

multi-point systems spacecraft control]. *Aviakosmicheskoe priborostroenie* [Aerospace Instrument], 2009, no. 5, pp. 8–11.

3. Goryachev N. V., Grab I. D., Petelin K. S., Trusov V. A., Kochegarov I. I., Yurkov N. K. Avtomatizirovannyi vybor sistemy okhlazhdeniya teplonagruzhennykh elementov radioelektronnykh sredstv [Automated selection of cooling system of thermally loaded elements of radioelectronic facilities]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2013, no. 4, pp. 136–143.

4. Kashcheev N. A. *Radiotekhnicheskie sredstva upravleniya kosmicheskimi apparatami* [Radio-technical means of spaceship control]. Moscow, 2005. 202 p.

5. Leonov A. I., Fomichev K. I. *Monoimpulsnaya radiolokatsiya* [Monoimpulse radiolocation]. Moscow, Radio i svyaz, 1984. 312 p.

6. Poltavskiy A. V., Yurkov N. K. Modifikatsiya modeli sistemy upravleniya podvizhnym obektom [Modification of model of moving object control]. *Nadezhnost i kachestvo slozhnykh sistem* [Reliability and quality of complex system], 2014, no. 1 (5), pp. 65–70.

7. Yurkov N. K., Danilov A. M., Garkina I. A., Lapshin Ye. V. Sistemnye metodologii, identifikatsiya sistem i teoriya upravleniya: promyshlennye i aerokosmicheskie prilozheniya [System methodologies, identification of systems and management theory]. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* [News of Higher Educational Institutions. Volga region. Engineering Sciences], 2009, no. 1, pp. 3–11.

8. Yurkov N. K., Bukharov A. Ye. Povyshenie radiolokatsionnogo kontrasta. Sistemnyy podkhod [Increase of radiolocation contrast. System approach]. Penza, Penza State Univ. Publ., 2013. 398 p.

9. Yakimov A. N. *Proektirovanie mikrovolnovykh anten s uchetom vneshnikh vozdeystviy* [Design of microwave antennas taking into account external effects]. Penza, Penza State Univ. Publ., 2004. 260 p.

10. Blinov A. V., Kanakov A. G., Trusov V. A., Yurkov N. K. Acceptance checking methods for uhf electronic components. *Measurement Techniques*, 2000, vol. 43, no. 10, pp. 895–901.

11. Botvinkin P. V. Information security of Glonass-based automated navigation systems for ground transportation monitoring and supervisory control. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2014, no. 3 (27), pp. 187–196.

УДК 616:519.2

ИСПОЛЬЗОВАНИЕ ПРИ ПРИНЯТИИ ДИАГНОСТИЧЕСКИХ РЕШЕНИЙ РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКОГО АНАЛИЗА ВЛИЯНИЯ ВИРУСНО-БАКТЕРИАЛЬНЫХ МАРКЕРОВ НА КЛИНИЧЕСКУЮ КАРТИНУ ЗАБОЛЕВАНИЯ

Статья поступила редакцию 18.11.2014, в окончательном варианте 09.12.2014.

Дедов Алексей Владимирович, кандидат медицинских наук, доцент, Астраханская государственная медицинская академия, 414000, Российская Федерация, г. Астрахань, ул. Бакинская, 121, e-mail: dedov1965.d@yandex.ru

Попов Георгий Александрович, доктор технических наук, профессор, Астраханский государственный технический университет, 414025, Российская Федерация, г. Астрахань, ул. Татищева, 16, e-mail: popov@astu.org

Статья посвящена формированию процедуры оценки степени взаимосвязи между заболеванием и вирусно-бактериальными маркерами, что может являться основой при принятии диагностических решений. Процедура реализована на основе методов математической статистики применительно к влиянию вирусно-бактериальных маркеров на признаки, характеризующие клиническую картину хронических диффузных заболеваний печени (ХДЗП). Исследования опираются на результаты обследований 165 пациентов хроническим гепатитом (ХГ) и циррозом печени (ЦП). Для решения пробле-

мы авторами был модифицирован применительно к многомерному случаю классический метод Манна – Уитни для проверки однородности двух выборок (многомерное обобщение). На основе предложенного метода выявлены наиболее значимые маркеры при 5% уровне статистической значимости. Их оказалось 8 из 26 исследованных маркеров. Для статистически значимых маркеров был проведен анализ их взаимосвязей с диагностическими признаками. Результаты анализа отражены в трех таблицах. В первой для каждого значимого маркера перечислены связанные с ним наиболее важные диагностические признаки. Во второй показаны взаимосвязи наиболее важных признаков со статистически значимыми маркерами. В третьей дается полная информация по корреляционной взаимосвязи маркеров между собой.

Ключевые слова: принятие диагностических решений, метод Манна – Уитни, многомерное обобщение, корреляционный анализ, вирусно-бактериальные маркеры, хронические заболевания печени, Прикаспийский регион

THE APPLICATION OF THE OUTCOMES OF THE STATISTICAL ANALYSIS OF THE VIRAL AND BACTERIAL MARKERS INFLUENCE ON THE CLINICAL PICTURE OF THE DISEASE IN DIAGNOSTIC DECISIONS

Dedov Aleksey V., Ph.D. (Medicine), Associate Professor, Astrakhan State Medical Academy, 121 Bakinskaya St., Astrakhan, 414000 Russian Federation, e-mail: dedov1965.d@yandex.ru

Popov Georgiy A., D.Sc. (Engineering), Professor, Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414025, Russian Federation, e-mail: popov@astu.org

The article is devoted to the creation of the procedure for the evaluation of the interrelation degree between the disease and the viral and bacterial markers that could be the basement of diagnostic decisions. This procedure was realised on the basis of mathematical, statistical methods with the special relation to the influence of viral and bacterial markers to the significant characteristics of clinical picture of chronic diffuse liver diseases (CDLD). The researches were based on the large diagnostic information collected during many years (165 patients). To solve this problem, the authors had used the multidimensional modification of the classical Mann – Whitney method for the checking the hypothesis of the uniformity of two choices. On the basement of the offered method the most significant markers were revealed with the 5 % significance level. Then 8 of 26 investigated markers were detected as the relevant ones. For statistically significant markers the analysis of their interrelations with the diagnostic signs was carried out. The results of the analysis were reflected in three tables. In the first one for each significant marker the most important diagnostic signs connected with it were listed. In the second table the interrelations between most important parameters with statistically significant markers were presented. In the third table the full information about the interrelation between markers themselves on the basis of their mutual correlations was given.

Keywords: diagnostic decisions, Mann – Whitney method, multidimensional modification, the correlation analysis, viral and bacterial markers, chronic diffuse liver diseases (CDLD), Near-the-Caspian-Sea region

Проблема взаимосвязи многообразных заболеваний, связанных с наличием в организме пациента различных вирусов и бактерий (как патогенных, так и нейтральных или даже полезных), а также антител к вирусно-бактериальным антигенам, представляет большой интерес. Выявление таких взаимосвязей позволяет выбирать более эффективные методы лечения, улучшать диагностический процесс.

