

# **Оценка диагностической информативности количественных признаков в биомедицинских исследованиях на основании описательных статистик и стандартизированной разности средних значений**

**Глазков А.А.<sup>\*1</sup>, Куликов Д.А.<sup>1,2</sup>, Глазкова П.А.<sup>1</sup>**

<sup>1</sup>ГБУЗ МО МОНИКИ им. М.Ф. Владимирского, г. Москва, Россия

<sup>2</sup>МГОУ, Московская область, г. Мытищи, Россия

**Аннотация.** Для анализа диагностических возможностей количественных признаков в биомедицинских исследованиях широко используется ROC-анализ. При этом анализе строится ROC-кривая, характеризующая частоты истинно положительных и ложно положительных результатов для различных градаций анализируемого признака. При этом в большом количестве публикаций, посвященных применению количественных методов диагностики, данный анализ не проводится, исследователи ограничиваются лишь анализом статистической значимости различий между группами. В мета-анализах, обобщающих исследования количественных признаков при каком-либо явлении, оцениваемым параметром является величина эффекта выражаемая через стандартизованную разность средних. Этот параметр не говорит напрямую о диагностических возможностях признака. Это приводит к затруднениям в оценке применимости описываемых в статьях методик для диагностики различных заболеваний. В статье описан подход к проведению ROC-анализа на основании кумулятивных функций нормального распределения. Этот метод может быть использован для оценки диагностической информативности количественного признака на основании средних арифметических значений и стандартных отклонений, медиан и квартилей, стандартизированной разности средних. На примере оценки диагностической информативности количественных признаков, анализируемых в ходе изучения состояния микроциркуляции при различных заболеваниях, нами приведены примеры работы предлагаемого подхода на моделированных данных, на данных из литературных источников и на собственных наблюдениях. Представленные в статье результаты могут быть использованы специалистами медицинского и биологического профиля для оценки диагностической информативности различных количественных признаков без доступа к первичным данным.

**Ключевые слова:** стандартизованная разность средних, диагностическая информативность, ROC-анализ, медицинская статистика, нормальные распределения, моделирование.

## **ВВЕДЕНИЕ**

Поиск новых диагностических критериев является актуальной задачей современной медицины. В наиболее простом и часто встречающемся виде задача сводится к изучению возможности классификации наблюдения в одну из двух групп на основании

---

\*A.Glazkov@monikiweb.ru

количественного признака: разделению обследованных на «больных» и «здоровых» на основании какого-либо изученного параметра [1].

Для анализа диагностических возможностей количественных признаков наиболее удобным методом является ROC-анализ [2, 3]. При этом анализе строится кривая, характеризующая частоты истинно положительных и ложно положительных результатов для различных градаций количественных признаков, после чего рассчитывается площадь под ROC-кривой (Area Under the Curve, *AUC*). Значения *AUC* ниже 0.5 интерпретируют как отсутствие диагностической информативности, от 0.5 до 0.7 – как слабую диагностическую информативность, от 0.7 до 0.8 – как умеренную, от 0.8 до 0.9 – как хорошую, от 0.9 до 1 – как отличную [4]. Однако, такая оценка проводится далеко не во всех исследованиях. Достаточно часто авторы ограничиваются лишь анализом статистической значимости различий между группами [5–7]. Это не позволяет делать вывод о применимости описываемых методик в диагностических целях. При отсутствии первичных данных исследователь лишён возможности проведения ROC-анализа и, как следствие простой оценки диагностической информативности метода.

Вторая проблема, обуславливающая актуальность настоящей работы связана с тем, что в мета-анализах и систематических обзорах, которые согласно общепринятой классификации обладают наибольшей убедительностью доказательств в парадигме «доказательной медицины», оцениваемой метрикой является величина эффекта [8]. Для количественных признаков эта метрика представляет из себя стандартизованную разность средних значений [9]. Как будет показано далее, существующие способы интерпретации стандартизованной разности средних значений не позволяют делать заключений о диагностических возможностях оцениваемого параметра: большая величина эффекта не всегда свидетельствует о высокой диагностической информативности.

Таким образом, целью настоящей работы была разработка подхода, позволяющего оценить диагностическую информативность метода на основании описательных статистик и по величине эффекта в подобных исследованиях.

## ПРЕДЛАГАЕМЫЙ ПОДХОД

### Анализ диагностических характеристик на основании выборочных оценок

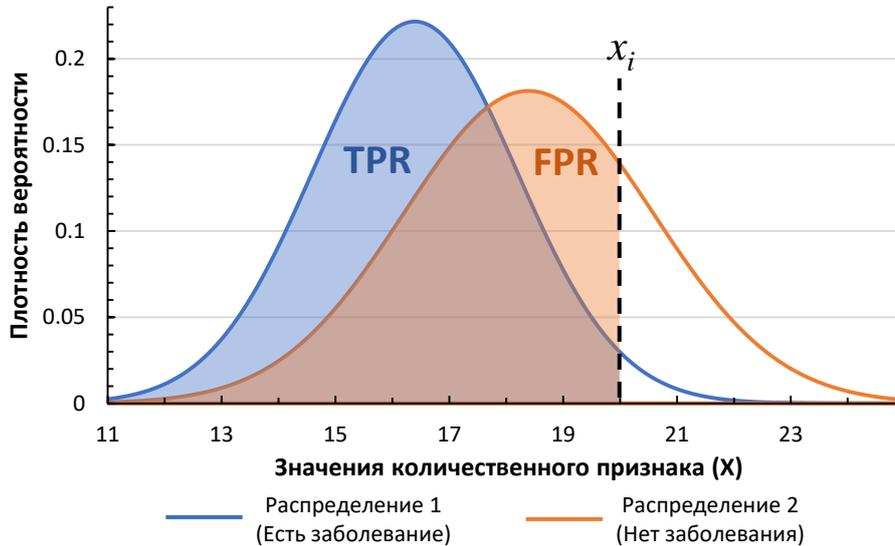
Стандартный способ проведения ROC-анализа предполагает расчёт долей истинно положительных *TPR* (True Positive Rate) и ложно положительных *FPR* (False Positive Rate) результатов классификации для различных пороговых градаций количественного признака, встречающихся в анализируемой выборке [10]. При этом если оцениваемый параметр  $x$  непрерывен и его распределение в двух группах известно, тогда *TPR* и *FPR* для каждого значения  $x$  могут быть оценены на основании кумулятивных функций распределений  $\Phi(x)$ . Пусть  $\Phi_1$  – кумулятивная функция распределения параметра у лиц с заболеванием,  $\mu_1$  – среднее арифметическое значение  $x$  в этой группе, а  $\Phi_2$  и  $\mu_2$  – кумулятивная функция распределения параметра и среднее арифметическое значение  $x$  у лиц без заболевания, тогда

