

# **К ВОПРОСУ О СОГЛАСОВАННОСТИ РЕЗУЛЬТАТОВ ПРИМЕНЕНИЯ ТЕХНИЧЕСКИХ СРЕДСТВ ДИАГНОСТИКИ В МЕДИЦИНЕ**

**А. А. Вальтер, А. В. Харламов**

*Саратовский национальный исследовательский  
государственный университет им. Н.Г. Чернышевского, Россия*  
E-mail: n.valter03@yandex.ru, harlamovav@info.sgu.ru

В статье дается сравнительный анализ показателей воспроизводимости результатов диагностики, проводимой с помощью технических средств на примере анализа значений тазового наклона, полученных прибором РадиАнт.

## **TO THE QUESTION OF THE CONSISTENCY OF THE RESULTS OF THE USE OF DIAGNOSTIC TOOLS IN MEDICINE**

**A. Valter, A. Kharlamov**

The article provides a comparative analysis of the reproducibility of diagnostic results carried out using technical means using the example of the analysis of pelvic incidence values obtained by the RadiAnt device.

Развитие медицинских технологий предполагает постоянное увеличение применения технических средств лечения и диагностики. Помимо прочего диагностическое оборудование должно обладать воспроизводимостью и согласованностью получаемых результатов. Воспроизжимость предполагает стабильность (повторяемость) результатов измерений, осуществляемых одним медицинским работником в разное время, например при анализе рентгеновских снимков, а согласованность предполагает одинаковость результатов измерений, выполняемых разными медицинскими работниками. В первом случае говорят о внутриэкспертной согласованности, а во втором о межэкспертной согласованности. Особенно актуально это становится при использовании в диагностике мобильных приложений [1].

Учет согласованности прогностических оценок осложняется еще и тем, что в сфере здравоохранения в общем и в медицине, в частности, отсутствуют единые признанные критерии, стандарты или рекомендации по обеспечению надежности и согласованности результатов.

В зависимости от шкалы представления показателей воспроизжимость (согласованность) измерений может быть оценена с помощью ряда статистических показателей, например, Каппа Коэна [5, 7], Каппа Флейсса [5, 6, 7, 8], коэффициент внутрикласовой корреляции ИСС [3], а также графиков Бланда-Альтмана [4]. По мнению авторов той же цели может служить t-критерий Стьюдента для парных измерений.

Формула для вычисления Каппы Коэна выглядит следующим образом:

$$k = 1 - \frac{1-p_0}{1-p_e}$$

$$p_0 = \frac{\text{число соглашений}}{\text{общее количество результатов}}$$

$$p_e = \frac{1}{N^2} \sum_k n_{k1} n_{k2},$$

где  $p_0$  – относительная частота совпадений между оценщиками,  $p_e$  – гипотетическая вероятность случайных совпадений,  $k$  – число категорий,  $N$  – число результатов,  $n_{ki}$  – количество соглашений между первым и вторым оценщиком в категории  $k$ .

Каппа Флэйсса вычисляется как отношение наблюдаемой степени согласия к ожидаемой степени согласия, учитывая вероятность случайных совпадений:

$$k = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e},$$

где  $\bar{P}$  – наблюдаемая доля совпадений,  $\bar{P}_e$  – ожидаемая доля совпадений. Для того, чтобы рассчитать  $\bar{P}$  необходимо знать  $P_i$ , а для  $\bar{P}_e$  –  $P_j$ :

$$P_j = \frac{1}{Nn} \sum_{i=1}^N n_{ij}$$

$$P_i = \frac{1}{n(n-1)} \sum_{j=1}^k n_{ij} (n_{ij} - 1)$$

$$\bar{P} = \frac{1}{N} \sum_{i=1}^N P_i$$

$$\bar{P}_e = \sum_{j=1}^k p_j^2,$$

где  $N$  – общее количество объектов,  $n$  – количество оценок по предмету и  $k$  – количество категорий.

Коэффициент внутриклассовой корреляции ИСС представлен следующей формулой:

$$r = \frac{1}{Ns^2} \sum_{n=1}^N (x_{n,1} - \bar{x})(x_{n,2} - \bar{x}),$$

где  $N$  – число парных значений,  $\bar{x}$  – выборочное среднее,  $s^2$  – дисперсия.

Сравнительный анализ характеристик согласованности был выполнен на эмпирических данных, предоставленных Центральным научно-исследовательским институтом Травматологии и Ортопедии имени Н.Н. Приорова (ЦИТО). Исходный массив данных представлял в обезличенном виде результаты измерений тазового наклона (PI), проведенные экспертами с помощью апробированного инструмента РадиАнт. Два эксперта провели по два измерения у 30 пациентов. Поочередно проводились измерения первым и вторым экспертами сначала одномоментно по времени, и повторно спустя две недели [2].

Показатель PI является количественным (представлен в абсолютной шкале), но для расчета статистик капп был переведен в категориальные шкалы.

Первая – это шкала Руссоли (три категории, равные интервалы, величина интервала 10) и вторая шкала (семь категорий, равные интервалы, величина интервала 5). Результаты вычислений представлены в табл. 1.

