

Оценка эффективности измерения данных

Для оценки эффективности методов измерения медико-биологических данных применяют ряд критериев, главными из которых являются:

Точность измерений – это соответствие результатов измерения истинному значению определяемой величины. Высокая точность измерения достигается при минимальных рандомизированных и систематических погрешностях.

Правильность измерений. Это качество измерения характеризует величину систематических погрешностей. Чем они меньше, тем более правильным оказываются измерения.


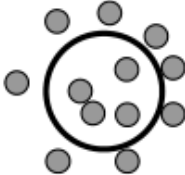
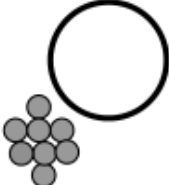
Показатели измерения	Нет ошибок	Рандомизированные ошибки	Систематические ошибки
«Охота за мишенью»			
Точность (достоверность)	Очень хорошая	Удовлетворительная	Плохая
Сходимость	Хорошая	Плохая	Очень хорошая

Рис.1.1. «Охота за мишенью». Связь между достоверностью (точностью) и сходимостью результатов, рандомизированными и систематическими ошибками исследований

Сходимость измерений. Данное качество измерения характеризует величину случайных ошибок. Чем они меньше, тем лучше сходимость измерения. Этот критерий показывает, насколько близки друг к другу измерения, выполненные в одинаковых условиях, т.е. в одной и той же лаборатории и на одном и том же приборе.

Воспроизводимость измерений. Этот критерий показывает, как близки между собою будут результаты измерений, выполненных в различных условиях, т.е. в различных лабораториях, на различных аппаратах и различными людьми.

На рис.1.1 представлено взаимоотношение систематических и рандомизированных ошибок, точности и сходимости измерений, условно обозначенное нами как «охота за мишенью». Как мы видим, *точность* измерений включает в себя в качестве обязательного критерия их *сходимость*. В то же время сходимые измерения из-за систематической ошибки могут оказаться *неточными*. Данное обстоятельство нужно обязательно учитывать при анализе всех медико-биологических данных.

Величину расхождения между измерениями, выполненными в одних и тех же условиях, обозначают как среднее квадратическое, или *стандартное, отклонение* (σ). Оно определяется по формуле:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}},$$

где $\sum (x_1 - \bar{x})^2$ – сумма квадратов разностей каждого из измерения со средними величинами χ и $\bar{\chi}$, n – количество измерений.

Для обозначения средней величины серии измерений (одного или нескольких объектов) пользуются показателем $\bar{x} \pm \sigma$. В таких границах («*значимых*»), то есть от $\bar{x} - \sigma$ до $\bar{x} + \sigma$, будут располагаться 68,3% от всех проведенных измерений. Существуют более широкие границы значений – «*зона предостережения*» ($\bar{x} \pm 2\sigma$), которая включает в себя 95,5% всех измерений, и «*контрольные границы измерений*» ($\bar{x} \pm 3\sigma$), в пределах которых находятся 99,7% измерений.

Для того чтобы оценить вариабельность измерений внутри одной серии, то есть оценить сходимость результатов, прибегают к определению *коэффициента вариации* (КВ) по следующей формуле:

$$KB \% = \frac{\sigma \cdot 100}{x},$$

где σ – стандартное отклонение, x – значение измерения.

Чем меньше коэффициент вариации, тем выше качество измерений. В медико-биологических исследованиях обычно допустимы измерения, у которых коэффициент вариации не превышает 10%.

Для того чтобы охарактеризовать точность, с которой определена какая-либо средняя величина, используют показатель «стандартная погрешность», или «стандартная ошибка» (m). Он составляет

$$m = \frac{\sigma}{\sqrt{n}},$$

где σ – стандартное отклонение, n – количество измерений.

Стандартная погрешность позволяет, кроме того, сопоставить результаты измерений, выполненных в различных лечебных учреждениях или произведенных различными методами исследования. Для этого нужно определить достоверность различия двух средних арифметических \bar{x}_A и \bar{x}_B . Выполняют это с помощью t – критерия Стьюдента. Критерий Стьюдента определяется по формуле

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{m_A^2 + m_B^2}},$$

где \bar{x} – среднее значение измерений, m – стандартная ошибка в сериях измерений A и B . Чем больше критерий Стьюдента, тем достовернее различие между изучаемыми сериями измерений.

В табл. 1.1 приведены максимально допустимые значения критерия Стьюдента для пятипроцентного риска ошибочных заключений, что вполне приемлемо в практической медицине.

Таблица 1.1

**Значения критерия Стьюдента
для 5% вероятности ошибки ($p < 0,05$)**

Число измерений	Значение t-критерия
2	12,71
3	4,30
4	3,18
5	2,78
6	2,57
7	2,45
8	2,37
9	2,31
10	2,26
11	2,23
12	2,20
13	2,18
14	2,16
15	2,15
...	...
30	2,04
100	1,98
1000	1,96

Значимость одного измерения, например, пульсового давления или радиоактивности пробы, определяют с помощью *относительной стандартной погрешности* (z) по формуле

$$z = \frac{100 \cdot \sqrt{N}}{N}$$

где N – количество отсчетов.

Из сказанного ясно, что чем большее количество отсчетов выполняется во время исследования, тем меньше стандартная относительная погрешность (говорят, что «лучше статистика» счета). Увеличение количества отсчетов, в частности увеличение числа зарегистрированных сигналов (квантов, частиц, эхо-сигналов и т. п.) на единицу изображения на мониторе или в единицу времени сопровождается уменьшением погрешности регистрации, что приводит к улучшению качества получаемых данных, будь то числовая характеристика процесса или изображение органа. Поэтому повысить качество любой диагностической системы можно двумя путями: либо увеличить интенсивность потока энергии, которая создает диагностические данные (например, увеличить количество вводимого в организм пациента с диагностической целью радиоактивного вещества), либо повысить чувствительность регистрирующих систем.

Для того чтобы оценить сходимость результатов измерений одним и тем же методом, но выполненных в двух различных лечебных учреждениях, или воспроизводимость результатов измерений одного и того же показателя, выполненных в одной лаборатории, но различными методами исследования, определяют *F-критерий*:

$$F = \frac{\sigma_A^2}{\sigma_B^2},$$

где σ – стандартное отклонение измерений, выполненных методами А и В.

Минимальная величина F-критерия, которая допустима в медицинской практике, показана в табл. 1.2.

Таблица 1.2

Минимально допустимая (критическая) величина критерия F

Число измерений в серии В	Число измерений в серии А				
	2	3	4	5	6
2	4052	4999	5403	5625	5674
3	98,49	99,01	99,17	99,25	99,30
4	34,12	30,82	29,46	28,71	28,24
5	21,20	18,00	16,69	15,98	15,52
6	16,26	13,27	12,06	11,39	10,97