$$\begin{cases} TPR(x) = \Phi_1(x), FPR(x) = \Phi_2(x), \text{ если } \mu_1 < \mu_2 \\ TPR(x) = 1 - \Phi_1(x), FPR(x) = 1 - \Phi_2(x), \text{ если } \mu_2 < \mu_1 \end{cases} \quad (1)$$

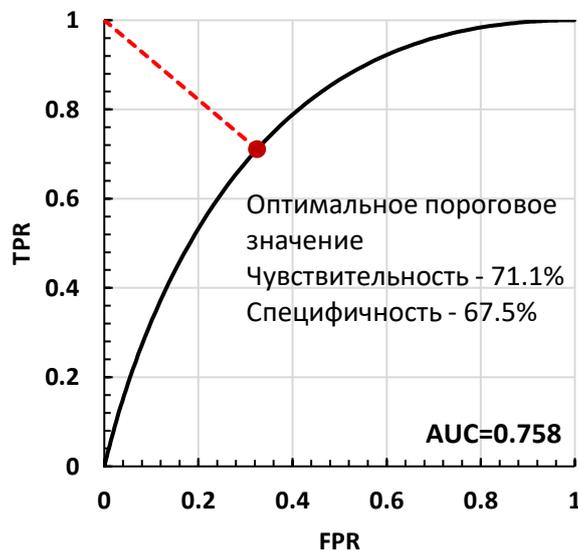
Известно, что кумулятивная функция нормального распределения, наиболее часто встречающегося в медицинских работах, зависит от среднего значения ( $\mu$ ) и среднеквадратического отклонения ( $\sigma$ ):

$$\Phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \quad (2)$$

Таким образом, зная выборочные оценки  $M_1$ ,  $SD_1$  и  $M_2$ ,  $SD_2$ , мы имеем возможность с найти пары значений  $TPR$  и  $FPR$  и построить ROC-кривую с заданным шагом  $dx$  в определённом диапазоне. На рисунке 1 на примере графиков плотности вероятности приведена иллюстрация такого анализа для выборочных оценок  $M_1 = 16$ ,  $SD_1 = 1.8$  (группа лиц с заболеванием),  $M_2 = 18$  и  $SD_2 = 2.2$  (группа лиц без заболевания).



**Рис. 1.** Пример анализа доли истинно положительных ( $TPR$ ) и ложноположительных ( $FPR$ ) результатов диагностики наличия заболевания у пациентов по количественному признаку  $X$ .



**Рис. 2.** ROC-кривая, построенная на основании предлагаемого подхода.

Несложной будет реализация алгоритма, рассчитывающего значения  $TPR$  и  $FPR$  с помощью (1) с заданным  $dx$ . На рисунке 2 показан пример ROC-кривой, построенной для приведённых выше выборочных оценок на основании предлагаемого подхода. Более того, с помощью данного подхода может быть определено и оптимальное пороговое значение с помощью анализа расстояния  $l$  до точки  $(0; 1)$  на графике ROC-кривой [11].

Для каждого  $x_i$   $l(x_i) = \sqrt{(1-TPR)^2 + FPR^2}$ , значение  $\{x: l(x) = \min_i l(x_i)\}$  будет являться оптимальным пороговым значением. Площадь под ROC-кривой также может быть рассчитана на основании полученных значений  $TPR$  и  $FPR$ , например, методом трапеций. Таким образом, оценка диагностических характеристик количественной переменной может быть проведена даже в ситуации, когда авторами оригинального исследования не был проведён ROC-анализ, а были только рассчитаны средние значения ( $M$ ) и стандартные отклонения ( $SD$ ). Ограничением данного подхода является тот факт, что не во всех статьях авторы, описывая свои результаты, приводят привычный набор из объема выборки, значений среднего и стандартного отклонения ( $n$ ,  $M$ ,  $SD$ ) в качестве описательных статистик количественных переменных. В ряде случаев вместо среднего и стандартного отклонения стало принято указывать медиану (она же второй квартиль) и первый и третий квартили  $m$ ,  $q_1$ ,  $q_3$ . В этих случаях для получения оценочных значений выборочного среднего  $M'$  и выборочного среднеквадратичного отклонения  $SD'$  может быть применён подход, описанный в статье X. Wan et al. [12]:

$$M' = \frac{q_1 + m + q_3}{3}, \quad SD' = \frac{q_3 - q_1}{2\Phi^{-1}\left(\frac{0.75n - 0.125}{n + 0.25}\right)}. \quad (3)$$

В формуле (3)  $\Phi^{-1}$  – это обратная кумулятивная функция нормального распределения, а  $n$  – объем выборки.

