

УДК 615.47:617-089

А.И. Бых¹, Т.В. Жемчужкина², Т.В. Носова³, В.А. Сухенький⁴¹ХНУРЭ, г. Харьков, Украина, bykh@kture.kharkov.ua;²ХНУРЭ, г. Харьков, Украина, zhemchuzhkina@rambler.ru;³ХНУРЭ, г. Харьков, Украина, E-mail: yasanosova@rambler.ru;⁴ХНУРЭ, г. Харьков, Украина, sva89kolomiets@gmail.com

ПОИСК ИНФОРМАТИВНЫХ КОЛИЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ ЭЛЕКТРОМИОГРАФИЧЕСКОГО СИГНАЛА. СООБЩЕНИЕ 3

В результате проведения исследований методом дискриминантного анализа был произведен отбор информативных количественных показателей ЭМГ сигнала и найдены коэффициенты канонических дискриминантных функций. Проведена классификация 65-ти ЭМГ сигналов в норме и патологии, точность которой составила 95,4%.

ЭЛЕКТРОМИОГРАММА, ДИСКРИМИНАНТНЫЙ АНАЛИЗ, СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ, ТЕРРИТОРИАЛЬНАЯ КАРТА

Введение

На сегодняшний день исследование электрической активности мышц становится все более популярным среди врачей-диагностов. Электромиография (ЭМГ) является единственным объективным и информативным методом исследования функционального состояния периферической нервной системы, патология которой в структуре неврологических заболеваний занимает ведущее место [1]. Электромиографические исследования позволяют не только установить характер заболевания, проводить его топическую диагностику, но и объективно контролировать эффективность лечения, прогнозировать время и этапы восстановления пациента.

Автоматизированные системы измерения и обработки медико-биологической информации, использующие современные программные средства, существенно расширяют диагностические возможности современной медицины. Это касается и электромиографии – метода исследования нервно-мышечной системы посредством регистрации электрических потенциалов мышц.

Все патологические характеристики ЭМГ определяются врачом по внешним признакам, то есть носят субъективный характер. Таким образом, постановка диагноза зависит от квалификации, опыта, интуиции и физиологического и эмоционального состояния врача. Несмотря на широкое применение анализа ЭМГ в клинической медицине, остается актуальным получение информативных показателей ЭМГ–сигнала для улучшения диагностирования патологий нервно-мышечной системы. Более подробное изучение и анализ ЭМГ позволят обеспечить не только дифференцирование патологий, но и обнаруживать эти патологии на более ранних стадиях, а также сократить время диагностирования.

Актуальной задачей является определение количественных характеристик ЭМГ–сигнала при патологических состояниях нервно-мышечной

системы и поиска наиболее точных методов классификации и прогнозирования заболеваний.

1. Постановка задачи

Целями данной работы являются:

- 1) отбор информативных показателей ЭМГ–сигналов дискриминантным методом;
- 2) классификация пациентов по диагностическим группам;
- 3) изучение эффективности использования выбранного метода путем сравнения полученных результатов с диагнозами, поставленными врачами-экспертами.

В сообщениях 1, 2 [2, 3] нами были рассмотрены количественные показатели интерференционного ЭМГ–сигнала, полученные методами статистического исследования биомедицинских сигналов. Было установлено, что рассчитанные характеристики являются статистически независимыми и могут быть использованы для дифференциальной диагностики состояния нервно-мышечной ткани. Поэтому их мы и будем использовать в качестве исходных данных.

Для решения поставленной задачи был выбран метод дискриминантного анализа. Метод отвечает нашим требованиям, поскольку его использование позволяет:

1. Ответить на вопрос о том, возможно ли, используя данный набор параметров, отличить одну группу от другой; насколько хорошо эти параметры помогают провести дискриминацию; и какие из них наиболее информативны.
2. Получить несколько функций, обеспечивающих возможность отнесения данного объекта к одной из групп. Эти функции называются классифицирующими и зависят от значений параметров таким образом, что появляется возможность отнести каждый объект к одной из групп.

