аналізатор	
5	Програма верифікації (протокол верифікації, тип зразків)	
6	Оцінка невизначеності	
7	Дати верифікаційного експерименту	
8	Метрологічна прослежуваність щодо методу дослідження та використаних контрольних зразків	

## 9. Результати верифікаційних експериментів:

## 9.1. Верифікація прецизійності та правильності (оцінка зсуву) за CLSI EP15-A3 та попередня оцінка невизначеності ISO/TS 20914

Матеріал для дослідження: контрольні зразки (КС) \_\_\_\_\_ пул зразків пацієнтів  
 Невизначеність (U,%) приписаних значень контрольного матеріалу \_\_\_\_\_

Функціональні характеристики методики	Вимоги до якості	Специфікації виробника (верифікаційний ліміт)						Результати верифікаційного експерименту						
		Рівень 1		Рівень 2		KC I		KC II						
		CI	$S_{Ri},\% / VL(S_{Ri})$	СII	$S_{Ri},\% / VL(S_{Ri})$	CI	$S_{Ri},\% / S_{wL},\% / VL(S_{wL})$	СII	$S_{Ri},\% / S_{wL},\% / VL(S_{wL})$	CI	$S_{Ri},\% / S_{wL},\% / VL(S_{wL})$	СII	$S_{Ri},\% / S_{wL},\% / VL(S_{wL})$	
Прецизійність														
Правильність (зсув)		-		-										
Оцінка відносної розширеної невизначеності U,%		-		-										

CI – концентрація аналіта в 1 рівні контрольного зразку, CII – концентрація аналіта в 1 рівні контрольного зразку, СII – концентрація аналіта в 1 рівні контрольного зразку, СII  
 $S_{Ri},\%$  – внутрішньосерійна прецизійність та її верифікаційний ліміт (VL)  
 $S_{wL},\%$  – внутрішньолабораторна прецизійність та її верифікаційний ліміт (VL)

### 9.2. Додаткові функціональні характеристики методики дослідження (за необхідністю)

За необхідністю приводяться дані щодо верифікації інших функціональних характеристик методики – референтних значень, діапазону вимірювання, діагностичних чутливості та специфічності та ін.

### 10. Висновок:

Висновок щодо верифікації функціональних характеристик методу та дозвіл щодо його використання в лабораторії

Верифікацію проведено: \_\_\_\_\_

(Відповідальна особа)\_\_\_\_\_

Звіт затвердив: \_\_\_\_\_

Зав. лабораторією\_\_\_\_\_

	Лабораторія	Індекс форми в СУЯ
	Звіт про валідацію методики дослідження	стор. _ 3 _
Процедура вимірювання:	Навести позначення та назву документованої процедури, короткий опис процедури	
Метод дослідження	Навести узагальнений метод дослідження (імуноферментний аналіз, наприклад)	
Призначення тесту (мета дослідження)	Навести клінічне використання результатів дослідження, яке планується	
Посилання	(Навести перелік нормативних посилань та літературних джерел, що були використані для валідації методики дослідження 1. ...	

## Результати валідації методики дослідження

Аналітичні характеристики методики дослідження	Процедура валідації	Отримані результати	Критерії прийнятності та їх виконання
Діапазон вимірювання	EP06		
Ліміт детекції	EP17		
Ліміт кількісного визначення	EP17		
Повторюваність	Реплікаційний експеримент EP05		
Проміжна прецизійність	Реплікаційний експеримент EP05		
Оцінка правильності	Порівняння методів EP09		
Вплив інтерференції	EP07		
Референтні інтервали	EP28		
...			

## Висновок щодо придатності методики дослідження

Відповідальна особа	Висновок щодо придатності для використання за призначенням	Дата	Завідуючий лабораторією	Висновок щодо придатності для використання в клінічній практиці	Дата