Кумулятивная функция распределения и её инвертированная форма реализованы в большинстве программ для статистического анализа данных (включая Excel), поэтому проведение расчётов по описанному выше алгоритму не должно представлять сложностей. Реализованный нами скрипт реализации такого алгоритма с помощью языка R V. 3.6.3 [13] представлен в Приложении (см. Дополнительные материалы к статье).

### Анализ диагностических характеристик на основании стандартизованной разности средних значений

Стандартизованная разность средних  $SMD$  (Standardized Mean Difference), или величина эффекта, используется в качестве сводной статистики в мета-анализе. Один из способов расчёта данной величины был предложен J. Cohen [9, 14]:

$$SMD = \frac{M_1 - M_2}{SD_{pooled}}; \quad (4)$$

$$SD_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}.$$

$M_1$ ,  $SD_1$  – это среднее значение и стандартное отклонение в группе численностью  $n_1$  человек, для которой изучается величина эффекта, а  $M_2$  и  $SD_2$  – соответствующие статистики в группе численностью  $n_2$  человек, выступающей в качестве контрольной. Предположим, что  $n_1 = n_2 = n$ . Выразим  $SD_1 = k \cdot SD_2$  ( $k$  – вещественное число). Тогда:

$$SD_{pooled} = \sqrt{\frac{(n-1) \cdot k \cdot SD_2^2 + (n-1)SD_2^2}{2(n-1)}} = SD_2 \sqrt{\frac{k+1}{2}}.$$

Ранее мы показали, что ROC-кривая для нормально распределённых количественных признаков может быть построена на основании кумулятивных функций распределения. То есть, мы можем оценить диагностическую информативность показателя  $SMD$ , приведя распределение в контрольной группе к стандартному нормальному

распределению ( $M_2 = 0$ ,  $SD_2 = 1$ ) и вычислив  $M_1$  и  $SD_1$  исходя из известных значений  $SMD$  и  $k$ . В таком случае  $SD_1 = k$ , а значение  $M_1$  может быть вычислено из  $SMD$ :

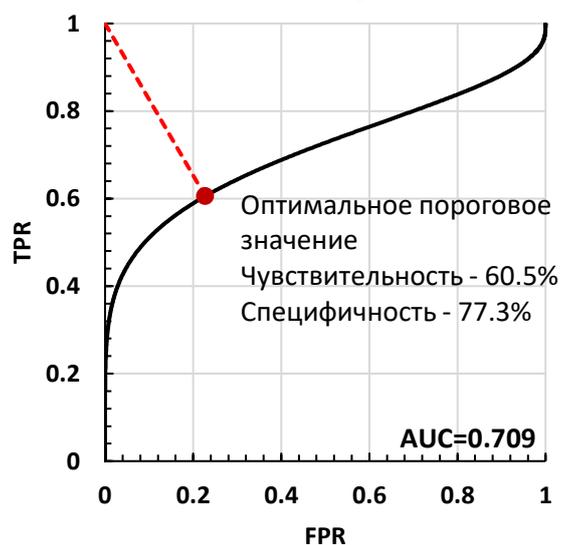
$$SMD = \frac{M_1 - 0}{1 \cdot \sqrt{\frac{k+1}{2}}} \Rightarrow M_1 = SMD \sqrt{\frac{2}{k+1}}.$$

Теперь зная  $M_1$ ,  $SD_1$ ,  $M_2$ ,  $SD_2$  мы можем оценить площадь под ROC-кривой, соответствующую определённым значениям  $SMD$  и  $k$  с помощью алгоритма, обоснованного выше и реализованного в Приложении.

## РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ ПРЕДЛАГАЕМОГО ПОДХОДА

### Оценка диагностических характеристик на основании выборочных оценок

Основной исследовательский интерес в этой работе связан с неинвазивным изучением микроциркуляторного русла и поиском количественных критериев, применимых для диагностики различных клинических состояний, таких, как осложнения сахарного диабета [15] или увеличение риска сердечно-сосудистых событий [16]. Поэтому в качестве примеров для оценки работы предлагаемых подходов будут использованы результаты изучения состояния микроциркуляции, полученные как в ходе анализа литературных источников, так и в ходе проведения собственных измерений.



**Рис. 3.** Результаты ROC-анализа данных, представленных в статье Cupisti et al. [17], на основании набора описательных статистик ( $M_1$ ,  $SEM_1$ ,  $n_1$ ,  $M_2$ ,  $SEM_2$ ,  $n_2$ ).

В работе Cupisti et al. [17] авторы исследовали увеличение кровотока (перфузии) на фоне пробы с ацетилхолином у контрольной группы, у пациентов с хронической почечной недостаточностью и у пациентов с артериальной гипертензией без поражения почек. Авторами были получены статистически значимые различия ( $p < 0.001$ ) между пациентами контрольной группы и лицами с артериальной гипертензией по уровню прироста перфузии (в %) на фоне ионофореза ацетилхолина. В контрольной группе показатель перфузии увеличился до  $683 \pm 92$  %, а у лиц с гипертензией до  $434 \pm 48$  % от исходного уровня. Так как перед авторами исследования стояли другие цели, оценка диагностической информативности прироста перфузии в ответ на ионофорез ацетилхолина не была проведена. Однако, нам стало интересно, с какой надёжностью на основании данного прироста можно оценить, есть у пациента артериальная гипертензия, или он относится к контрольной группе. Так как в качестве отклонений авторы указали