Целью данной работы является формирование процедуры выработки рекомендаций при диагностике болезней на основе оценки влияния вирусно-бактериальных маркеров на клиническую картину на примере хронических заболеваний печени. Предлагаемые авторами методические подходы могут быть использованы как при решении других медико-

биологических проблем, так и в иных, не связанных с медициной, областях, при которых обрабатываются количественные данные, организованные в многомерные массивы.

Общая характеристика проблематики исследования вирусно-бактериальных маркеров. Для решения вышеуказанной проблемы необходимо прежде всего оценить характер влияния вирусно-бактериальных факторов на диагностическую картину заболевания. Настоящая работа и посвящена рассмотрению данной задачи. Анализ проводится применительно к хроническим диффузным заболеваниям печени (ХДЗП), а именно циррозу печени (ЦП) и хроническому гепатиту (ХГ). Болезни печени являются одной из ведущих причин смерти лиц среднего возраста в большинстве экономически развитых стран, включая Российскую Федерацию. Патология печени занимает 5-ое место в структуре заболеваемости среди всех видов болезней желудочно-кишечного тракта в России, Южном федеральном округе и Астраханской области как части Прикаспийского региона. На особенности течения заболеваний печени в этом регионе оказывают влияние характер и качество пищевых продуктов и воды, климатические, социально-экономические факторы, а также склонность к злоупотреблению алкоголем, курение, психосоциальные стрессы [10]. Общеизвестна роль вирусов гепатитов В и С как ведущих этиологических агентов, ответственных за развитие хронического гепатита (ХГ) и цирроза печени (ЦП). Следует отметить, что вирус гепатита D является эндемичным для Астраханской области и вносит значительный вклад в формирование всей группы хронических диффузных заболеваний печени (ХДЗП) [4]. Но лишь в немногих работах отмечается роль цитомегаловируса, вируса Эпштейн-Барр, герпес-вирусов, иных вирусов при ХГ и ЦП, причем чаще всего они рассматриваются как кофакторы, не имеющие самостоятельного значения. Участие бактериального фактора (патогенной и условно- патогенной микрофлоры) в патогенезе заболеваний печени изучено в ещё меньшей степени – имеются лишь единичные работы, указывающие на возможную роль в патогенезе ХДЗП хламидий, микоплазм [12, 14, 16].

Поскольку у авторов имеется достаточный набор результатов обследования пациентов по этим заболеваниям, формализованный в числовом виде, то в качестве основного метода анализа были выбраны методы математической статистики [6].

Ранее задача построения регрессионных моделей для установления диагноза по указанным заболеваниям рассматривалась авторами в работе [5] – были построены уравнения регрессии, в которых выходной (объясняемой) переменной был диагноз, представленный в числовой форме (0 – нет заболевания, 1 – есть заболевание). Результаты проведенного в этой работе анализа используются ниже. Статей по анализу влияния различных микроорганизмов на процесс течения определенных заболеваний достаточно много – укажем в том числе на работы [1, 3, 4]. Однако в большинстве исследований авторы исследуют разнообразные зависимости между выявлением того или иного микроорганизма и развитием заболевания, с использованием различных методов статистического анализа (корреляционный, факторный, вычисления относительного риска (RR) и др.). В данной работе предлагается процедура сравнения особенностей влияния вирусно-микробных факторов на клиническую картину, или «образ» болезни – без анализа возможных деталей патогенеза. Затем на этой основе формируется ответ на вопрос, существуют или нет достоверные различия между группами больных при наличии или отсутствии определенных вирусных или бактериальных антигенов; имеет ли место «модификация» клинической картины и какова степень ее выраженности. Совокупность полученных ответов и предлагается использовать в дальнейшем при формировании диагностического решения.

Для решения задачи взаимосвязи клинического диагноза и вирусно-бактериальных маркеров в работе предлагается использовать методы многомерного статистического анализа.

Исходные данные результатов обследования пациентов. В настоящей работе были использованы результаты обследований 96 лиц с ЦП и 69 больных с ХГ – всего 165 пациентов с ХДЗП. Средний возраст больных, наблюдавшихся в 2000–2014 гг., составил $48,8 \pm 1,7$ лет, где 1,7 – среднеквадратическое отклонение. Все пациенты были обследованы согласно национальным стандартам Российской Федерации с созданием базы данных, включающей как «паспортные данные», так и формализованные результаты физикальных, лабораторных, инструментальных методов исследования. Кроме того, в базу были включены сведения об изучаемых сывороточных маркерах вирусных и бактериальных инфекций (*Chlamidophyla pneumoniae*, *Mycoplasma pneumoniae*, *Helicobacter pylori*, вирус простого герпеса (HSV-1), вирус Эпштейна-Барр (ВЭБ), цитомегаловирус (ЦМВ), вирус гепатита А (HAV)) и антитела к атогенной и условно-патогенной микрофлоре (*E.coli*, *Proteus mirabilis*, *Candida albicans*, *Streptococcus spp.*, *Bacteroides*, *Klebsiella pneumoniae*) классов IgG, IgM, IgA или суммарные антитела методом ИФА (качественным, полуколичественным или количественным). Всего в базе данных у каждого больного учитывалось 135 параметров (смотри [5]).

Описание процедуры статистического анализа данных. Задача выявления статистической взаимосвязи между диагностическими признаками заболевания и вирусно-бактериальными маркерами в области математической статистики может быть сведена к задаче проверке гипотезы на однородность двух многомерных выборок. Именно для конкретного вирусно-бактериального маркера, данные по которому имеются в результатах обследования пациентов, формируются две выборки. Первая выборка – наборы результатов обследований по всем показателям больных, у которых зафиксировано либо наличие данного маркера (бинарный признак – «есть или нет маркер»), либо – при количественном определении (концентрация превышающая норму). Вторая выборка – аналогичные наборы по больным, у которых маркера либо не было, либо его концентрация не превышала некоторое пороговое значение. (В ряде случаев было возможным выделение дополнительной, третьей группы со сниженной относительно «донорской» концентрацией антител, определяемых количественно).

Рассматривается гипотеза H_0 : можно ли считать, что обе выборки подчиняются одному тому же вероятностному распределению, то есть неразличимы с точки зрения их вероятностных свойств. Справедливость данной гипотезы будет означать, что наличие изучаемого конкретного вирусно-бактериального маркера не влияет на особенности протекания заболевания печени. Альтернативной является гипотеза H_1 о несовпадении распределений, которым подчиняются выборочные значения для пациентов по следующим двум выборкам: пациентов, имеющих высокое содержание данного маркера; пациентов, у которых данный маркер практически отсутствует.

