

МЕДИЦИНА И ВЕТЕРИНАРИЯ

УДК 519.22

ИДЕНТИФИКАЦИЯ СОСТОЯНИЯ БОЛЬНЫХ ЛЕПТОСПИРОЗОМ ПОСРЕДСТВОМ ФАКТОРНОГО И ДИСКРИМИНАНТНОГО АНАЛИЗА

© 2008 г. А.А. Халафян, В.Н. Городин, Д.В. Савенко

At first time secret factors were uncovered by means of factor analysis, fully defining the state leptospirosis patients. They investigated structure of correlation between important index in total sufficiently describing state patients. They wrote program, defining degree of heaviness leptospirosis patients for different stage flow illness.

Главными целями факторного анализа являются сокращение числа переменных (редукция данных) и оп-ределение структуры взаимосвязей между переменными, т.е. классификация переменных [1]. Сокращение достигается путем выделения скрытых общих факторов, объясняющих связи между наблюдаемыми признаками (переменными) объекта, т.е. вместо исходного набора переменных появится возможность анализировать данные по выделенным факторам, число которых значительно меньше исходного числа переменных.

Лептоспироз – острая инфекционная болезнь. Полиморфизм клинической симптоматики, поражение жизненно важных органов являются причиной высокого процента летальности. Учитывая большой процент летальных исходов, актуальным является своевременная и адекватная оценка тяжести течения заболевания. Задача усложняется из-за большого количества показателей, характеризующих состояние больного, – свыше 50. Поэтому целесообразно применение факторного анализа для выявления скрытых показателей, описывающих состояние больного. Исследования проводились с использованием данных, состоящих из результатов наблюдений над больными безжелтушным лептоспирозом, находившимися на лечении в клинической инфекционной больнице г. Краснодара за период 1975–1996 г. Характерной особенностью лептоспироза является высокая динамичность изменения состояния больных в течение первых 3 недель заболевания. Поэтому анализ данных был проведен программой STATISTICA 6 по базам данных больных 1-й (404 чел.), 2-й (301 чел.) и 3-й (144 чел.) недель заболевания. Для каждого больного, кроме клинико-лабораторных показателей, известна форма течения заболевания – легкая, среднетяжелая, тяжелая.

По критерию Кэттеля было определено число факторов, которое надо выделить, – 2. В таблице приведены факторные нагрузки при числе факторов, равном 2, и выбранных для дальнейшего анализа переменных, для которых факторные нагрузки по какому-либо фактору больше, чем 0,7.

Из таблицы следует, что группе переменных: креатинин плазмы, клубочковая фильтрация (КФ), концентрационный индекс, канальцевая реабсорбция (КР), среднемолекулярные пептиды (СМП), индекс интоксикации Гриневы (ИИ Гриневы), эффективная концентрация альбумина (ЭКА) соответствуют значительно большие факторные нагрузки (сильные корреляции)

для фактора 1 и меньшие для фактора 2. И, наоборот, для группы переменных α -, β -, γ -глобулины соответствуют значительно меньшие факторные нагрузки (слабые корреляции) для фактора 2 и большие для фактора 1, что является подтверждением правильного выбора числа факторов.

Результаты факторного анализа

Переменные	Факторные нагрузки.	
	Фактор 1	Фактор 2
Креатинин плазмы	-0,924478	0,011084
КФ	0,861810	0,073840
Конц. индекс	0,904381	0,104031
КР	0,888109	0,112869
СМП	-0,869751	0,006675
ИИ Гринева	-0,847122	-0,056597
ЭКА	0,815255	-0,072418
α -глобулины	-0,162932	0,918989
β -глобулины	-0,003877	0,954917
γ -глобулины	-0,082329	0,939858
Общ. дис.	5,376082	2,677367
Доля общ.	0,537608	0,267737

Данные результаты были получены методом главных компонент.

По линейным уравнениям регрессии программа вычисляет значения факторов для больных. Анализ формы течения болезни показал, что по значениям факторов для каждого больного можно идентифицировать состояние больного, т.е. вместо 53 показателей, характеризующих состояние больного, можно использовать 2 показателя – фактор 1 и фактор 2, вычисляемых по значениям 9 показателей, характеризующих болезненное состояние.

На первой неделе заболевания ведущим клиническим синдромом является синдром интоксикации, его лабораторное отражение – увеличение пула (содержания) СМП, повышение ИИ Гриневы, снижение ЭКА. Выраженная интоксикация и гипотензия (снижение артериального давления) разной выраженности, вплоть до шока, приводят к органной дисфункции и недостаточности. При этом признаками дисфункции почек являются увеличение концентрации креатинина в сыворотке крови, снижение КФ и КР, что влечет за собой и изменение концентрационного индекса. Имеющаяся диспротеинемия (гипоальбуминемия и изменение со-

держания белковых фракций – α -, β -, γ -глобулины) свидетельствует о наличии острофазового ответа.

Поэтому фактор 1 назовем синдромом интоксикации и органной дисфункции, фактор 2 – синдромом острофазового ответа.

Факторный анализ данных по 2-й неделе течения заболевания позволил выделить 2 фактора, описывающих состояние больных лептоспирозом: фактор 1 – назовем его синдромом эндогенной интоксикации и органной дисфункции, состоящий из переменных: креатинин плазмы, КФ, СМП, билирубин, мочевины, ИИ Гринева, ЭКА, РСА (резерв связывания альбумина), индекс (ЭКА/ОКА) и фактор 2 – назовем его синдромом воспалительной реакции периферической крови, состоящий из переменных: сегментоядерные нейтрофилы, лимфоциты, лейкоцитарный индекс (ЛИ), кровяно-клеточный показатель (ККП).