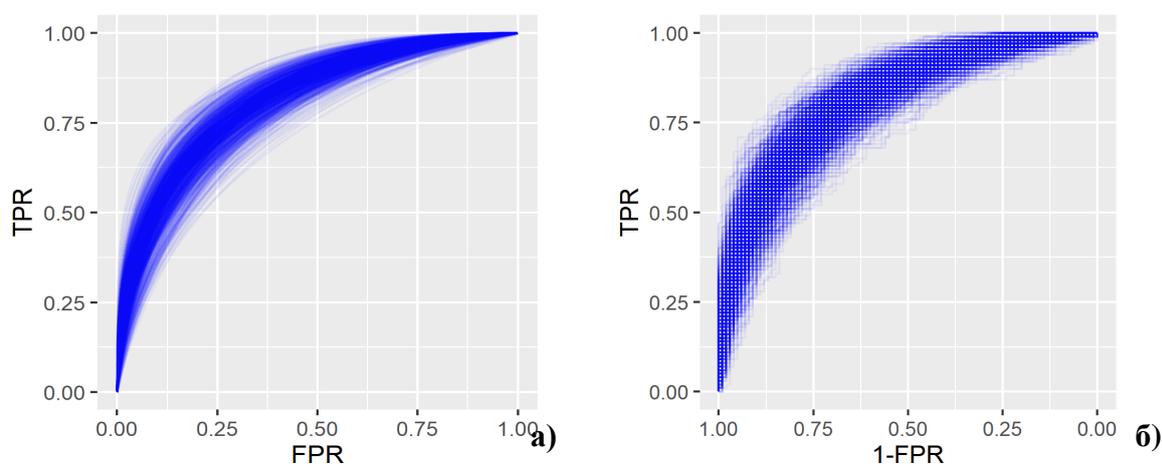
не стандартное отклонение ( $SD$ ), а стандартную ошибку среднего значения ( $SEM$ ), стандартное отклонение было вычислено как  $SD = SEM \cdot \sqrt{n}$ . Далее с шагом в 1 % была проведена оценка диагностической информативности прироста перфузии с помощью алгоритма, описанного ранее и реализованного в Приложении. Результаты построения ROC-кривой представлены на рисунке 3.

Мы видим, что несмотря на высокий уровень статистической значимости ( $p < 0.001$ ), параметр обладает лишь умеренной диагностической информативностью, и, если мы захотим использовать его для постановки диагноза «артериальная гипертензия», мы упустим 29.5 % пациентов с заболеванием, а у 22.7 % ложно предположим его наличие.

Полученный результат демонстрирует возможность использования предложенного нами подхода для оценки информативности количественных признаков по описательным статистикам без возможности провести первичный анализ данных.

### Сравнение предлагаемого подхода со стандартным способом проведения ROC-анализа

Для проведения ROC-анализа с помощью стандартных решений, необходим доступ к первичным данным [10]. Для того, чтобы понять, насколько отличающуюся от «стандартной» оценку даёт предлагаемый в настоящей работе подход был проведён следующий эксперимент. Для моделирования были использованы данные о распределении параметра «эндотелиальная активность» у лиц с ишемической болезнью сердца и у здоровых добровольцев старше 50 лет из статьи A. Fronek et al. [5]. На основании описательных статистик, представленных в статье, было сгенерировано значение показателя «эндотелиальная активность» в двух популяциях численностью по 100 000 наблюдений каждая: лица с заболеванием (ишемическая болезнь сердца) ( $\mu_1 = 61$ ,  $\sigma_1 = 34.3$ ) и лица без заболевания (здоровые добровольцы старше 50 лет) ( $\mu_2 = 103.3$ ,  $\sigma_2 = 32.5$ ). Эти популяции выступали в качестве модельных генеральных совокупностей. Из каждой популяции была сделана случайная выборка по 100 человек. Далее для выборок были рассчитаны  $M_1$ ,  $SD_1$ ,  $M_2$ ,  $SD_2$  и проведён ROC-анализ с помощью предлагаемого подхода и стандартным способом (реализованным в библиотеке «pROC» v.1.16.2 [18]). Случайное извлечение выборок было проведено 1000 раз. По результатам расчётов было построено 2 графика (рис. 4). На одном графике представлены 1000 ROC-кривых, построенных традиционным способом, а на втором – с применением нового алгоритма (прозрачность каждой ROC-кривой – 95 %).



**Рис. 4.** Результаты моделирования многократного проведения двух способов ROC-анализа на выборках, извлекаемых из двух генеральных совокупностей: **а)** ROC-кривые, построенные на основании выборочных оценок, **б)** ROC-кривые, построенные с помощью библиотеки pROC.

Также 1000 раз была рассчитана площадь под ROC-кривой ( $AUC$ ). 5 % квантиль  $AUC$  на основании выборочных оценок составил 0.762, а на основании стандартной оценки – 0.763, 50 % квантиль – 0.813 и 0.816, 95 % квантиль – 0.860 и 0.861, соответственно.

Малый уровень расхождений на графиках и при анализе  $AUC$  демонстрирует отсутствие систематической ошибки при проведении оценки с помощью предлагаемого подхода.

### Анализ данных с выборочным распределением, отличным от нормального

В большинстве медицинских исследований авторы предполагают нормальность распределений анализируемых переменных и используют в качестве параметров описательной статистики количественных переменных средние арифметические значения и стандартные отклонения (или стандартную ошибку среднего). Однако, всё чаще наблюдается ситуация, когда авторы предполагают отличие формы распределения от нормального и тогда в качестве описательных статистик используются медианы и квартили. Тем не менее существует подход, позволяющий вычислить средние арифметические значения и стандартные отклонения на основании рассчитанных медиан и квартилей (3) для включения этих данных в мета-анализы и систематические обзоры.