Таким образом, нам удастся получить принципиально новые результаты: как научные (уточнение информативных характеристик для проведения

дальнейших исследований), так и практические (возможность классификации ЭМГ-сигналов по группам нормы и разной степени патологии). Это, в свою очередь, позволит сократить время диагностирования и обеспечить не только дифференцирование патологий, но и обнаруживать эти патологии на более ранних стадиях.

2. Решение задачи

С помощью дискриминантного анализа на основании некоторых признаков (независимых переменных) индивидуум может быть причислен к одной из двух (или к одной из нескольких) заданных заранее групп. Ядром дискриминантного анализа является построение так называемой канонической дискриминантной функции

$$d_{km} = \beta_0 + \beta_1 x_{1km} + \dots + \beta_p x_{pkm},$$

где: d_{km} – значение дискриминантной функции для m -го объекта в группе k ; x_{pkm} – значение дискриминантной переменной x_1 для m -го объекта в группе k , соответствующих рассматриваемым случаям; $\beta_0 \dots \beta_1$ – неизвестные коэффициенты.

Целью анализа является определение таких коэффициентов, с помощью которых дискриминантные функции дали бы возможность с максимальной точностью провести разделение по группам [4].

Наиболее общим принципом применения дискриминантного анализа является включение в исследование по возможности большего числа переменных с целью определения тех из них, которые наилучшим образом разделяют выборки между собой. Пошаговая процедура дискриминантного анализа для отбора переменных основывается на F -критериях однофакторного дисперсионного анализа: “ F -включения” и “ F -исключения”. Значение F -статистики для переменной указывает на ее статистическую значимость при дискриминации между совокупностями и является мерой вклада признака в предсказание членства в группах. Тогда в качестве результата успешного анализа можно сохранить только наиболее информативные переменные модели, то есть те переменные, чей вклад в дискриминацию больше остальных [5].

В исследовании мы использовали количественные статистические характеристики 65-ти ЭМГ-сигналов (снятых с нескольких мышц пациентов в норме и патологии), полученные нами ранее [2, 3], а именно: минимум и максимум выборки амплитуд (MIN , MAX), вариационный размах ($Delta$), математическое ожидание (M), среднее квадратическое отклонение (SKO), дисперсия (D), коэффициент вариации (V), асимметрия (As), эксцесс (Ex), мода (Mo), амплитуда моды (AMo), амплитуда моды в процентах (AMo_perc), треугольный индекс (TI), индекс наклона гистограммы (N), индекс

напряженности (IN), значение периодограммы в нулевой частоте ($P0$), средняя спектральная мощность (SP), значение спектра в нулевой частоте ($S0$), эффективная ширина спектра (wS), средняя мощность периодограммы (PP), эффективная ширина периодограммы (wP), погрешность аппроксимации амплитуд волн ЭМГ нормальным распределением ($deltaNorm$), погрешность аппроксимации длительностей волн ЭМГ экспоненциальным распределением ($deltaExp$).

Очевидно, что классификация переменных будет осуществляться тем лучше, чем меньше рассеяние точек относительно центроида внутри группы и чем больше расстояние между центроидами групп. Один из методов поиска наилучшей дискриминации данных заключается в нахождении такой канонической дискриминантной функции d , которая максимизировала бы отношение межгрупповой вариации к внутригрупповой:

$$\lambda = B(d) / W(d),$$

где B межгрупповая и W внутригрупповая матрицы рассеяния наблюдаемых переменных от средних [5, 6].

