

У статті запропоновано методику визначення систолічної дисфункції міокарда у підлітків, отриману в результаті дискримінантного аналізу клінічних, інструментальних і лабораторних показників

Ключові слова: дискримінантний аналіз, систолічна дисфункція міокарда

В статтє предложена методика определения систолической дисфункции миокарда у подростков, полученная в результате дискриминантного анализа клинических, инструментальных и лабораторных показателей

Ключевые слова: дискриминантный анализ, систолическая дисфункция миокарда

The method for determination of systolic myocardial dysfunction in adolescents resulting from discriminant analysis of clinical, instrumental and laboratory parameters is offered in this article

Keywords: discriminant analysis, systolic dysfunctions

МЕТОДИКА ОПРЕДЕЛЕНИЯ СИСТОЛИЧЕСКОЙ ДИСФУНКЦИИ МИОКАРДА У ПОДРОСТКОВ

Е. В. Высоцкая

Кандидат технических наук, доцент*

Л. И. Рак

Кандидат медицинских наук, ведущий научный сотрудник

Отделение педиатрии и реабилитации

ГУ «Институт охраны здоровья детей и подростков АМН

Украины»

пр. 50-летия ВЛКСМ, 52-а, г. Харьков, Украина, 61153

Контактный тел.: (057) 62-11-50

А. П. Порван

Кандидат технических наук, научный сотрудник*

Контактный тел.: (057) 702-13-64

О. А. Сватенко*

*Кафедра биомедицинских электронных устройств и систем

Харьковский национальный университет радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

Контактный тел.: (057) 702-13-64

E-mail: diagnost@kture.kharkov.ua

1. Введение

Всем известно, что за последнее время сильно возросло количество людей, имеющих проблемы с сердечно-сосудистой системой. Все чаще болезни сердца диагностируют не только у взрослых, но и у детей и подростков. На фоне такой опасной патологии, как хроническая сердечная недостаточность (ХСН), включающей в себя врожденные пороки сердца, его множественные малые структурные аномалии, кардиомиопатии физического напряжения и другие заболевания, создаются условия, приводящие к ухудшению функциональной способности сердца, снижению физической активности и трудоспособности пациента, ухудшению качества жизни и, как следствие, значительно ее укорочению.

Развитие ХСН - это патофизиологически необратимый процесс. Механизмы ее формирования, выполняющая адаптационную функцию на начальных этапах, постепенно становятся дезадаптационными и включают активацию нейрогуморальных систем регуляции, системы цитокинов и оксидативного стресса, ремоделирование миокарда, которые имеют тесную взаимосвязь. Однако у подростков эти механизмы в комплексе до сих пор не изучались.

В связи с вышесказанным особенное значение имеет выявление влияния этих показателей на работу сердечно-сосудистой системы (ССС) подростков и, соответственно, диагностика систолической дисфункции миокарда (СДМ), которая является одним

из ранних проявлений доклинической стадии ХСН [1, 2].

2. Анализ проблемы и постановка задачи исследования

Наиболее близким по своему функциональному назначению к предлагаемому методу является метод диагностики субклинической сердечной недостаточности у детей с патологией миокарда [3], в котором с помощью эхокардиографии исследуют фракцию выброса и ударный объем сердца до и после физической нагрузки в 20 приседаний за 30 секунд и дополнительно проводят биохимические исследования для определения суточной экскреции катехоламинов с мочой. При этом снижение показателей фракции выброса и ударного объема сердца после физической нагрузки хотя бы на 1-2% и уровня катехоламинов в моче: для адреналина – ниже 18,0 нмоль/сут, для норадреналина – ниже 80 нмоль/сут, свидетельствует о субклинической стадии сердечной недостаточности. Недостатком такого метода является невозможность определения типа реакции на физическую нагрузку у пациентов с патологией миокарда и выявления СДМ при отсутствии существенных изменений показателей нейрогуморальной регуляции или их незначительных колебаний.