Для того, чтобы проверить, применим ли в таких случаях предлагаемый нами подход и насколько сильно результаты ROC-анализа, проведённого по  $M$  и  $SD$ , отличаются от результатов ROC-анализа, проведённого на основании первичных данных, мы проанализировали наблюдения, полученные нами в ходе измерения микроциркуляции на фоне тепловой пробы у лиц с микроциркуляторными нарушениями, вызванными сахарным диабетом (СД). В исследование включено две группы обследуемых: в первой группе 54 пациента с СД и микроциркуляторными нарушениями, во второй – 54 условно здоровых добровольца без признаков микроциркуляторных нарушений.

Всем обследуемым было проведено исследование кожной микроциркуляции крови методом лазерной доплеровской флоуметрии на фоне проведения тепловой пробы. Было оценено два показателя: относительный прирост перфузии в момент максимальной гиперемии (LTH) и площадь под микроциркуляторной кривой за первые две минуты нагрева (AUMC).

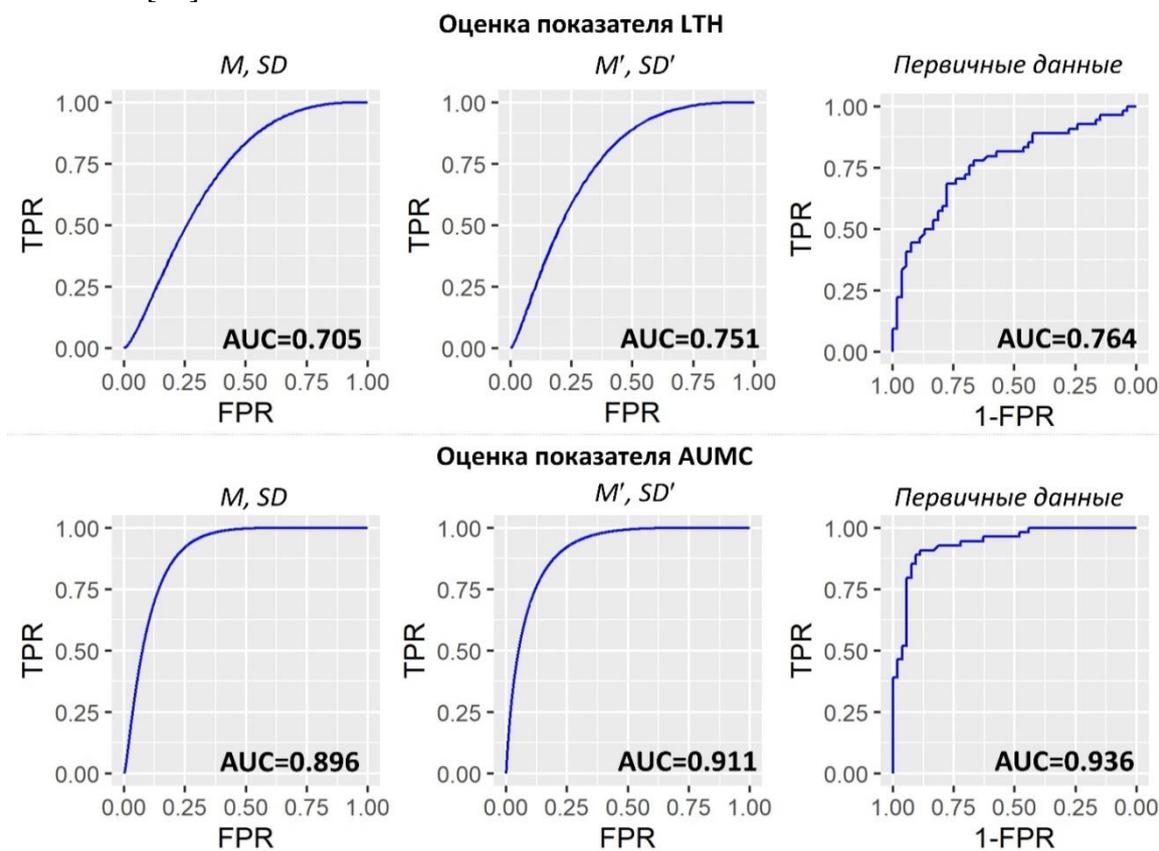
Для обоих показателей в первой и второй группах были рассчитаны средние арифметические значения ( $M$ ), стандартные отклонения ( $SD$ ), медианы ( $m$ ) и квартили ( $q_1$ ,  $q_3$ ). Также была проведена оценка средних арифметических значений и стандартных отклонений ( $M'$ ,  $SD'$ ) по формуле (3). Результаты оценки перечисленных параметров в двух группах приведены в таблице 1.

**Таблица 1.** Результаты оценки описательных статистик в группах для показателей LTH и AUMC

Параметр описательной статистики	Показатель LTH		Показатель AUMC	
	Группа 1 ( $n = 54$ )	Группа 2 ( $n = 54$ )	Группа 1 ( $n = 54$ )	Группа 2 ( $n = 54$ )
$M$	4.7	7.5	828.9	1690.8
$SD$	2.9	4.3	323.2	600.1
$m$	3.96	6.74	813	1619
$q_1$	2.91	4.72	620	1259
$q_3$	5.67	8.85	1032	1868
$M'$	4.18	6.77	821.6	1582.2
$SD'$	2.10	3.15	314.4	464.2

На основании данных, представленных в таблице 1, был проведён ROC-анализ в трёх вариантах: 1) на основании  $M$  и  $SD$ ; 2) на основании  $M'$  и  $SD'$ ; 3) на основании первичных данных (с помощью библиотеки «rROC»). Результаты ROC-анализа представлены на рисунке 5.