Классификация объектов в дискриминантном методе производится с помощью функции расстояния. В качестве меры сходства можно использовать евклидово расстояние между объектами. Чем меньше расстояние между объектами, тем больше сходство. Однако в тех случаях, когда переменные коррелированы, измерены в разных единицах и имеют различные стандартные отклонения, полезнее применить не евклидово расстояние, а выборочное расстояние Махаланобиса

$$D^2 = \sum_{k=1}^p n_k \times (x_k - \bar{x}) \times C^{-1} \times (x_k - \bar{x}).$$

При использовании функции расстояния, объект относят к той группе, для которой расстояние D^2 наименьшее. На содержательном уровне его можно интерпретировать как взвешенную сумму расстояний от вектора средних каждой группы x_k до общего вектора средних \bar{x} . Если гипотеза H_0 верна, а объем выборки стремится к ∞ (достаточно большой), то D^2 может быть аппроксимирована F -распределением. Другим, в некоторых случаях более точным способом проверки гипотезы H_0 , является использование U -статистики Уилкса (она же – лямбда Вилкса), которая вычисляется как отношение детерминантов (\det) матрицы внутригрупповой ковариации W и полной ковариационной матрицы T

$$U = \det(W) / \det(T).$$

В некоторых работах для классификации используется обобщенное расстояние Махаланобиса V (обобщение величины D^2). Эта мера, известная как V -статистика Рао (аппроксимация статистики

U-Уилкса с помощью F-распределения), измеряет расстояния от каждого центроида группы до главного центроида с весами, пропорциональными объему выборки соответствующей группы. Она применима при любом количестве классов и может быть использована для проверки гипотезы H0. Отметим, что при включении или исключении переменных U-статистика имеет распределение χ-квадрат с числом степеней свободы, равным количеству групп -1, умноженное на число переменных, включенных (исключенных) на этом шаге. Если изменение статистики не значимо, то переменную можно не включать. Если после включения новой переменной V-статистика оказывается отрицательной, то это означает, что включенная переменная ухудшает разделение центроидов.

В ходе проведения дискриминантного анализа на каждом шаге (максимальное число шагов равно 48), включалась переменная, минимизирующая индикатор “лямбда Уилкса”. Максимальная значимость F-включения переменной составляла 0,05, а минимальная значимость F-исключения – 0,10.

3. Основные результаты

В результате проведения анализа были определены переменные, которые вносят значительный статистический вклад в дискриминацию элементов по диагностическим группам. Таким образом, делаем вывод, что информативными параметрами являются: *M, As, Ex, Mo, AMo, TI, N, IN, P0, SP, SO, wS, wP, deltaNorm, deltaExp*, т.е. 15 из 23 выбранных для анализа параметров.

Полученные нормированные канонические дискриминантные функции имеют вид:

$$d_1 = -1.931 \times M - 1.223 \times As + 2.726 \times Ex + 0.172 \times AMo + 0.183 \times Mo + 2.137 \times TI + 0.931 \times N - 0.644 \times IN - 1.596 \times SP - 0.055 \times SO + 1.542 \times wS - 3.048 \times P0 + 0.295 \times wP - 3.274 \times deltaNorm - 0.304 \times deltaExp;$$

$$d_2 = 0.624 \times M + 0.439 \times As - 1.31 \times Ex - 4.105 \times AMo + 1.431 \times Mo - 0.169 \times TI - 0.654 \times N + 0.588 \times IN + 0.713 \times SP + 0.839 \times SO + 2.391 \times wS - 1.902 \times P0 + 0.861 \times wP + 3.396 \times deltaNorm + 0.64 \times deltaExp;$$

$$d_3 = 1.334 \times M + 0.198 \times As + 0.211 \times Ex + 0.778 \times AMo + 1.062 \times Mo + 0.054 \times TI + 0.669 \times N - 0.226 \times IN + 0.606 \times SP + 0.382 \times SO + 1.315 \times wS - 2.313 \times P0 + 1.909 \times wP - 0.917 \times DeltaNorm + 0.302 \times deltaExp;$$

Из рис. 1 видно, что показатели расположены отдельно по диагнозам. Это говорит о правильном выборе значимых количественных показателей электромиографического сигнала.

На территориальной карте (рис. 2), показано разделение на области, которые означают принадлежность к группе. При этом в пределах границ

соответствующей области вероятность отнесения к данной группе выше, чем для других групп. На границах областей вероятности для граничащих групп одинаковы. Символы, используемые в территориальной карте, приведены в табл. 1.