Поскольку в медицинской практике существует проблема нехватки данных в связи со сложностью их получения, возникает вопрос заполнения пропусков

–импутирования [4, 5]. Условно методы импутирования можно разделить на простые и сложные.

К простым (неитеративным) методам, основанным на элементарных арифметических операциях, расчете расстояний между объектами, относят: метод HotDeck, заполнение пропусков средним арифметическим, регрессионное моделирование пропусков и подбор в группе.

Самым простым методом является заполнение средним значением (модой, медианой или средним арифметическим) по имеющимся данным. Он не требует применения специального программного обеспечения, однако уменьшает дисперсию признака.

С помощью метода ближайшего соседа пропуски заполняются различными значениями, полученными в результате оценивания расстояния между центроидами рассматриваемого набора данных. Необходимы значительные вычислительные затраты. К тому же существует вероятность неточных предсказаний при условии, что пропуск данных не имеет закономерности.

При использовании метода многомерной регрессии строится модель линейной зависимости переменной, в которой необходимо заполнить пропуски, от ряда других имеющихся признаков. Регрессионные коэффициенты для каждого из предикторов находятся методом наименьших квадратов в массиве с полными данными. Недостатком метода является возможность пропуска не только значения предсказываемой переменной, но и значения предикторов – предсказание непосредственно на основе коэффициентов уравнения оказывается невозможным. Сложность метода еще и в том, что исследователю необходимо выбрать переменные, коррелирующие с рабочей переменной.

Сложные алгоритмы (итеративные) предполагают оптимизацию некоторого функционала, отражающего точность расчета подставляемых на место пропуска значений. Их делят на глобальные и локальные.

Особенностью локальных алгоритмов является то, что в оценивании (предсказании) каждого пропущенного значения участвуют полные наблюдения, находящиеся в некоторой окрестности предсказываемого объекта. К данной группе относятся алгоритмы Zet и Zet Braid.

Глобальные алгоритмы для оценивания каждого пропущенного значения оперируют всеми объектами рассматриваемой совокупности. К ним относятся метод Бартлетта, EM-оценивания (EM – expected value maximization) и Resampling (метод попарного сравнения).

Метод Бартлетта включает в себя три итерации. На первой итерации пропуски заполняются некоторым начальным значением. На второй итерации для преобразованной переменной строится регрессионная модель. На заключительном этапе на основе полученного регрессионного уравнения предсказываются новые значения для пропусков.

Суть алгоритма Resampling состоит в том, что значения для пропусков выбираются из имеющихся случайным образом, с возвращением (когда значение может использоваться еще раз после выбора) или без него. После этого на всем массиве строится регрессионная модель.

EM-метод позволяет не только восстановить пропущенные значения посредством двухэтапного итеративного алгоритма, но и оценить средние значения для количественных переменных.

Исходя из описанного выше, разработка методики определения СДМ у подростков должна учитывать проблему нехватки данных.

Для построения математической модели диагностирования той или иной патологии миокарда на современном этапе развития науки широко применяются методы корреляционного и регрессионного анализов. Это, как показал опыт, далеко не всегда позволяет выявить сущность исследуемых явлений и, более того, не дает гарантий в отношении надежности результатов.

Недостатки корреляционного анализа связаны с тем, что он позволяет лишь констатировать наличие некой связи между параметрами или переменными, но не может доказать причинно-следственный тип этой связи.

При использовании регрессионного анализа синтезированные модели могут оказаться неточными, а их усложнение приводит к избыточности, что, в свою очередь, снижает эффективность их использования.

Методы кластерного анализа имеют широкие возможности, но их использование предполагает значительные затраты времени и сил для проведения соответствующих расчетов. Сложность состоит и в том, что полученные результаты группирования интерпретируются не всегда правильно. Особенно, если информативные признаки, характеризующие объекты исследования, достаточно разнородны.