## 9. БІБЛІОГРАФІЯ

1. A practical guide to validation and verification of analytical methods in the clinical Laboratory. Joachim Pum, 2019 Advances in Clinical Chemistry <https://doi.org/10.1016/bs.acc.2019.01.006>
2. CLSI EP26 User Evaluation of Between-Reagent Lot Variation, 1st Edition
3. CLSI C24 Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions, 4th Edition
4. CLSI EP06 Evaluation of Linearity of Quantitative Measurement Procedures, 2nd Edition
5. CLSI EP07 Interference Testing in Clinical Chemistry, 3rd Edition
6. CLSI EP09 Measurement Procedure Comparison and Bias Estimation Using Patient Samples, 3rd Edition
7. CLSI EP12 User Protocol for Evaluation of Qualitative Test Performance
8. CLSI EP15 User Verification of Precision and Estimation of Bias, 3rd Edition
9. CLSI EP19 A Framework for Using CLSI Documents to Evaluate Clinical Laboratory Measurement Procedures, 2nd Edition
10. CLSI EP28 Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory, 3rd Edition
11. CLSI EP29 Expression of Measurement Uncertainty in Laboratory Medicine, 1st Edition
12. EFLM Recommendation for the review of biological reference intervals in medical laboratories DOI 10.1515/cclm-2016-0793
13. EURACHEM/CITAC Guide QUAM:2012.P1 Quantifying Uncertainty in Analytical Measurement Third Edition Editors S L R Ellison (LGC, UK) A Williams (UK), EURACHEM, 2012
14. ISO/TS 20194:2019 Medical laboratories – Practical guidance for the estimation of measurement uncertainty
15. James O. Westgard, Sten A. Westgard, Basic Quality Management Systems, 2014
16. Tietz textbook of clinical chemistry and molecular diagnostics / [edited by] Nader Rifai; senior editors, Andrea R. Horvath, Carl T. Wittwer, sixth edition, Elsevier 2018.
17. Validation and verification of examination procedures in medical laboratories: opinion of the EFLM Working Group Accreditation and ISO/CEN standards (WG-A/ISO) on dealing with ISO 15189:2012 demands for method verification and validation <https://doi.org/10.1515/cclm-2019-1053>

18. ДСТУ EN ISO 15189:2015 Лабораторії медичні. вимоги до якості та компетентності (EN ISO 15189 :2012, IDT) Київ, ДП «Укр-НДНЦ», 2015.
19. Настанова Eurachem «Придатність аналітичних методів для конкретного застосування. Настанова для лабораторій з валідації методів та суміжних питань»: за ред. Б. Магнуссона та У. Ернемарка: переклад другого видання 2014 р. – К.: ТОВ «Юрка Любченка», 2016. – 92 с.
20. Термінологія аналітичного вимірювання. Вступ до VIM 3: за ред. В. Барвік та Е. Прічард: переклад першого видання настанови Eurachem 2011 р. – К.: ТОВ «Юрка Любченка», 2015. – 82 с.
21. ISO 15189:2012 implementation: an applied guide for medical laboratories Dennis Mok, Elizabeth Lim, Kathryn Eckersley, Lidia Hristov and Christina Kirsch Australian Journal of Medical Science November 2013 Vol. 34 No. 4
22. А.Котуза Н.Вінарська Загальні вимоги підготовки медичної лабораторії до акредитації відповідно до вимог ДСТУ 18015189:2015 «Практика управління медичним закладом. Спецвипуск» I півріччя 2017 року (12) ТОВ «Пресс-Альянс»
23. Andrew C. Don-Wauchope. Lot change for reagents and calibrators. Clinical Biochemistry Volume 49, Issues 16–17, November 2016, Pages 1211-1212
24. Thompson, Simon, and Douglas Chesher. «Lot-to-Lot Variation.» The Clinical biochemist. Reviews vol. 39,2 (2018): 51-60.
25. Simundic AM, Kackov S, Miler M, Fraser CG, Petersen PH. Terms and symbols used in studies on biological variation: the need for harmonization. Clin Chem 2015;61:438–9. <https://doi.org/10.1373/clinchem.2014.233791>
26. Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. Crit Rev Clin Lab Sci 1989; 27:409 – 37. <https://doi.org/10.3109/10408368909106595>
27. Sandberg S, Carobene A, Bartlett B, Coskun A, Fernandez-Calle P, Jonker N, Díaz-Garzón J, Aarsand Aasne K. Biological variation: recent development and future challenges. Clin Chem Lab Med 2023;61:741–750. <https://doi.org/10.1515/cclm-2022-1255>
28. Carobene, A, Campagner, A, Uccheddu, C, Banfi, G, Vidali, M, Cabitza, F. The multicenter European Biological Variation Study (EuBIVAS): a new glance provided by the Principal Component Analysis (PCA), a machine learning unsupervised algorithms, based on the basic metabolic panel linked measurands. Clin Chem Lab Med 2022;60:556–68. <https://doi.org/10.1515/cclm-2021-0599>

29. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) // JAMA. – 2001. – 285. – P.2486-2497
30. ISO 15189:2022(en) Medical laboratories – Requirements for quality and competence