Имеется ряд особенностей в исходных данных, которые также повлияли на выбор методов решения проблемы. Прежде всего, в исходных данных, которые формировались на протяжении многих лет, имеются большие пропуски, причем эти пропуски разбросаны по данным достаточно нерегулярно. Основные причины этих пропусков: данные по ряду показателей стали собираться и фиксироваться относительно недавно, поэтому в более ранних наблюдениях данные по этим показателям отсутствуют. Кроме того, анализ ряда показателей требует значимых усилий и финансовых затрат, что ограничило круг пациентов, для которых эти анализы были проведены. Наконец, по отдельным показателям данные отсутствуют по различным текущим причинам, связанным с конкретным пациентом, конкретными условиям работы медицинского учреждения. Существуют методы восстановления пропущенных данных (см., например, [3, 7, 8, 9, 15]). Однако эти методы эффективны для случаев, когда пропуски имеют негрупповой, разбросанный характер и их количество не очень велико, что не выполняется для использованных в работе данных.

Далее многие из исходных показателей имеют субъективное содержание – эти параметры были получены в основном на основе опроса и физикального обследования пациентов.

Перечисленные особенности наложили определенные ограничения на выбор методов проверки гипотез. Один из возможных подходов мог бы опираться на проверку одномерных гипотез об однородности по каждому показателю, и если хотя бы по одному из показателей гипотеза об однородности не выполняется, то можно считать, что и исходные наборы не являются однородными. Имеется достаточно много методов проверки гипотезы об однородности для одномерных выборок: для нормально распределенных наборов – критерии Стьюдента, F -критерий, критерий Крамера – Уэлча, непараметрические критерии Смирнова, типа омега-квадрат (Лемана – Розенблатта), Вилкоксона (Манна – Уитни), Ван-дер-Вардена, Сэвиджа, хи-квадрат и др. [9, 11, 13]. Пример анализа многомерных данных приведен в [2]. Однако при таком подходе (на основе анализа одномерных выборок) игнорируется возможная взаимосвязь между признаками, то есть признаки рассматриваются как независимые, что, вообще говоря, некорректно. Кроме того, поскольку в задаче, рассматриваемой в данной работе, число признаков (как указано в предыдущем разделе) равно 135, то практическое воплощение описанного подхода в рамках данной работы трудно реализуемо ввиду ее трудоемкости. Поэтому целесообразно рассмотреть многомерные методы анализа однородности. Наиболее эффективными методами проверки гипотезы об однородности многомерных данных являются методы, разработанные для анализа нормально распределенных данных. В этом случае проблема проверки однородности сводится к проверке двух гипотез о равенстве ковариационных матриц обоих выборок и равенстве их векторов средних значений. Однако в рассматриваемом случае данные методы оказались неприменимыми, поскольку они опираются на наличие полного набора данных. Кроме того, предположение о нормальности неубедительно для субъективных данных, принимающих два значения: есть симптом (или признак) или его нет. Поэтому целесообразно использовать непараметрические методы, среди которых одним из наиболее известных является метод Манна – Уитни, модифицированный авторами с учетом особенностей имеющихся данных. Отметим, что в качестве многомерного обобщения критерия Манна – Уитни рассматривается критерий Краскела – Уоллиса. Однако с его помощью можно проверить только равенство медиан многомерных наблюдений, что недостаточно для вывода об однородности выборок. Поэтому возникает проблема модификации критерия с целью возможности его использования для анализа многомерных данных, описанных выше.

Предлагается следующий способ модификации процедуры на основе КМИ. Пусть имеется набор показателей X_1, X_2, \dots, X_L , и два набора наблюдений $\{x_{ij}^{(1)}\}$ ($i = \overline{1; N_1}; j = \overline{1; L}$) и $\{x_{ij}^{(2)}\}$ ($i = \overline{1; N_2}; j = \overline{1; L}$) по этим показателям, где N_1 и N_2 – объемы наблюдений по вектору (X_1, X_2, \dots, X_L) в первом и втором наборах. При этом отдельные наблюдения в каждом из наборов могут отсутствовать. Далее, в общем случае предполагается, что признаки имеют определенные веса, характеризующие важность каждого признака; а именно – признак X_j имеет вес w_j , причем веса пронормированы так, что

$$\sum_{j=1}^L w_j = 1$$
 (то есть каждый весовой коэффициент был разделен на сумму всех коэффициентов). Если веса отсутствуют либо не представляют интереса, то считаем все признаки равнозначными, и полагаем $w_j = 1/L$ для всех $j = \overline{1; L}$. Отметим, что в качестве весовых ко-

эфициентов могут быть использованы пронормированные значения среднего арифметического коэффициентов корреляции (КК) между рассматриваемым признаком и другими признаками совместной (объединенной) совокупности. Задан также определенный уровень значимости α достоверности вывода по гипотезе H_0 ; полагаем $\alpha = 0,95$ – наиболее распространенное в практических приложениях значение уровня значимости.

Предлагаемый алгоритм состоит из следующих этапов:

1. В тех позициях индексов i, j , по которым наблюдения отсутствуют, записываем в матрице недостижимое число. Например, полагаем $x_{ij}^{(1)} = x_{\max} + 1$, либо $x_{ij}^{(2)} = x_{\max} + 1$ (либо другое «очень большое» значение), если соответствующее значение отсутствует, где $x_{\max} = \max_{i,j,k} x_{ij}^{(k)}$, i – порядковый номер большого в имеющемся наборе наблюдений, j – порядковый номер рассматриваемого показателя. Напомним, что все имеющиеся наблюдения неотрицательны, и поэтому записанные значения будут наибольшими значениями в таблице данных. В результате наборы наблюдений не будут иметь пропусков.

2. Для каждого признака X_j составляем единый ранжированный ряд из наборов наблюдений $\{x_{ij}^{(1)}\} (i = \overline{1; N_1})$ и $\{x_{ij}^{(2)}\} (i = \overline{1; N_2})$ сопоставляемых выборок, расставив их элементы по степени нарастания признака и приписав меньшему значению меньший ранг r_{ij} ($i = \overline{1; N}$), где $N = N_1 + N_2$ – общее количество всех наблюдений (и их рангов) в объединенной совокупности.

3. Находим сумму рангов T_{1j} , пришедшихся на долю элементов первой выборки, и отдельно сумму T_{2j} – на долю элементов второй выборки; то есть $T_{1j} = \sum^1 r_{ij}$ и $T_{2j} = \sum^2 r_{ij}$, где индексы «1» и «2» у знаков сумм означают, что суммирование берется только по элементам первой или второй выборок соответственно. Затем определяем большую из двух ранговых сумм $T_{j,\max} = \max(T_{1j}, T_{2j})$, и N_{\max} – объем соответствующей выборки, то есть $N_{\max} = N_1$, если $T_{1j} \geq T_{2j}$, и $N_{\max} = N_2$, если $T_{1j} < T_{2j}$.