В отличие от корреляционного, в дисперсионном анализе исследователь исходит из предположения, что одни переменные выступают как влияющие (именуемые факторами или независимыми переменными), а другие (результативные признаки или зависимые переменные) – подвержены влиянию этих факторов. Применяется только тогда, когда известно, что распределение результативного признака является нормальным.

Поэтому при построении диагностической модели СДМ целесообразно использовать математический аппарат, учитывающий все недостатки вышеописанных методов вероятностного моделирования, и дающий возможность быстро и качественно классифицировать исследуемые объекты и синтезировать адекватную линейную математическую модель. Таким методом является метод построения дискриминантных функций, который имеет ряд преимуществ: учитывается вариабельность параметров, рассматривается совокупность всех клинических и параклинических показателей, взятых со своими коэффициентами, которые отражают удельный вес влияния каждого показателя на диагноз [6].

Таким образом, целью нашей работы является изучение механизмов формирования ХСН в комплексе на основе комбинации основных клинических параметров и разработка методики определения систолической дисфункции миокарда у подростков.

Использование предлагаемой методики даст возможность выявлять заболевание на самых ранних стадиях развития, что будет способствовать своевременному качественному лечению и предотвращению снижения трудоспособности пациентов.

В основу предлагаемой методики был положен математический метод дискриминантных функций.

3. Экспериментальные исследования

Методика определения СДМ по данным клинических и инструментальных показателей заключается в следующем:

1. Сбор материалов, необходимых для проведения исследования СДМ на начальных стадиях, который включает в себя проведение эхокардиографии, биохимических и иммунологических исследований, определение анамнеза жизни пациента, его жалоб и др.

2. Импутирование пропущенных данных с использованием EM-метода [7].

3. Определение общих параметров выборки.

4. Составление схемы кодировки анкеты врача.

5. Ввод данных в компьютер.

6. Модификация и отбор данных.

7. Кодирование качественных показателей согласно схеме, разработанной в п.4.

8. Описательный анализ данных (построение линейных распределений).

9. Определение предельной ошибки выборки.

10. Определение необходимого объема выборки пациентов:

$$n_{\omega} = \frac{z^2 \cdot \omega \cdot (1 - \omega)}{\Delta_{\omega}^2}, \quad (1)$$

z – значение коэффициента доверия Стьюдента при заданном значении надежности $\gamma = 0,95$;

ω – пациенты с адекватной реакцией ССС на физическую нагрузку, %

Δ_{ω} – предельная ошибка выборки.

11. Расчет канонических коэффициентов дискриминантных функций b_i , $i = 1, \dots, n$. Целью является определение таких коэффициентов, чтобы по значениям дискриминантной функции можно было с максимальной четкостью провести разделение по группам.

12. Определение количества дискриминантных функций:

$$DF = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n + a, \quad (2)$$

где x_1 и x_2 – значения переменных, соответствующих рассматриваемым случаям, $b_1 \dots b_n$ и a – коэффициенты, которые оцениваются с помощью дискриминантного анализа.

13. Определение коэффициентов канонической корреляции [8].

14. Оценка значимости дискриминантных функций с помощью λ -статистики Уилкса. λ -статистика Уилкса – это мера различий между классами по нескольким дискриминантным переменным. Чем ближе значение λ к 0, тем лучше различие классов, а чем ближе к 1, тем различие хуже (классы совпадают).

15. Расчет критериального значения χ^2 [7]. Сравнивается расчетный χ^2 с критическим, с заданным числом степеней свободы. Число степеней свободы зависит от количества дискриминантных функций p и количества классов g :

$$dfk = (p - (k - 1))(g - (k - 1) - 1). \quad (3)$$

Если $\chi^2 > \chi_{кр}^2$, то значимость гипотезы подтверждается.

16. Построение территориальной карты на основе значений дискриминантных функций. На территориальной карте показывается 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.

17. Получение информации о фактической и прогнозируемой группах для каждого объекта.

18. Определение группы, имеющей наибольшую прогнозируемую вероятность включения данного пациента.