ROC-анализ на основании  $M'$  и  $SD'$  оказался более приближенным к анализу на основании первичных данных, чем анализ на основании  $M$  и  $SD$ . Это связано с тем, что на среднее арифметическое значение и стандартное отклонение при их расчёте стандартным способом сильное влияние оказывают высокие значения реактивности микроциркуляторного русла. Особенно сильно этот эффект заметен во второй группе (контрольная группа). В таких случаях авторы прибегают к расчёту медиан и квартилей, поскольку они менее подвержены влиянию таких значений. При этом количество статей, в которых использованы медианы и квартили ежегодно увеличивается. Количество медицинских статей в базе данных PubMed, у которых в названии или в аннотации были упомянуты медианы и квартили, в 1990 году составляло 14, в 2000 – 99, в 2010 – 409, в 2020 – 1131 [19].



**Рис. 5.** Результаты проведения ROC-анализа показателей LTH и AUMC тремя различными способами.

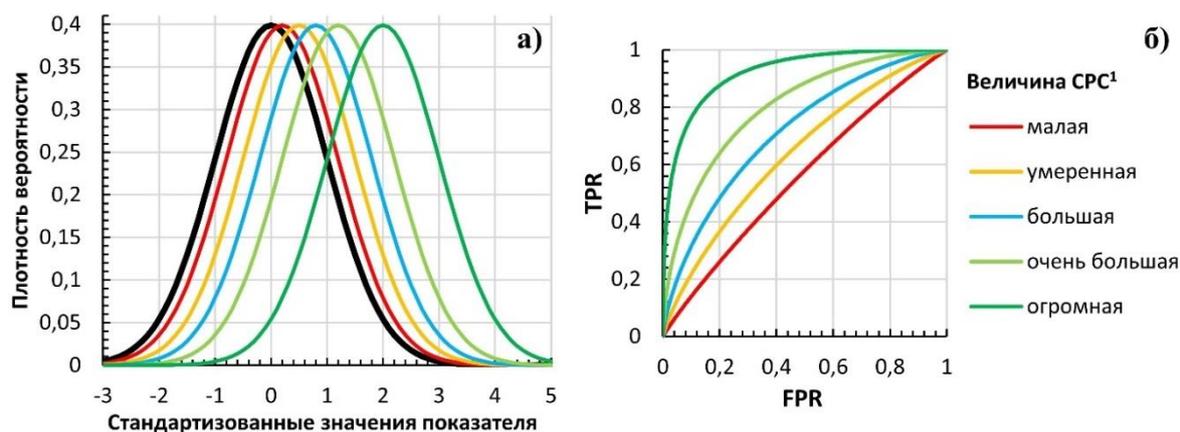
Таким образом, предлагаемый подход может быть применён и в ситуации, когда распределение оцениваемых авторами признаков несколько отличается от нормального и вместо  $M$  и  $SD$  авторами приведены медианы и квартили.

### Анализ диагностических характеристик на основании стандартизованной разности средних значений

Выше был описан подход оценки диагностической информативности на основании стандартизованной разности средних, который используется в мета-анализах как показатель величины эффекта при изучении количественных признаков. Согласно широко используемой интерпретации, предложенной J. Cohen, величина эффекта, равная

0.2, является малой, 0.5 – умеренной, 0.8 – большой [9]. При этом в более современных работах была предложена дополнительная градация: 1.2 – очень большая, 2 – огромная [20]. Однако, если рассчитать значения  $AUC$  для обозначенных величин эффекта и сопоставить эту интерпретацию с интерпретацией диагностической информативности по ROC-анализу, мы получим неоднозначную картину (рис. 6).

В таблице 2 представлены результаты расчёта значений  $AUC$  для различных градаций величины эффекта и для различных  $k$ . Мы видим, что малая и умеренная величина эффекта соответствует слабой диагностической информативности. Большая величина эффекта будет соответствовать только умеренной и слабой диагностической информативности. Для огромной величины эффекта мы получим  $AUC \geq 0.9$  (соответствующий отличной диагностической информативности [4]) только в случаях, если  $k \approx 1.12$  и меньше.



**Рис. 6.** Результаты анализа диагностической информативности количественного показателя на основании стандартизованной разницы средних (CPC). **а)** Чёрной линией обозначен стандартизованный график плотности вероятности условной «контрольной» группы. Цветными линиями обозначены графики плотности вероятности, соответствующие различным градациям CPC по [10] и [14]. **б)** Цветными линиями обозначены ROC-кривые, рассчитанные на основании описанного выше подхода и соответствующие различным градациям CPC.

**Таблица 2.** Оценка значений  $AUC$  в зависимости от величины эффекта ( $SMD$ ) и соотношения стандартных отклонений  $k$

Значение $k$ ( $SD_2/SD_1$ )	Значение $SMD$ , величина эффекта по [9], [20]				
	0.2 малая	0.5 умеренная	0.8 большая	1.2 очень большая	2 огромная
0.5	0.580	0.696	0.794	0.891	0.979
0.57	0.576	0.687	0.782	0.879	0.974
0.67	0.571	0.674	0.765	0.861	0.964
0.8	0.564	0.658	0.744	0.837	0.949
1	0.555	0.637	0.713	0.801	0.920
1.25	0.547	0.615	0.680	0.759	0.879
1.5	0.540	0.598	0.654	0.724	0.838
1.75	0.534	0.584	0.632	0.694	0.801
2	0.529	0.572	0.615	0.669	0.767

Примечания: 1.  – отличная диагностическая информативность ( $AUC \geq 0.9$ );  
 2.  – хорошая диагностическая информативность ( $0.8 \leq AUC < 0.9$ );  
 3.  – умеренная диагностическая информативность ( $0.7 \leq AUC < 0.8$ );  
 4.  – слабая диагностическая информативность ( $0.5 < AUC < 0.7$ ).