Факторный анализ данных по 3-й неделе течения заболевания позволил выделить 2 фактора, описывающих состояние больных лептоспирозом: фактор 1 – назовем его синдромом органных и метаболических нарушений, состоящий из переменных: КФ, концентрационный индекс, КР, Na (натрий), К (калий), Cl (хлор), билирубин, мочевины, триглицериды, β -липопротеиды, и фактор 2 – назовем его синдромом диспротеинемии, состоящий из переменных: альбумин, ЭКА, РСА, индекс ЭКА/ОКА, связывающая способность плазмы (ССП).

Таким образом, факторный анализ позволил для первых трех недель заболевания выявить скрытые факторы, наиболее полно характеризующие состояние больных; показал, что для каждой недели характерны свои показатели, описывающие состояние больных.

Для описания состояния больных в файлах данных присутствует переменная – форма течения болезни: легкая, среднетяжелая, тяжелая. После выделения значимых показателей можно говорить о компьютеризации классификации состояния больных по сочетанию возможных значений этих показателей. Для решения этой задачи был использован дискриминантный анализ.

Для больных 1-й недели из переменных, составляющих факторы 1 и 2, дискриминантным анализом были выделены наиболее информативные показатели, по которым можно успешно классифицировать форму тяжести течения заболевания: КФ, конц. индекс, КР, СМП, ИИ Гринева, ЭКА. Все показатели принадлежат фактору 1. Процедура дискриминации больных по формам течения заболевания завершилась успешно – значение лямбды Уилкса равно 0,096 и близко к 0, из 404 больных программа неверно классифицировала только 11, общий процент правильной классификации составил 97,2.

Из диаграммы рассеяния канонических корней (рисунки) видно, что расстояние между группами больных легкой и среднетяжелой формами заболевания значительно меньше, чем расстояние между группами больных легкой и тяжелой формами. Точки, соответствующие больным с легкой формой, соприкасаются с точками, соответствующими больным группы *среднетяжелая*, которые в свою очередь соприкасаются с точками, соответствующими больным группы *тяжелая*.

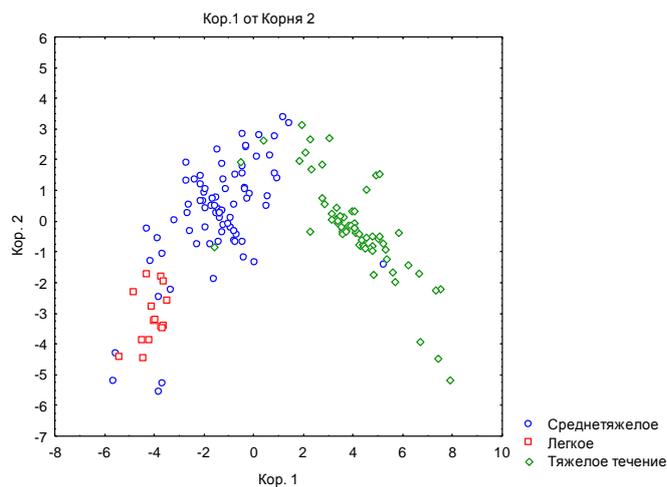


Диаграмма рассеяния канонических корней

Для больных лептоспирозом на 2-й неделе заболевания из переменных, составляющих факторы 1 и 2, дискриминантным анализом были выделены только 2 показателя, по которым можно успешно классифицировать форму течения болезни, – КФ, ИИ Гринева. Оба показателя принадлежат фактору 1. Процедура дискриминации больных по формам течения заболевания завершилась менее успешно – значение лямбды Уилкса равно 0,16. Из 301 больного программа неверно классифицировала 12, общий процент правильной классификации составил 96.

Для больных на 3-й неделе заболевания из переменных, составляющих факторы 1 и 2, дискриминантным анализом были выделены показатели, по которым можно успешно классифицировать форму тяжести течения заболевания, – КФ, К, Cl, триглицериды, β -липопротеиды, альбумин. Они также принадлежат фактору 1. Процедура дискриминации больных по формам течения заболевания завершилась более успешно, чем для больных первых двух недель, – значение лямбды Уилкса равно 0,0029 и близко к 0. Из 143 больных программа верно классифицировала всех больных.

На основании приведенных результатов исследований можно сделать вывод о возможности применения дискриминантного анализа для определения, и, что особенно важно, прогнозирования тяжести течения заболевания [2].

Для возможности применения метода в практической медицине были записаны Analysis Macros – макросы модуля дискриминантный анализ. На основе списков больных – баз данных для каждой недели заболевания, макросов и написанных на языке Visual Basic интерфейса была разработана программа верификации степени тяжести больных. Интерфейс программы прост и не требует специального обучения для работы с ними [3].

Предлагаемый способ оценки тяжести состояния больных позволит снизить частоту ошибок при верификации степени тяжести заболевания; определить промежуточные состояния больного и своевременно вносить коррективы в проводимой терапии; своевременно выработать оптимальные для каждого больного стратегию и тактику лечения.

Литература

1. Халафян А.А. Статистический анализ данных. STATIS-TICA 6.0. Краснодар, 2005.
2. Халафян А.А., Савенко Д.В., Городин В.Н. // Обзорение прикладной и промышленной математики. 2006. № 2. С. 359–360.
3. Халафян А.А., Савенко Д.В., Городин В.Н. Идентификация состояния больных лептоспирозом (*DILEPTS*): Свидетельство об официальной регистрации программы для ЭВМ № 2007610599. Федеральная служба по интеллектуальной собственности, патентам и товарным знакам. Зарегистрировано 7.02.07.

**Кубанский государственный университет,
Специализированная клиническая инфекционная больница, г. Краснодар**

10 июля 2007 г.