4. Определяем значение U-критерия Манна – Уитни по каждому признаку на основе соотношения ($1 \leq j \leq L$):

$$U_j = N_1 \cdot N_2 + N_{\max} \cdot (N_{\max} + 1) / 2 - T_{j,\max}.$$

Затем находим взвешенную сумму

$$U = \sum_{j=1}^L w_j U_j. \quad (1)$$

5. Если объем хотя бы из одной выборок мал (меньше 20), то для нахождения критического уровня $U_{кр}$ при уровне значимости α распределения величины U необходимо составить таблицу значений величины U для данных объемов выборок N_1 и N_2 . Для этого воспользуемся таблицей значений критерия Манна – Уитни для одномерной выборки [2]. Именно, в определенном (например, лексиграфическом) порядке перебираются всевозможные значения величин $\{U_1, U_2, \dots, U_L\}$, и из таблицы одномерных распределений КМИ при каждом наборе значений $\{U_1, U_2, \dots, U_L\}$ вычисляется значение U по формуле (1), а также

произведение соответствующих им вероятностей. Полученное произведение добавляется в качестве слагаемого к вероятности соответствующего значения U . В начальный момент значения всех вероятностей равны 0. Отметим, что в рассматриваемой в работе задаче значения N_1 и N_2 больше 20. Поэтому в рассматриваемой конкретной задаче выбор критического уровня $U_{кр}$ осуществляется на основе процедуры, описываемой ниже в пункте 6.

6. Если оба значения N_1 и N_2 больше 20, то для нахождения $U_{кр}$ можно воспользоваться асимптотическим приближением для U . Именно при справедливости гипотезы H_0 статистика U при достаточно большом объеме выборочных данных $N_1 \geq 20$ и $N_2 \geq 20$ имеет практически нормальное распределение со средним

$$M(U) = \sum_{j=1}^L w_j M(U_j) = \sum_{j=1}^L w_j N_1 \cdot N_2 / 2 = L \cdot N_1 \cdot N_2 / 2$$

и дисперсией

$$D(U) = \sum_{j=1}^L w_j^2 D(U_j) = \sum_{j=1}^L w_j^2 N_1 \cdot N_2 \cdot (N_1 + N_2 + 1) / 12 = \left(\sum_{j=1}^L w_j^2 \right) \cdot N_1 \cdot N_2 \cdot (N_1 + N_2 + 1) / 12.$$

Внутри формул точки как знаки умножения ставить не целесообразно.

В частности, если все $w_j = 1/L$, то $D(U) = N_1 \cdot N_2 \cdot (N_1 + N_2 + 1) / 12$. Отсюда определяем, что для нахождения $U_{кр}$ можно воспользоваться таблицами стандартного нормального распределения. А именно

$$U_{кр} = L \cdot N_1 \cdot N_2 / 2 + U_{кр}^{cm} \cdot \sqrt{\left(\sum_{j=1}^L w_j^2 \right) \cdot N_1 \cdot N_2 \cdot (N_1 + N_2 + 1) / 12},$$

где $U_{кр}^{cm}$ находится из таблиц стандартного нормального распределения как квантиль уровня значимости α .

7. Значение критерия U , полученное в результате обработки имеющихся наблюдений, сравнивается $U_{кр}$. Если полученное значение U меньше или равно $U_{кр}$, то признается наличие существенного различия между уровнем признака в рассматриваемых выборках и принимается гипотеза H_1 . Если же полученное значение U больше $U_{кр}$, то принимается нулевая гипотеза H_0 . Достоверность различий тем выше, чем меньше значение U .

Таким образом, модифицированная процедура проверки гипотезы об однородности двух многомерных выборок, основанная на критерии Манна – Уитни, построена. Наиболее трудоемкой ее частью, требующей дальнейшей детализации, является построение таблиц значений для U – этап 5 процедуры. Однако в задаче, рассматриваемой в данной работе, этап 5 оказывается не востребуемым, поскольку объем имеющийся выборки достаточно большой и позволяет воспользоваться асимптотическими методами, описанными в пункте 6. Поэтому задача формирования конкретного алгоритма формирования таблиц значений для U в работе не рассматривается.

Анализ степени зависимости между признаками и маркерами. Описанная выше процедура была использована для оценки степени зависимости маркеров с признаками – как с отдельными признаками, так и со всей их совокупностью. При этом в качестве весового коэффициента w_j для каждого признака была выбрана величина, пропорциональная КК по Спирмену между этим признаком и рассматриваемым маркером. В данной задаче КК по Спирмену более предпочтителен по сравнению с классическим КК по Пирсону по двум причинам. Во-первых, данные, приведенные в исходных таблицах наблюдений (прежде всего, данные субъективного характера), имеют по ряду признаков значительные погрешности,

что, безусловно, отражается и на величине классического КК. Но, в отличие от классического, КК по Спирмену вычисляется только на основе порядкового расположения оценок без непосредственного учета их абсолютных значений – при искажении оценок их порядок обычно не изменяется.

Во-вторых, в отличие от классического коэффициента, КК по Спирмену при линейном преобразовании признака может изменять свои значения. Указанное свойство в контексте данной работы важно при учете пола пациента (в работе 0 соответствует женскому полу, а 1 – мужскому): классический КК выдает одинаковые значения как для признака «пол мужской», так и для признака «пол женский», поскольку они линейно связаны – их сумма равна 1. В то же время КК по Спирмену для мужского и женского пола различен, что, по-видимому, точнее передает содержательное понимание вопроса взаимосвязи маркера с полом пациента.

Отметим, однако, при больших объемах выборок значения классического КК и КК по Спирмену отличаются в среднем на 5 %, что укладывается в рамки погрешности исходных данных. То есть различия между значениями КК по Пирсону и Спирмену могут быть объяснены только погрешностями в исходных данных, и поэтому эти коэффициенты могут считаться равными (с точностью до погрешностей в исходных данных).

Результаты оценок рассмотренных выше КК по Спирмену приведены в табл. 1, где в первом столбце оставлены только те маркеры, для которых выполняется предположение о зависимости этих маркеров со всей совокупностью признаков при 5%-ом уровне значимости – таких маркеров оказалось восемь (из 26), они перечислены в первом столбце табл. 1. При этом каждый из этих маркеров оказался существенно (то есть имеет большое абсолютное значение коэффициента связи признака с маркером) взаимосвязан с большим числом признаков: число таких признаков для каждого из перечисленных в таблице маркеров соответственно равно 36, 45, 57, 57, 79, 17, 52, 61. В рамках данной статьи невозможно полностью перечислить все признаки для каждого маркера. Поэтому в таблице для каждого маркера приведены только пять наиболее значимых признаков, то есть имеющих наибольшее значение коэффициента связи признака с маркером. Во втором столбце приведены оценки значимости каждого маркера для 5%-ого уровня; в третьем столбце перечислены указанные пять признаков, а в последнем (четвертом) – оценки их значимости.