Таким образом, используя предлагаемый подход, мы можем рассчитать, что, например, при  $k = 1$  порогу  $AUC$  для умеренной диагностической информативности ( $AUC > 0.7$ ) соответствует  $SMD \geq 0.75$ , для хорошей ( $AUC > 0.8$ ) –  $SMD \geq 1.2$ , для отличной ( $AUC > 0.9$ ) –  $SMD \geq 1.83$ . То есть интерпретация с позиции оценки величины эффекта не всегда позволяет сделать вывод о возможности применения показателя в практических целях.

Рассмотрим пример мета-анализа, проведённого D. Fuchs et al. (2017) [21]. Авторы провели систематический обзор и мета-анализ, посвящённый измерению состояния микроциркуляции методами лазерной доплеровской флоуметрии и лазерной спекл-контрастной визуализации на фоне локальной тепловой пробы у пациентов с сахарным диабетом. В аннотации статьи авторы пишут, что согласно определению J. Cohen, в своём исследовании они наблюдают большую величину эффекта, равную (95 % ДИ)  $-0.78$  ( $-1.1, -0.51$ ). Авторы делают заключение о том, что методика оценки локальной тепловой гиперемии перспективна в качестве диагностического инструмента для выявления диабет-ассоциированных микроциркуляторных нарушений.

Однако, если мы обратимся к таблице 2 и к рассчитанным ранее пороговым значениям  $SMD$  для различных градаций диагностической информативности, мы увидим, что указанный  $SMD$  и границы 95 % ДИ для него соответствуют значениям  $AUC < 0.8$ , что свидетельствует лишь об умеренной диагностической информативности проанализированного подхода. То есть, несмотря на большую величину эффекта, диагностическая информативность большой не является.

Таким образом, для того, чтобы оценить возможность применения количественного признака в качестве диагностического недостаточно использования величины эффекта, выраженной в виде  $SMD$ : необходимо проведение дополнительного анализа, для оценки диагностической информативности. В представленной статье описан алгоритм и способ проведения такого анализа.

## ЗАКЛЮЧЕНИЕ

В статье предложен способ оценки диагностической информативности количественных переменных без доступа к первичным данным путём проведения ROC-анализа на основании кумулятивных функций распределений. Предложены алгоритмы проведения расчётов на основании средних арифметических значений и стандартных отклонений, медиан и квартилей, стандартизированной разности средних значений.

На основании собственных наблюдений, моделирования, анализа литературных данных была продемонстрирована применимость предлагаемого подхода на примере оценки диагностической информативности количественных признаков, анализируемых в ходе изучения состояния микроциркуляции. Было показано, что стандартная градация «величины эффекта» не может быть напрямую интерпретирована для диагностической информативности изучаемого количественного признака. Были предложены значения  $SMD$ , соответствующие различным градациям диагностической информативности. При равенстве дисперсий  $AUC > 0.7$  будет наблюдаться при  $SMD \geq 0.75$ ,  $AUC > 0.8$  – при  $SMD \geq 1.2$ ,  $AUC > 0.9$  – при  $SMD \geq 1.83$ .

Представленные в статье результаты могут быть использованы специалистами медицинского и биологического профиля для оценки диагностической информативности различных количественных признаков без доступа к первичным данным, например, при анализе литературных источников, в которых авторы не провели соответствующих расчётов или при изучении мета-анализов и систематических обзоров.

Исследование выполнено при финансовой поддержке гранта Президента Российской Федерации, внутренний номер гранта МК-1786.2020.7 (соглашение № 075-15-2020-354).

The reported study was funded by the grant of the President of the Russian Federation, grant number МК-1786.2020.7 (agreement No. 075-15-2020-354).