19. Определение группы, имеющей вторую по величине вероятность (после прогнозируемой):

- $P(G = g | D = d)$ – вероятность наблюдаемого значения дискриминантной функции (D), если задана принадлежность объекта к группе (G);

- $P(G = g | D > d)$ – вероятность принадлежности объекта к группе (G) при данной величине дискриминантной функции (D).

Если мы предположим, что каждый объект должен принадлежать к одной из групп, то можно вычислить вероятность принадлежности для любой из групп. Вероятность того, что объект X является членом класса G , равна:

$$P(G_g | X) = \frac{P(X | G_g)}{\sum_{i=1}^k P(X | G_i)}, \quad (4)$$

где величина $P(G_g | X)$ – дает вероятность того, что объект принадлежит классу g , а величина $P(X | G_g)$ – оценивает долю объектов в этом классе, которые отстоят от центра дальше, чем X .

20. Определение группы, к которой фактически принадлежит пациент.

21. Построение графика распределения пациентов с различными состояниями сердечно-сосудистой системы.

22. Формирование диагностического заключения.

Предлагаемая методика была апробирована на 137 подростках (10-18 лет) с различной реакцией ССС на физическую нагрузку.

Была проведена общая статистическая обработка показателей состояния сердечно-сосудистой системы и систем нейрогуморальной регуляции.

Так как при диагностике СДМ используется большое количество показателей, получение которых не всегда возможно, то мы применили EM-метод импутирования [7]. Заполнение пропусков позволяет с необходимым набором показателей морфофункционального состояния сердца и других биохимических констант построить диагностическую и прогностическую модели СДМ.

Был рассчитан оптимальный объем выборки.

Всех пациентов разделили на 3 группы:

1) подростки с адекватной реакцией на физическую нагрузку (1 группа – 52 подростка);

2) подростки с неадекватной реакцией на физическую нагрузку (2 группа – 23 подростка);

3) контрольная группа (3 группа – 62 здоровых подростка).

Все пациенты рассматривались как совокупность объектов с варьирующимися качественными и количественными характеристиками, на основании которых и определялась принадлежность объекта к группе. Этот подход позволил нам определить для новых объектов из той же совокупности вхождение в ту или иную из вышеперечисленных групп.

Было выявлено, что наименьший процент составляют подростки с неадекватной реакцией на физическую нагрузку, поэтому объем выборки для создания математической модели определялся с учетом этого показателя.

Была рассчитана предельная ошибка выборки $\Delta\omega = 5\%-6\%$. При надежности $\gamma = 0,95$ и предельной ошибке выборки $\Delta\omega = 5\%-6\%$ необходимый объем выборки составляет 119-168 человек.

Для определения значимых для формирования диагноза свойств, было проанализировано 83 признака. Все признаки были закодированы и поставлены в соответствии 83-мерному вектору, что учитывает отсутствие, наличие, направленность и величину каждого признака. В 83-мерном пространстве при диагностике 3-х состояний сердечно-сосудистой системы были получены 3 области: точки, присущие группе здоровых подростков; группе с адекватной реакцией на физическую нагрузку; группе с неадекватной реакцией на физическую нагрузку.

Рассчитали значения диагностических коэффициентов и получили 12 значимых для определения диагноза признаков.

В случае дифференциации 3-х состояний сердечно-сосудистой системы необходимы 2 дискриминантные функции:

$$DF1 = 0,061 \cdot X1 + 0,123 \cdot X2 - 0,03 \cdot X3 + 0,058 \cdot X4 - 0,09 \cdot X5 + 0,004 \cdot X6 - 2,346 \cdot X7 + 0,026 \cdot X8 + 0,022 \cdot X9 - 0,035 \cdot X10 + 1,643 \cdot X11 - 0,005 \cdot X12 + 0,989; \tag{5}$$

$$DF2 = 0,023 \cdot X1 + 0,155 \cdot X2 - 0,013 \cdot X3 + 0,052 \cdot X4 + 0,138 \cdot X5 + 0,003 \cdot X6 + 3,797 \cdot X7 + 0,718 \cdot X8 + 0,0001 \cdot X9 - 0,004 \cdot X10 + 0,265 \cdot X11 + 0,007 \cdot X12 - 16,452.$$

Мерой удачного разделения на группы служат корреляционные коэффициенты между рассчитанными значениями дискриминантных функций и показателями принадлежности к группе (табл. 1).