Таблица 1

**Результаты вычислений Каппа Коэна и Флейсса**

	Каппа Коэна		Каппа Флейсса	
	1 шкала	2 шкала	1 шкала	2 шкала
Один эксперт, два измерения	0,64	0,41	0,69	0,40
Два эксперта, одно измерение	0,69	0,41	0,69	0,40

Для негруппированных данных вычислялся коэффициент внутриклассовой корреляции ICC (Intraclass correlation). Соответствующие значения даны в табл. 2. В четвертой строке даны значения парного коэффициента корреляции Пирсона, при этом значения коэффициентов различаются только в третьем знаке. В последней строке таблицы представлены *p*-значения соответствующих значений статистик *t*-критерия Стьюдента для парных наблюдений. Проверка гипотезы об однородности результатов расчетов показала их совпадения, только у второго эксперта можно отметить некоторое смещение при втором измерении.

Таблица 2

**Результаты расчета показателей воспроизводимости**

	Внутриэкспертная воспроизводимость		Межэкспертная воспроизводимость	
	Первый эксперт	Второй эксперт	Первое измерение	Второе измерение
<b>ICC</b>	0,910	0,953	0,853	0,801
<b>r</b>	0,911	0,957	0,858	0,816
<b>p</b>	0,29	0,03	0,23	0,49

Графики Бланда-Альтмана для всех случаев показывают отсутствие тренда, и практически все измерения находятся в пределах референтного интервала. Аналогичная картина наблюдается для диаграмм рассеяния. На рисунках приведены примеры графика и диаграммы для первого измерения между экспертами (рис.1 и рис. 2)

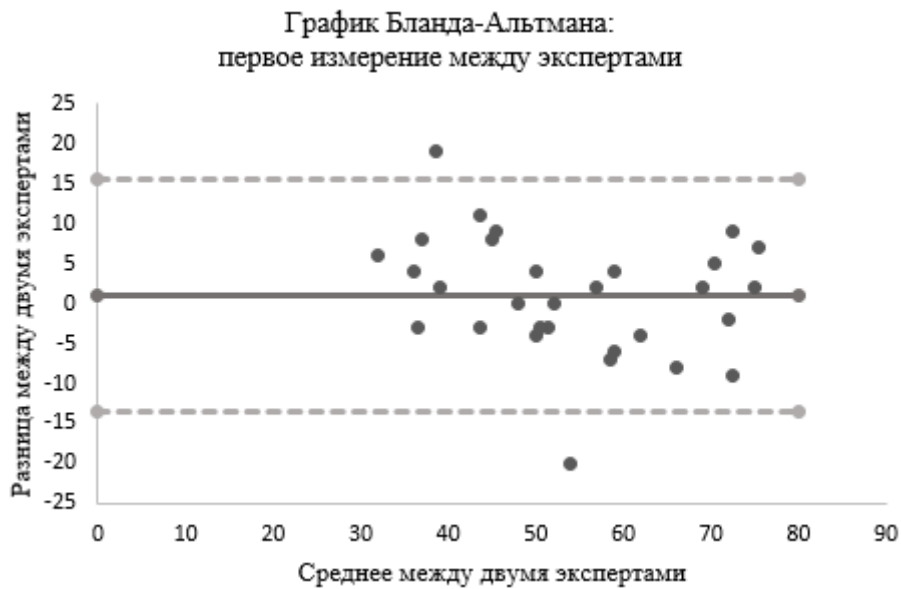


Рис. 1. График Бланда-Альтмана



Рис. 2. Диаграмма рассеяния

Таким образом по результатам расчетов можно сделать следующие выводы:

оказатели воспроизводимости совпадают для показателей капп и ICC, что свидетельствует об их согласованности.

применение разных статистик дает несопоставимые результаты, при этом значения ICC больше, чем значения Капп.

3. Шкалирование с разным числом категорий изменяет значения показателей, при этом включение/исключение границ не изменяет результат. классический статистический подход согласуется с упрощенными медицинскими практиками, при этом позволяет получать больше достоверной информации, по сравнению с отдельными индикаторами.

## СПИСОК ИСТОЧНИКОВ

1. *Kottner J., Audigeo L., Brorson S., Donner A., Gajewski B. J., Asbjørn Hróbjartsson, Robertsg C., Shoukri M., David L. Streiner* Guidelines for Reporting Reliability and Agreement Studies (GRRAS) // *Journal of Clinical Epidemiology*. 2011. № 64. С. 96-106.
2. *Ivanov D. V., Kirillova I. V., Kossovich L. Yu., Likhachev S. V., Polienko A. V., Kharlamov A. V., Shulga A. E.* Comparative analysis of the SpinoMeter mobile application and Surgimap system for measuring the sagittal balance parameters: inter-observer reliability test // *Genij Ortopedii*. 2021. № 1. P. 96-106.
3. *Koch G. G.* Intraclass correlation coefficient. In Samuel Kotz and Norman L. Johnson (ed.). // *Encyclopedia of Statistical Sciences*. New York: John Wiley & Sons. 1982. Vol. 4. P. 213-217.
4. *Bland J. M., Altman D. G.* Statistical methods for assessing agreement between two methods of clinical measurement // *Lancet (London, England)*. 1986. Vol. 1. P. 307-310.
5. *Gwet K. L.* Computing inter-rater reliability and its variance in the presence of high agreement // *The British journal of mathematical and statistical psychology*. 2008. P. 29-48.
6. *Shrout P. E., Fleiss J. L.* Intraclass correlations: uses in assessing rater reliability // *Psychol Bull*. 1979 Vol. 86 (2). P. 420-428.
7. *McHugh M. L.* Interrater reliability: the kappa statistic // *Biochem Med (Zagreb)*. 2012. Vol. 22 (3). P. 76-82.
8. *Fleiss J. L.* Statistical methods for rates and proportions. Second Edition / New York: John Wiley. 1981. 336 p.