Таблица 1

Оценки значимости маркеров и наиболее важных признаков

№	МАРКЕР	Число достоверно связанных с маркером признаков (N)	Коэффициент связи набора признаков с маркером (1,645 – табличный квантиль 95 %)	Примеры наиболее важных, статистически значимых факторов	Оценка фактора (Коэффициент связи признака с маркером)
1	Chlamydo- phylla pneumonia (ChP) IgA	40	1,77	Гемоглобин (г/л)	3,41
				Билирубин общий (мкМоль/л)	-4,459
				Альбумин (г/л)	3,809
				Альбумин-глобулиновый индекс	3,6

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии № 4 (28) 2014
ОБРАБОТКА СИГНАЛОВ И ДАННЫХ, РАСПОЗНАВАНИЕ ОБРАЗОВ,
ВЫЯВЛЕНИЕ ЗАКОНОМЕРНОСТЕЙ И ПРОГНОЗИРОВАНИЕ

				УЗИ. Желчные протоки (норма – 0, уплотнены – 1, иное – 2)	–9,32
2	Herpes simplex virus I (HSV1) IgM	44	1,953	Фибриноген А(г/л)	9,681
				Фибриноген В (0 – отрицателен, 1 – положителен)	–8,167
				Диаметр v.lienalis (мм)	–9,598
				УЗИ. Структура поджелудочной железы (0 – норма, 1 – не норма)	5,245
				ФГС желудка (1-2-3-4: гастрит – гастродуоденит – эрозии – язвы)	5,321
3	Cytomagalovirus (CMV) IgM	57	2,059	Цирроз печени (диагноз)	–5,134
				Фибриноген В (0 – отрицателен, 1 – положителен)	–9,078
				УЗИ. Толщина стенки желчного пузыря (мм)	–5,558
				Диаметр v.lien (мм)	–10,664
				УЗИ. Структура поджелудочной железы (0 – норма, 1 – не норма)	4,685
4	Суммарные антитела к E.Coli (норма 14,1±0,42 мкг/мл)	57	1,69	Цирроз печени (диагноз)	–4,09
				Гепатомегалия, определенная физикально	–4,975
				Мочевина (ммоль/л)	–5,455
				УЗИ. Толщина стенки желчного пузыря (мм)	–7,679
				ФГС желудка (1-2-3-4: гастрит – гастродуоденит – эрозии – язвы)	8,194
5	Суммарные антитела к Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	79	1,933	Асцит	–26,87
				Гепатомегалия, определенная физикально	–29
				Лейкоциты крови(*10 ¹²)	–29,217
				Тимоловая проба (норма 0–4 ед.)	–36,77
				Диаметр v.porta (мм)	–20,367
6	Суммарные антитела к Proteus spp. (норма 11,53±0,56 мкг/мл)	17	1,826	Гемоглобин (г/л)	–4,045
				Палочкоядерные лейкоциты (%)	2,576
				АСТ (мМоль/ч*л)	–1,915
				Мочевина (ммоль/л)	–16,102
				УЗИ. Длина селезенки (мм)	–5,528
7	Суммарные антитела к Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	52	1,67	Лейкоциты крови (*10 ⁹ /л)	–4,54
				Билирубин прямой (мкМоль/л)	–3,667
				Глюкоза крови (мМоль/л)	–5,745
				Бета-липопротеиды (ед)	–3,917
				Структура печени (0-1-2-3-4: однородная – неоднородная – диффузно-неоднородная – резко неоднородная – иное)	–4,053
				Диаметр v.lienalis (мм)	–3,797
				ФГС желудка (1-2-3-4: гастрит – гастродуоденит – эрозии – язвы)	–3,173
8	Суммарные антитела к Candida spp. (норма 24,7±0,72 мкг/мл)	61	1,717	Возраст	–4,621
				Цирроз печени (диагноз)	–4,123
				Альбумин-Глобулиновый индекс	–4,155
				УЗИ Толщина стенки желчного пузыря (мм)	–5,172
				ФГС желудка (1-2-3-4: гастрит – гастродуоденит – эрозии – язвы)	–4,25

Табл. 1 позволяет сделать следующие выводы:

1) заключения по всем приведенным маркерам, за исключением Протея, опираются на достаточно большой объем данных – более 40 наблюдений, что позволяет говорить о высокой надежности результатов, приведенных в табл. 1;

2) состав признаков по разным маркерам достаточно разнороден и разнообразен – на 40 позициях в четвертом столбце перечислены (с некоторыми повторениями) 26 признаков – это указывает на комплексный, многообразный характер зависимостей между маркерами и признаками; более того, оценка значений перечисленных 26 признаков позволит выявить наличие или отсутствие каждого из перечисленных восьми значимых маркеров;

3) подавляющая доля связей между маркерами и признаками (столбец 5) отрицательна, то есть увеличение количества маркера сопровождается уменьшением количественной оценки признака, и наоборот;

4) перечисленные в столбце 4 признаки относятся к числу общепринятых методов, включенных в Стандарты диагностики ХГ и ЦП, что обеспечивает практическую возможность реализации процедур выявления маркеров на основе измерения или оценки признаков.

Выше на основе корреляционного анализа были выявлены наиболее значимые маркеры и проведен анализ их зависимости от различных признаков – основные результаты анализа приведены в табл. 1. Представляет интерес также проведение дополняющего анализа, описывающего влияние каждого из исходных признаков на выделенные выше восемь наиболее значимых маркеров.

Для удобства проведения указанного анализа полученные ранее результаты исследования корреляционных зависимостей представлены в виде перечисления признаков, каждому из которых был сопоставлен список значимых маркеров, с которыми у рассматриваемого признака корреляционная связь существенна. Оказалось, что почти все признаки существенно связаны хотя бы с одним восьми значимых маркером. При этом 6 признаков существенным образом связаны только с одним из значимых маркеров, 16 – с двумя маркерами, 24 – с тремя, 16 – с четырьмя, 11 – с пятью, 18 – с шестью, 10 – с семью и 2 – с восемью маркерами. Таким образом, с точки зрения информативности признаков, то есть возможности выявления как можно большего числа маркеров при измерении признака, имеется 2 признака с максимально возможной информативностью и 10 – с информативностью, выявляющей 7 (то есть 87,5 %) маркеров. Указанные наиболее информативные признаки, ассоциированные с 7-ым и 8-ым маркерами, перечислены в табл. 2. Здесь в первом столбце указаны порядковые номера, во втором приведены наименования признаков, в третьем перечислены маркеры, выявляемые на основе данного признака, а в последнем (четвертом) – коэффициенты значимости для каждой пары «признак – маркер».