Таблица 1

Характеристика дискриминантных возможностей функций

Функция	Собственное значение	% дисперсии	Общий %	Каноническая корреляция
1	4,026	81,9	81,9	0,895
2	0,890	18,1	100,0	0,686

Приведенные в таблице данные свидетельствуют, что первая функция имеет наибольшие возможности, что дискриминируют. Чтобы сравнить функции, мы провели расчеты процентов дисперсии каждой функции.

Таким образом, в приведенной системе уравнений первая функция содержит 81,9 % общих дискриминантных возможностей. Учитывая значение канонического коэффициента первой функции (0,895), можно сделать вывод о существовании высокой связи между диагнозом и значениями первой дискриминантной функции.

Также нами была обнаружена высокая связь между диагнозом и значениями второй дискриминантной функции (0,686). Качественная оценка силы связи r_{xy} величин X и Y была сделана на основании шкалы Чеддока.

Оценка значимости дискриминантных функций также была проверена λ -статистикой Уилкса (табл. 2).

Таблица 2

Определение значимости дискриминантных функций методом λ -статистики Уилкса

Тест функций	Лямбда Уилкса	χ^2	Степени свободы, df	Значимость
1	0,032	523,746	20	0,000
2	0,442	123,819	9	0,000

λ -статистика Уилкса – это мера различий между классами по нескольким переменным (дискриминантным переменным). Чем ближе значение λ к 0, тем лучше различие классов, а чем ближе к 1, тем различие хуже (классы совпадают).

Рассчитали критериальное значение χ^2 , которое сравнивается с критическим с заданным числом степеней свободы. Если значение χ^2 больше критического, то значимость подтверждается. В нашем случае, уровень значимости равен 0,000 ($P < 0,001$), то есть различие между группами значимое.

На основе значений обеих дискриминантных функций построена территориальная карта распределения больных в зависимости от реакции на физическую нагрузку.

Анализ полученных данных показал, что из 52 пациентов первой группы (подростки с адекватной реакцией на физическую нагрузку) двое были ошибочно отнесены ко второй группе. Из 23 пациентов второй группы (подростки с неадекватной реакцией на физическую нагрузку) по двое были ошибочно отнесены к первой и третьей группе. Из 62 пациентов третьей группы (здоровые подростки) только 1 был ошибочно отнесен к первой группе.

Распределение пациентов с различной реакцией сердечно-сосудистой системы на физическую нагрузку показано на рис. 1, где по оси X располагаются значения дискриминантной функции 1 (Function1(DF1)), а по оси Y – значения дискриминантной функции 2 (Function2(DF2)).

Анализ результатов классификации показал, что прогноз для исследованных групп сделан верно (корректно классифицировано в общем 94,9% всех больных систолической дисфункцией миокарда).

Соответственно, точность предсказания СДМ для обследуемых пациентов может достигать 94,9% в том случае, если выборка пациентов будет идентична тем пациентам, данные которых послужили основой для прогноза.

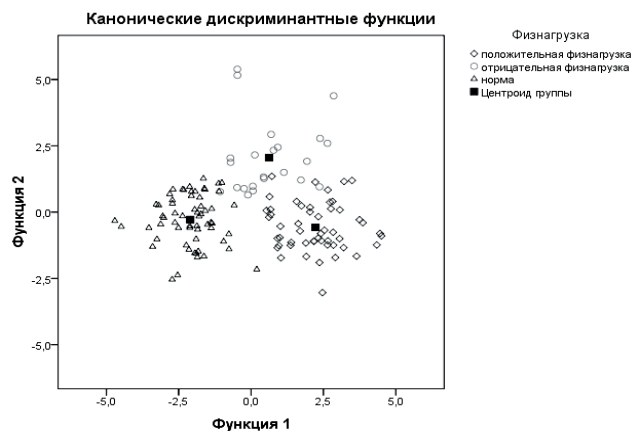


Рис. 1. Распределение пациентов с различной реакцией сердечно-сосудистой системы на физическую нагрузку

