

УДК 004.67

## ВЫБОР НАИЛУЧШЕГО ПРЕДИКТОРА В МЕДИЦИНСКИХ ИССЛЕДОВАНИЯХ С ПОМОЩЬЮ ROC-АНАЛИЗА

П. Р. ИВАНОВА, В. С. ДАШКО

Научный руководитель В. А. ЛИВИНСКАЯ, канд. физ.-мат. наук, доц.  
Белорусско-Российский университет  
Могилев, Беларусь

Основной целью данной работы являлось определение наилучшего предиктора патологии, используя ROC-анализ, и придание количественного измерения результатам классификации для каждого фактора с помощью нахождения доверительных интервалов.

С использованием программного обеспечения Реоспектр (Нейрософт) были получены определенные физические характеристики реографической кривой, описывающей динамику сопротивления живой ткани за определенный временной промежуток: амплитуды артериальной части волны  $A_{арт}$ ; время распространения пульсовой волны от сердца  $Q_x$ ; систолический максимум реограммы  $A_{сист}$ ; время восходящей части волны  $\alpha$ ; время быстрого кровенаполнения  $\alpha_1$ ; время медленного кровенаполнения  $\alpha_2$ ; время общей систолы  $T_{общ}$ ; длительность катакроды  $T_{кат}$  [1].

С помощью языка R из отдельных файлов в текстовом формате, содержащих информацию о каждом пациенте, был сформирован датасет из девяти столбцов, восемь из которых являлись физическими параметрами кривой, а девятый отвечал за принадлежность пациента к одной из двух групп: пациенты с наличием одного из трех патологических состояний: синдрома полиорганной дисфункции (ОАО СПОД), пациенты с циррозом печени (ОАО печень), пациенты с наличием хронических заболеваний желудка и поджелудочной железы (ОАО гастро), а также контрольная группа – практически здоровые. Отнесение пациента к одной из двух групп описывается бинарной переменной, принимающей два значения: 1 – пациент болен; 0 – пациент здоров. Под реографической кривой была рассчитана площадь для каждого пациента. Далее выявление значимых различий в группах на основании различных предикторов (различных характеристик реографической кривой и площади под ней) осуществлялось с помощью непараметрического статистического критерия Краскела – Уоллиса в среде R.

Для дальнейшего подтверждения возможности использования площади под реограммой в качестве предиктора состояния больного проводился ROC-анализ (табл. 1), при выполнении которого для фактора, выступающего классификатором (в данном случае это площадь под реографической кривой), рассчитываются специальные метрики, являющиеся характеристиками качества разбиения – чувствительность (относительный показатель верно классифицированных положительных случаев) и специфичность (доля неверно

классифицированных отрицательных случаев), по которым и строится ROC-кривая, представленная на рис. 1.

Табл. 1. Результаты ROC-анализа

| Тестовая переменная | AUC   | Стд. ошибка <sup>a</sup> | Ошибка первого рода | Асимптотический 95-процентный доверительный интервал |                 |
|---------------------|-------|--------------------------|---------------------|--|-----------------|
|                     |       |                          |                     | Нижняя граница                                       | Верхняя граница |
| Площадь по РК       | 0,763 | 0,053                    | 0,000               | 0,660  | 0,867           |
| alf1                | 0,598 | 0,061                    | 0,138               | 0,478  | 0,719           |
| alf2                | 0,632 | 0,059                    | 0,047               | 0,517  | 0,746           |
| Tkat                | 0,663 | 0,058                    | 0,014               | 0,550  | 0,776           |
| RI                  | 0,690 | 0,061                    | 0,004               | 0,570  | 0,810           |
| DIA                 | 0,584 | 0,061                    | 0,204               | 0,465  | 0,703           |
| Vmaks               | 0,801 | 0,046                    | 0,000               | 0,710  | 0,892           |

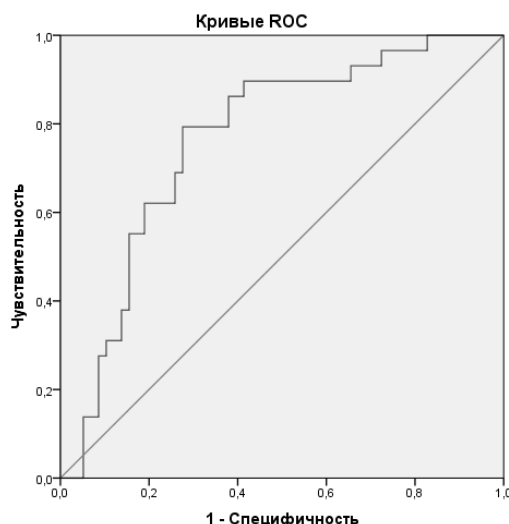


Рис. 1. ROC-кривая для показателя «площадь» под реографической кривой как предиктора болезни

Идеальный диагностический тест должен иметь Г-образную форму характеристической кривой, проходящей через верхний левый угол, в котором доля истинно положительных случаев равна 1, а доля ложноположительных случаев равна 0. В рассматриваемом случае диагностический тест является эффективным, т. к. напоминает букву «Г», а также находится выше диагонали графика. Разбиение признается качественным, если площадь под кривой (AUC) принимает значение больше, чем 0,5.

#### СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Интегративный показатель состояния артериального печеночного кровотока у пациентов при критических состояниях / С. А. Точило [и др.] // Вестн. Витеб. гос. мед. ун-та. – 2019. – Т. 18, № 3. – С. 52–60.