Таблица 2

Значимое влияние каждого из признаков на выделенные маркеры

№	Признаки	Значимые маркеры (антитела к соответствующему микроорганизму)	Степень зависимости
1	Гемоглобин (г/л)	Chlamydomyella pneumonia (ChP) IgA	3,41
		Herpes simplex virus I (HSV1) IgM	-3,281
		Cytomagalovirus (CMV) IgM	-3,969
		E. Coli (норма 14,1±0,42 мкг/мл)	-2,381
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-8,378
		Protacis spp. (норма 11,53±0,56 мкг/мл)	-4,045
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-3,809
		Candida spp. (норма 24,7±0,72 мкг/мл)	-3,006
2	Эритроциты (*10 ¹² /л)	Chlamydomyella pneumoniae (ChP) IgA	3,307
		Herpes simplex virus I (HSV1) IgM	-3,5

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии № 4 (28) 2014
ОБРАБОТКА СИГНАЛОВ И ДАННЫХ, РАСПОЗНАВАНИЕ ОБРАЗОВ,
ВЫЯВЛЕНИЕ ЗАКОНОМЕРНОСТЕЙ И ПРОГНОЗИРОВАНИЕ

		Cytomagalovirus (CMV) IgM	-4,146
		E. Coli (норма 14,1±0,42 мкг/мл)	-2,343
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-9,238
		Protaeus spp. (норма 11,53±0,56 мкг/мл)	-3,835
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-3,464
		Candida spp. (норма 24,7±0,72 мкг/мл)	-3,429
3	Палочкоядерные лейкоциты в общем анализе крови (%)	Chlamydohylla pneumoniae (ChP) IgA	-2,592
		Herpes simplex virus I (HSV1) IgM	1,995
		Cytomagalovirus (CMV) IgM	2,113
		E. Coli (норма 14,1±0,42 мкг/мл)	2,207
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-37,333
		Protaeus spp. (норма 11,53±0,56 мкг/мл)	2,576
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-1,739
4	Тимоловая проба (норма 0-4 ед)	Chlamydohylla pneumoniae (ChP) IgA	-2,373
		Herpes simplex virus I (HSV1) IgM	3,588
		Cytomagalovirus (CMV) IgM	2,88
		E. Coli (норма 14,1±0,42 мкг/мл)	2,089
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-36,77
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-2,008
		Candida spp. (норма 24,7±0,72 мкг/мл)	-1,903
5	Альбумин сыворотки крови (г.л)	Chlamydohylla pneumoniae (ChP) IgA	3,809
		Herpes simplex virus I (HSV1) IgM	-3,002
		Cytomagalovirus (CMV) IgM	-2,953
		E. Coli (норма 14,1±0,42 мкг/мл)	-2,057
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-7,307
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-3,225
		Candida spp. (норма 24,7±0,72 мкг/мл)	-3,287
6	Альбумин-глобулиновый индекс	Chlamydohylla pneumoniae (ChP) IgA	3,6
		Herpes simplex virus I (HSV1) IgM	-3,741
		Cytomagalovirus (CMV) IgM	-3,327
		E. Coli (норма 14,1±0,42 мкг/мл)	-2,109
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-8,721
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-1,985
		Candida spp. (норма 24,7±0,72 мкг/мл)	-4,155
7	Мочевина (ммоль/л)	Herpes simplex virus I (HSV1) IgM	-13,16
		Cytomagalovirus (CMV) IgM	-12,89
		E. Coli (норма 14,1±0,42 мкг/мл)	-5,455
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-6,928
		Protaeus spp. (норма 11,53±0,56 мкг/мл)	-16,102
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-9,295
		Candida spp. (норма 24,7±0,72 мкг/мл)	-21,213
8	УЗИ –толщина стенок желчного пузыря (мм)	Chlamydohylla pneumoniae (ChP) IgA	-2,7
		Herpes simplex virus I (HSV1) IgM	-5,216
		Cytomagalovirus (CMV) IgM	-5,558
		E. Coli (норма 14,1±0,42 мкг/мл)	-7,679
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-15,81
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-3,905
		Candida spp. (норма 24,7±0,72 мкг/мл)	-5,172
9	Диаметр v. porta (мм)	Chlamydohylla pneumoniae (ChP) IgA	-2,202
		Herpes simplex virus I (HSV1) IgM	-4,233

		Cytomagalovirus (CMV) IgM	-4,769
		E.Coli (норма 14,1±0,42 мкг/мл)	-3,89
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-20,367
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-1,69
		Candida spp. (норма 24,7±0,72 мкг/мл)	-2,545
10	УЗИ – поперечник селезенки (мм)	Chlamydophylla pneumonia (ChP) IgA	-1,849
		Herpes simplex virus I (HSV1) IgM	-2,646
		Cytomagalovirus (CMV) IgM	-3,662
		E.Coli (норма 14,1±0,42 мкг/мл)	-2,113
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-19,616
		Protaceus spp. (норма 11,53±0,56 мкг/мл)	-5,528
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-2,662
11	УЗИ хвост поджелудочной железы (норма до 25 мм)	Chlamydophylla pneumonia (ChP) IgA	-2,197
		Herpes simplex virus I (HSV1) IgM	-4,374
		Cytomagalovirus (CMV) IgM	-4,463
		E.Coli (норма 14,1±0,42 мкг/мл)	-16,296
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-3
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-12,006
		Candida spp. (норма 24,7±0,72 мкг/мл)	-8,083
12	ФГС – исследование пищевода (нет патологии – 0, эзофагит – 1, эрозии – 2, язвы – 3)	Chlamydophylla pneumonia (ChP) IgA	2,792
		Herpes simplex virus I (HSV1) IgM	4,632
		Cytomagalovirus (CMV) IgM	2,905
		E.Coli (норма 14,1±0,42 мкг/мл)	7,584
		Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	-7,833
		Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	-1,8
		Candida spp. (норма 24,7±0,72 мкг/мл)	-2,125

Из табл. 2 также следует, что наиболее стабильно по перечисленным признакам распознается маркер *Pseudomonas spp*, хотя по отдельным признакам есть и более значимые маркеры. Далее, подавляющая доля зависимостей (75 из 90–83 %) имеет отрицательный характер, что подтверждает сделанный выше вывод по результатам анализа табл. 1.

Представляет интерес также задача оценки степени зависимости между значимыми маркерами, для чего был проведен соответствующий корреляционный анализ зависимостей между маркерами, результаты которого приведены в табл. 3. Здесь К.П. – значение КК по Пирсону, Зн. – оценка значимости КК (p), N – объем выборки, на основе которой выполнены расчеты и сделаны выводы.

По результатам расчетов, приведенных в табл. 3, заключаем, что степень зависимости между маркерами изменяется в широких пределах: между 8-ым и 7-ым маркерами КК меньше 6,5 %, в то время как между 4-ым и 5-ым маркерами он превышает 87 %. Уровень достоверности приведенных в таблице результатов в целом достаточно высокий, объемы выборок, на которые опираются расчеты большие – не меньше 35 наблюдений. Однако часть выводов опираются на выборки объемом 12 наблюдений – практически минимально приемлемый объем для получения достоверных результатов. Была проведена оценка статистической значимости полученных КК с использованием критерия χ^2 при 5 % уровне значимости. Результаты расчетов доказали значимость полученных оценок КК.