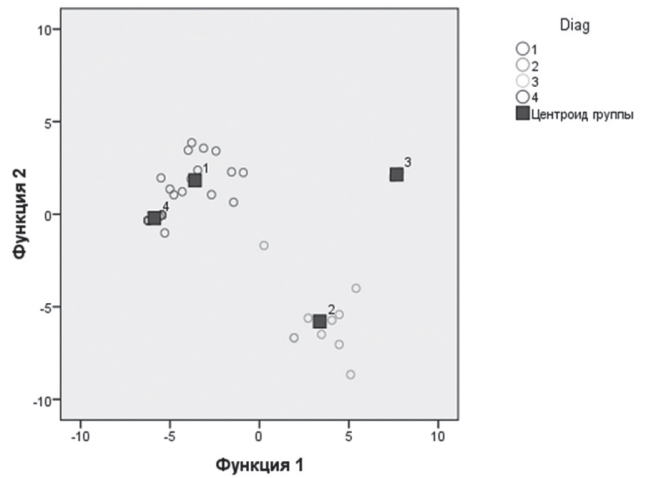


Рис. 1. Размещение центроидов групп и показателей в пространстве признаков

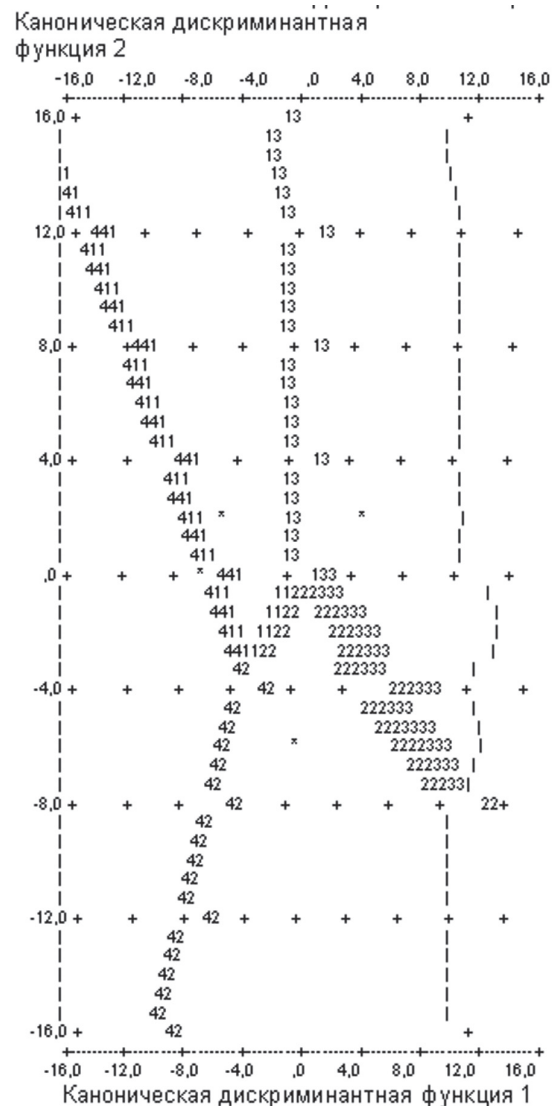


Рис. 2. Территориальная карта

В результате проведения дискриминантного анализа был определен набор диагностически значимых количественных показателей ЭМГ-сигнала, а именно: математическое ожидание, асимметрия, эксцесс, мода, амплитуда моды, триангулярный индекс, индекс наклона гистограммы, индекс напряженности, значение периодограммы в нулевой частоте, средняя спектральная мощность, значение спектра в нулевой частоте, эффективная ширина спектра, эффективная ширина периодограммы, погрешность аппроксимации амплитуд волн ЭМГ нормальным распределением, погрешность аппроксимации длительностей волн ЭМГ нормальным распределением.