### СПИСОК ЛИТЕРАТУРЫ

1. Бойко И.Ю., Анисимов Д.С., Смолякова Л.Л., Рязанов М.А. Подход к отбору значимых признаков при решении биомедицинских задач бинарной классификации данных с микрочипов. *Математическая биология и биоинформатика*. 2020. Т. 15. № 1. С. 4–19. doi: [10.17537/2020.15.4](https://doi.org/10.17537/2020.15.4)
2. Alemayehu D., Zou K.H. Applications of ROC analysis in medical research: recent developments and future directions. *Academic Radiology*. 2012. V. 19. № 12. P. 1457–1464. doi: [10.1016/j.acra.2012.09.006](https://doi.org/10.1016/j.acra.2012.09.006)
3. Nakas C.T., Reiser B. Editorial for the special issue of “Statistical Methods in Medical Research” on “Advanced ROC analysis”. *Statistical Methods in Medical Research*. 2018. V. 27. № 3. P. 649–650. doi: [10.1177/0962280217742536](https://doi.org/10.1177/0962280217742536)
4. Hosmer D.W., Lemeshow S., Sturdivant R.X. *Applied Logistic Regression*. John Wiley & Sons, 2013. 528 p.
5. Fronek A., Allison M. Noninvasive evaluation of endothelial activity in healthy and diseased individuals. *Vascular and Endovascular Surgery*. 2014. V. 48. № 2. P. 134–138. doi: [10.1177/1538574413508229](https://doi.org/10.1177/1538574413508229)
6. Ovadia-Blechman Z., Avrahami I., Weizman-Shammai E., Sharir T., Eldar M., Chouraqui P. Peripheral microcirculatory hemodynamic changes in patients with myocardial ischemia. *Biomedicine & Pharmacotherapy*. 2015. V. 74. P. 83–88. doi: [10.1016/j.biopha.2015.07.011](https://doi.org/10.1016/j.biopha.2015.07.011)
7. Куликов Д.А., Глазков А.А., Ковалева Ю.А., Балашова Н.В., Куликов А.В. Перспективы использования лазерной доплеровской флоуметрии в оценке кожной микроциркуляции крови при сахарном диабете. *Сахарный диабет*. 2017. Т. 20. № 4. С. 279–285. doi: [10.14341/DM8014](https://doi.org/10.14341/DM8014)
8. Glass G.V. Integrating Findings: The Meta-Analysis of Research. *Review of Research in Education*. 1977. V. 5. № 1. P. 351–379. doi: [10.2307/1167179](https://doi.org/10.2307/1167179)
9. Cohen J. A power primer. *Psychological Bulletin*. 1992. V. 112. № 1. P. 155–159. doi: [10.1037/0033-2909.112.1.155](https://doi.org/10.1037/0033-2909.112.1.155)
10. Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006. V. 27. № 8. P. 861–874. doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010)
11. Perkins N.J., Schisterman E.F. The inconsistency of “optimal” cut-points using two ROC based criteria. *American Journal of Epidemiology*. 2006. V. 163. № 7. P. 670–675. doi: [10.1093/aje/kwj063](https://doi.org/10.1093/aje/kwj063)
12. Wan X., Wang W., Liu J., Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology*. 2014. V. 14. № 135. P. 1–13.
13. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, 2013. URL: <http://www.R-project.org/>. (accessed 09.12.2020).
14. Lee D.K. Alternatives to P value: Confidence interval and effect size. *Korean Journal of Anesthesiology*. 2016. V. 69. № 6. P. 555–562. doi: [10.4097/kjae.2016.69.6.555](https://doi.org/10.4097/kjae.2016.69.6.555)
15. Kulikov D., Glazkov A., Dreval A., Kovaleva Y., Rogatkin D., Kulikov A., Molochkov A. Approaches to improve the predictive value of laser Doppler flowmetry in detection of microcirculation disorders in diabetes mellitus. *Clinical Hemorheology and Microcirculation*. 2018. V. 70. № 2. P. 173–179. doi: [10.3233/CH-170294](https://doi.org/10.3233/CH-170294)
16. Глазкова П.А., Терпигорев С.А., Куликов Д.А., Иванова Н.А., Глазков, А.А. Пути повышения диагностической значимости метода лазерной доплеровской

- флоуметрии при оценке кожной микроциркуляции у пациентов с артериальной гипертензией. *Артериальная гипертензия*. 2019. Т. 25. № 1. С. 74–83. doi: [10.18705/1607-419X-2019-25-1-74-83](https://doi.org/10.18705/1607-419X-2019-25-1-74-83)
17. Cupisti A., Rossi M., Placidi S., Fabbri A., Morelli E., Vagheggin G., Meola M., Barsotti G. Responses of the Skin Microcirculation to Acetylcholine in Patients with Essential Hypertension and in Normotensive Patients with Chronic Renal Failure. *Nephron*. 2000. V. 85. № 2. P. 114–119. doi: [10.1159/000045643](https://doi.org/10.1159/000045643)
  18. Robin X., Turck N., Hainard A., Tiberti N., Lisacek F., Sanchez J.C., Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011. V. 12. № 1. P. 1–8. doi: [10.1186/1471-2105-12-77](https://doi.org/10.1186/1471-2105-12-77)
  19. *PubMed Database*. URL: <https://pubmed.ncbi.nlm.nih.gov/> (accessed 09.12.2020).
  20. Sawilowsky S.S. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*. 2009. V. 8. № 2. P. 597–599.
  21. Fuchs D., Dupon P.P., Schaap L.A., Draijer R. The association between diabetes and dermal microvascular dysfunction non-invasively assessed by laser Doppler with local thermal hyperemia: a systematic review with meta-analysis. *Cardiovascular Diabetology*. 2017. V. 16. № 11. P. 1–12. doi: [10.1186/s12933-016-0487-1](https://doi.org/10.1186/s12933-016-0487-1)

Рукопись поступила в редакцию 19.10.2020, переработанный вариант поступил 02.12.2020.  
Дата опубликования 17.12.2020.

===== INFORMATION AND COMPUTER =====  
===== TECHNOLOGIES IN BIOLOGY AND MEDICINE =====

## Assessing Diagnostic Accuracy of Quantitative Data in Biomedical Studies Using Descriptive Statistics and Standardized Mean Difference

Glazkov A.A.<sup>1</sup>, Kulikov D.A.<sup>1,2</sup>, Glazkova P.A.<sup>1</sup>

<sup>1</sup>Moscow Regional Research and Clinical Institute, Moscow, Russia

<sup>2</sup>Moscow Region State University, Moscow Region, Mytishchi, Russia

**Abstract.** ROC analysis is the most used method for analyzing the diagnostic accuracy of quantitative data in biomedical research. ROC analysis generates a curve describing the frequencies of true positive and false positive results for different degrees of the analyzed variable. However, in many publications devoted to the application of quantitative diagnostic methods, this analysis is not carried out: researchers report only analysis of statistical significance for the groups difference. In meta-analyses, the estimated parameter is the effect size expressed through standardized mean difference. The article describes the approach, which allows performing ROC analysis using cumulative normal distribution functions for studied and controlling groups. The proposed approach can be used to evaluate the diagnostic accuracy of quantitative variables on the base of one of the sets of descriptive statistics (mean and standard deviation, or median and quartiles) or the value of standardized mean difference. Examples of application of the proposed approach on model data, on data from literature sources, as well as on the authors' own observations are given as an example of assessment of diagnostic accuracy of quantitative variables analyzed in the microcirculation studies in various diseases.

The results presented in the article can be used by medical and biological specialists to assess the diagnostic accuracy of various quantitative variables without access to primary data.

**Key words:** *standardized mean difference, diagnostic accuracy, ROC-analysis, statistics in medicine, normal distributions, simulation.*