Таблица 3

Анализ корреляционных зависимостей между маркерами

№	Наименование маркера		Chlamydophylla pneumoniae (ChP) IgA	Herpes simplex virus I (HSV1) IgM	Cytomagalovirus (CMV) IgM	Pseudomonas spp. (норма 14,2±1,4мкг/мл)	Protacaeus spp. (норма 11,53±0,56 мкг/мл)	Klebsiella pneumoniae (норма 24,9±0,88мкг/мл)	E.Coli (норма 14,1±0,42 мкг/мл)	Candida spp. (норма 24,7±0,72 мкг/мл)							
			К.П.	Зн.	N	К.П.	Зн.	N	К.П.	Зн.	N	К.П.	Зн.	N	К.П.	Зн.	N
1	Chlamydophylla pneumoniae (ChP) IgA	К.П.	1	0,417	0,453	–	0,131	–0,177	0,321	–0,164							
		Зн.	–	<0,001	<0,001	<0,001	0,454	0,310	0,060	0,347							
		N	80	80	80	12	35	35	35	35							
2	Herpes simplex virus I (HSV1) IgM	К.П.	0,417	1	0,693	0,259	0,537	0,052	0,590	–0,298							
		Зн.	<0,001	–	<0,001	0,416	0,001	0,768	<0,001	0,082							
		N	80	80	80	12	35	35	35	35							
3	Cytomagalovirus (CMV) IgM	К.П.	0,453	0,693	1	0,180	0,562	–0,075	0,493	–0,307							
		Зн.	0	0	–	0,576	0	0,667	0,003	0,073							
		N	80	80	80	12	35	35	35	35							
4	Pseudomonas spp. (норма 14,2±1,4 мкг/мл)	К.П.	0	0,259	0,180	1	–0,052	0,202	–0,082	–0,115							
		Зн.	–	0,416	0,576	–	0,872	0,528	0,800	0,722							
		N	12	12	12	12	12	12	12	12							
5	Protacaeus spp. (норма 11,53±0,56 мкг/мл)	К.П.	0,131	0,537	0,562	–0,052	1	0,072	0,572	–0,238							
		Зн.	0,454	0,001	<0,001	0,872	–	0,682	<0,001	0,168							
		N	35	35	35	12	35	35	35	35							
6	Klebsiella pneumoniae (норма 24,9±0,88 мкг/мл)	К.П.	–0,177	0,052	–0,075	0,202	0,072	1	0,145	0,356							
		Зн.	0,310	0,768	0,667	0,528	0,682	–	0,405	0,036							
		N	35	35	35	12	35	35	35	35							
7	E.Coli (норма 14,1±0,42 мкг/мл)	К.П.	0,321	0,590	0,493	–0,082	0,572	0,145	1	–0,064							
		Зн.	0,060	<0,001	0,003	0,800	<0,001	0,405	–	0,713							
		N	35	35	35	12	35	35	35	35							
8	Candida spp. (норма 24,7±0,72 мкг/мл)	К.П.	–0,164	–0,298	–0,307	–0,115	–0,238	0,356	–0,064	1							
		Зн.	0,347	0,082	0,073	0,722	0,168	0,036	0,713	–							
		N	35	35	35	12	35	35	35	35							

Отметим дополнительно, что авторами в качестве альтернативы были проведены расчеты для случая, когда полуколичественные оценки признаков, опирающиеся на отнесение признака к одному из заданных уровневых значений (например, 0 – 1 – 2 – 3 – 4) были заменены на качественные (бинарные) оценки – наличие или отсутствие признака. Полученные в результате коэффициенты отличались от исходных не более чем на 1,5–2 %. Указанный результат позволяет сделать вывод о возможности замены количественных или полуколичественных оценок многих признаков на их качественные оценки. Последние получить значительно проще, чем количественные и полуколичественные оценки тех же признаков.

Заключение. В работе получены следующие результаты.

1. Исходная задача оценки влияния вирусно-бактериальных маркеров на клиническую картину хронических диффузных заболеваний печени сведена к задаче в области ма-

тематической статистики, именно к задаче проверке гипотезы на однородность двух многомерных выборок.

2. Предложена алгоритмическая процедура проверки гипотезы на однородность двух многомерных выборок, обобщающая метод Манна – Уитни, предназначенный для проверки однородности одномерных выборок.

3. Описанная процедура проверки гипотезы об однородности использована для анализа указанных в работе данных по ХДЗП. Проведен анализ полученных результатов, сформулированы выводы и рекомендации.

4. Полученные в работы результаты применительно к диффузным заболеваниям печени позволяют сделать вывод о практической эффективности разработанных в работе процедур выявления зависимостей заболеваниями и вирусно-бактериальными маркерами. Разработанные методы могут быть использованы при анализе и других заболеваний для повышения качества диагностических решений.

Список литературы

1. Большев Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – 3-е изд. – Москва : Наука, 1983. – 474 с.
2. Брумштейн Ю. М. Одно- и многомерные временные ряды: анализ возможных методов оптимизации отсчетов и оценки характеристик / Ю. М. Брумштейн, М. В. Иванова // Прикаспийский журнал: управление и высокие технологии. – 2012. – Т. 20, № 4. – С. 35–45.
3. Гаек Я. Теория ранговых критериев : пер. с англ. / Я. Гаек, З. Шидак. – Москва : Наука, 1971. – 376 с.
4. Дедов А. В. Клинико-патогенетическое значение антигенов HLA и бета-2 микроглобулина при хронических диффузных заболеваниях печени вирусной этиологии : автореф. дис. ... канд. мед. наук / А. В. Дедов. – Астрахань, 1995. – 24 с.
5. Дедов А. В. Построение математических моделей предварительного диагноза заболеваний печени на основе методов регрессионного анализа / А. В. Дедов, Г. А. Попов // Вестник Астр. техн. ун-та. Серия: Управление, вычислительная техника и информатика. – 2014. – № 4. – С. 124–136.
6. Леонов В. П. Статистика в кардиологии. 15 лет спустя / В. П. Леонов // Медицинские технологии. Оценка и выбор. – 2014. – № 1. – С. 17–28.
7. Маркин А. В. Метод автоматического восстановления значений в потоках данных на основе взвешенной модели / А. В. Маркин, М. В. Щербаков // Прикаспийский журнал: управление и высокие технологии. – 2013. – № 3. – С. 49–54.
8. Медик В. А. Руководство по статистике в медицине и биологии : в 2 т. / В. А. Медик, Б. Б. Фишман, М. С. Токмачев ; под ред Ю. М. Комарова. – Москва : Медицина, 2001. – Т. 2. Прикладная статистика здоровья. – 352 с.
9. Реброва О. Ю. Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTIKA / О. Ю. Реброва. – Москва : МедиаСфера, 2002. – 312 с.
10. Сердюкова Т. В. Мониторинг общей заболеваемости органов пищеварения в Астраханской области по отдельным нозологическим формам с 2006–2010 гг. / Т. В. Сердюкова, Н. Н. Курьянова, М. А. Сердюков // Астраханский медицинский журнал. – 2012. – Т. 7, № 2. – С. 116–119.
11. Холлендер М. Методы непараметрической статистики : пер. с англ. / М. Холлендер, Д. А. Вулф. – Москва : Финансы и статистика, 1983. – 518 с.
12. Bolukbas F. F. High rate of seropositivity of Chlamydia pneumoniae IgA in male patients with nonalcoholic steatohepatitis / F. F. Bolukbas, C. Bolukbas, F. Zeyrek, M. Aslan, H. I. Bahcecioglu, I. Ozardali // Dig Dis Sci. – 2005. – Vol. 50, № 6. – P. 1141–1145.
13. Bruijne de M. H. J. On the use of Cox regression in the presence of an irregularly observed time-dependent covariate / Bruijne de M. H. J., le Cessie S., Kluine-Nelemas H. C., van Houwelingen H. C. // Statistic in medicine. – 2001. – Vol. 20, № 24. – P. 3817–3829.
14. Cisneros-Herreros J. M. Hepatitis due to herpes group viruses / J. M. Cisneros-Herreros, M. Herrero-Romero // Enferm. Infecc. Microbiol. Clin. – 2006. – Vol. 24, № 6. – P. 392–397.