Таблица 1

Символы, используемые в территориальной карте

Символ	Группа	Метка
1	1	Диагноз: норма
2	2	Диагноз: слабые нарушения
3	3	Диагноз: сильные нарушения
4	4	Диагноз: денервация
*	центр группы	

В дискриминантном анализе процедура классификации используется для определения принадлежности к той или иной группе случайно выбранных объектов, которые не были включены при выводе дискриминантной и классифицирующих функций. Для проверки точности классификации применим классифицирующие функции к тем объектам, по которым они были получены. По количеству правильно классифицированных объектов можно оценить точность процедуры классификации. Результаты такой классификации представлены в табл. 2.

Таблица 2

Классификационная таблица

		Результаты классификации					
		Диагноз	Предсказанная принадлежность к группе, %				Итого
			1	2	3	4	
Группа	1	87,5	0,0	0,0	12,5	100,0	
	2	9,1	90,9	0,0	0,0	100,0	
	3	0,0	0,0	100,0	0,0	100,0	
	4	0,0	0,0	0,0	100,0	100,0	

Выводы

При помощи найденных дискриминантных функций была проведена классификация ЭМГ-сигналов в норме и при разной степени патологии. Фактическая и прогнозируемая группы не совпали

для 3 из 65 объектов. Таким образом, точность классификации при данном наборе дискриминантных переменных составляет 95,4%.

Высокая эффективность использования выбранного метода обеспечивает возможность широкого применения полученных результатов в практической диагностике, благодаря параметризации и переводу анализа в количественную область, возможности проводить классификацию произвольных объектов по найденным классификационным функциям.

Список литературы: 1. Гехт, Б.М. Теоретическая и клиническая электромиография [Текст] / Б.М. Гехт. – Л.: Наука, 1990. – 229 с. 2. Бых, А.И. Поиск информативных количественных показателей электромиографического сигнала. Сообщение 1 [Текст] / А.И. Бых, Т.В. Жемчужкина, Т.В. Носова // Бионика интеллекта. – 2007. – № 1 (66). – С. 18–125. 3. Бых, А.И., Поиск информативных количественных показателей электромиографического сигнала. Сообщение 2. [Текст] / А.И. Бых, Т.В. Жемчужкина, Т.В. Носова // Бионика интеллекта. – 2008. – № 1 (68). – С. 111–116. 4. Ким, Дж.-О. Факторный, дискриминантный и кластерный анализ [Текст] / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др. М.: 1989. 5. Енюков, И.С. Факторный, дискриминантный и кластерный анализ: сборник работ под ред. Енюкова И. С. – М.: Финансы и статистика, 1989. – 215 с. 6. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности [Текст] / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607 с.

Поступила в редколлегию 20.01.2012

УДК 615.47:617-089

Пошук інформативних кількісних показників електроміографічного сигналу. Повідомлення 3 / А.І. Бих, Т.В. Жемчужкіна, Т.В. Носова, В.А. Сухенький // Біоніка інтелекту: наук.-техн. журнал. – 2012. – № 1 (78). – С. 68-71.

Робота присвячена пошуку інформативних кількісних показників ЕМГ-сигналів для диференціальної діагностики стану нервово-м'язової системи. Для відбору діагностично значущих статистичних параметрів застосовується метод дискримінантного аналізу.

Табл. 2. Іл. 2. Бібліогр.: 6 найм.

UDC 615.47:617-089

Searching for informative quantitative parameters of electromyographic signal / A.I. Bykh, T.V. Zhemchuzhkina, T.V. Nosova, V.A. Suchenkyi // Bionics of Intelligense: Sci. Mag. – 2012. – № 1 (78). – P. 68-71.

The work is dedicated to the searching informative quantitative parameters of EMG signals for the differential diagnostics of the condition of the neuromuscular system. The method of discriminant analysis has been used for selection of diagnostically significant statistical parameters.

Tab. 2. Fig. 2. Ref.: 6 items.