4. Выводы

Предложенная методика позволяет определить наличие или отсутствие систолической дисфункции миокарда на ранней стадии развития с адекватной или неадекватной реакцией на физическую нагрузку у подростков. При этом неадекватная реакция на физическую нагрузку свидетельствует о наличии систолической дисфункции миокарда.

Методика имеет большое практическое значение. Внедрение ее в клиническую практику позволит не только объяснить механизм патологического процесса, правильно оценить прогноз заболевания, но и разработать эффективные лечебно-профилактические мероприятия.

Литература

1. Воронков, Л.Г. Первичная профилактика сердечной недостаточности – один из приоритетов современной кардиологии [Текст] / Л.Г.Воронков. // Укр. кардіол. журнал. - 2004. - № 4. - С. 9-13.
2. Подростковая медицина [Текст] / под ред. Л.И. Левиной, А.М. Куликова. – СПб.: Питер, 2006. – 544 с.
3. Пат. 23297 Украина, МПК А 61 В 10/00, G 01 N 33/48. Спосіб діагностики субклінічної серцевої недостатності у дітей з патологією міокарда / Богмат Л.Ф., Рак Л.І., заявитель и патентообладатель Институт охраны здоровья детей и подростков академии медицинских наук Украины. - № u200608054. заявл. 17.07.06, опубл. 25.05.07, бюл. №7. – 3 с.
4. Злоба, Е. Статистические методы восстановления пропущенных данных [Текст] / Е. Злоба, И. Яцкив // Computer Modeling & New Technologies. – 2004. – Vol. 6. – P. 55–56.
5. Douglas, L.A. Missing data / A.L. Douglas, L. Marsh, M. Smithson et al. / SAGE, 2001. – P. 58.
6. Nenadic, Z. Information discriminant analysis: feature extraction with an information-theoretic objective / Z. Nenadic. // IEEE T. Pattern Anal. – 2007. – Vol. 29 (8). – P. 1394-1407.
7. McLachlan, G.J., Thriyambakam, K. The EM Algorithm and Extensions / G.J. McLachlan, K.Thriyambakam. – N.Y.: John Wiley & Sons, Inc., 1946. – P. 274.
8. Факторный, дискриминантный и кластерный анализ [Текст] / Дж. Ш. Ким, Ч.У. Мюллер, У.Р. Клекка и др. / под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.

Розглянуто питання дослідження проектів на транспорті. Запропонована класифікація проектів з урахуванням особливостей транспортної діагностики

Ключові слова: проект, класифікація, діагностика, транспорт

Рассмотрены вопросы исследования проектов на транспорте. Предложена классификация проектов с учетом особенностей транспортной диагностики

Ключевые слова: проект, классификация, диагностика, транспорт

The questions of projects research on transport are considered. The classification of projects taking into account features of transport diagnostics is offered

Keywords: project, classification, diagnostics, transport

УДК 656:681.518.5

КЛАССИФИКАЦИЯ ПРОЕКТОВ НА ТРАНСПОРТЕ С ПОЗИЦИИ ТРАНСПОРТНОЙ ДИАГНОСТИКИ

А. Н. Горяинов

Кандидат технических наук, доцент
Харьковской национальной академии городского хозяйства
ул. Революции, 12, г. Харьков, Украина, 61002

1. Введение

Формирование теории транспортной диагностики затрагивает вопросы описания объектов ди-

агностирования на транспорте. Согласно [1], одним из видов таких объектов могут выступать проекты на транспорте. На сегодняшний день в научной литературе вопросы диагностирования проектов