15. Marshall G. Prospective prediction in the presence of missing data / G. Marshall, B. Warner, S. MaWhinney, K. Hammermeister // *Statistics in Medicine*. – 2002. – Vol. 21, № 4. – P. 561–570.

16. Montaña-Loza A. Seropositivity for Chlamydia pneumoniae in patients with primary biliary cirrhosis [Article in Spanish] / A. Montaña-Loza, E. Vázquez-Ballesteros // *Gastroenterol. Hepatol.* – Vol. 29, № 3. – P. 113–117.

References

1. Bolshev L. N., Smirnov N. V. *Tablitsy matematicheskoy statistiki* [The tables of mathematical statistics], 3rd ed. Moscow, Nauka, 1983. 474 p.

2. Brumshteyn Yu. M., Ivanova M. V. Odno- i mnogomernye vremennyye ryady: analiz vozmozhnykh metodov optimizatsii otschetov i otsenki kharakteristik [Mono- and multidimensional timely series: The analysis of possible methods of optimization of calculations and evaluation of characteristics]. *Prikaspiyskiy zhurnal: upravlenie i upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2012, vol. 20, no. 4, pp. 35–45.

3. Gack Ya., Shidak Z. *Teoriya rangovykh kriteriev* [Range criteria theory]. Moscow, Nauka, 1971. 376 p.

4. Dedov A. V. *Kliniko-patogeneticheskoe znachenie antigenov HLA i beta-2 microglobulina pri khronicheskikh diffuznykh zabolovaniyakh pecheni* [Clinical and pathogenetic role of HLA antigens and beta-2 microglobulin in chronic diffuse liver diseases]. Astrakhan, 1995. 24 p.

5. Dedov A. V., Popov G. A. Formirovanie modeley predvaritelnogo diagnoza zabolovaniy pecheni na osnove metodov regressionnogo analiza [The forming of the preliminary diagnosis models of liver diseases on the basis of regression analysis methods]. *Vestnik Astrakhanskogo tekhnicheskogo universiteta. Seriya: upravlenie, vychislitel'naya tekhnika i informatika* [Bulletin of Astrakhan Technical University], 2014, no. 4, pp. 124–136.

6. Leonov V. P. *Statistika v kardiologii. 15 let spustya* [Statistics in cardiology. 15 years later]. *Meditsinskie tekhnologii. Otsenka i vybor* [Medical Technologies. Evaluation and Choice], 2014, no. 1, pp. 17–28.

7. Markin A. V., Shcherbakov M. V. Metod avtomaticheskogo vosstanovleniya znacheniy v potokakh dannykh na osnove vzveshennoy modeli [The method of automatical restoration of the values in the data flows on the basis of weighted model]. *Prikaspiyskiy zhurnal: upravlenie i upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2013, no. 3, pp. 49–54.

8. Medik V. A., Fishman B. B., Tokmachev M. S. *Rukovodstvo po statistike v meditsine i biologii. T. 2. Prikladnaya statistika zdorovya* [The manual for the statistics in the medicine and biology. Vol. 2. The applied statistics of the health], in 2 vol. Moscow, Meditsina, 2001. 352 p.

9. Rebrova O. Yu. *Statisticheskii analiz meditsinskikh dannykh. Primenenie paketa prikladnykh programm STATISTIKA* [Statistical analysis of medical data. The usage of applied packs of programs «STATISTIKA»]. Moscow, Mediasfera, 2002. 312 p.

10. Serdyukova T. V., Kuryanova N. N., Serdyukov M. A. Monitoring obshhey zabolvaemosti organov pishhevareniya v Astrahanskoy oblasti po otdelnym nozologicheskim formam s 2006–2010 gg [The monitoring of total morbidity of digestive organs diseases according to the specific nosologic forms in 2006–2010]. *Astrahanskiy meditsinskij zhurnal* [Astrakhan Medical Journal], 2012, vol. 7, no. 2, pp. 116–119.

11. Hollender M., Vulf D. A. *Metody neparametricheskoy statistiki* [The methods of non-parametric statistics]. Moscow, Finansy i statistika, 1983. 518 p.

12. Bolukbas F. F., Bolukbas C., Zeyrek F., Aslan M., Bahcecioglu H. I., Ozardali I. // High rate of seropositivity of Chlamydia pneumoniae IgA in male patients with nonalcoholic steatohepatitis. *Dig Dis Sci.*, 2005, vol. 50, no. 6, pp. 1141–1145.

13. Bruijine de M. H. J., le Cessie S., Kluine-Nelemas H. C., van Houwelingen H. C. On the use of Cox regression in the presence of an irregularly observed time-dependent covariate. *Statistic in Medicine*, 2001, vol. 20, no. 24, pp. 3817–3829.

14. Cisneros-Herreros J. M., Herrero-Romero M. Hepatitis due to herpes group viruses. *Enferm. Infecc. Microbiol. Clin.*, 2006, vol. 24, no. 6, pp. 392–397.

15. Marshall G., Warner B., MaWhinney S., Hammermeister K. Prospective prediction in the presence of missing data. *Statistics in Medicine*, 2002, vol. 21., no. 4, pp. 561–570.

16. Montaña-Loza A., Vázquez-Ballesteros E. Seropositivity for Chlamydia pneumoniae in patients with primary biliary cirrhosis [Article in Spanish]. *Gastroenterol. Hepatol.*, vol. 29, no. 3, pp. 113–117.