

УДК 616.12-008.331, 004.65

Е.В. ВЫСОЦКАЯ, А.П. ПОРВАН, Д.А. АМБРОСОВ, Т.Н. АМБРОСОВА

Харьковский национальный университет радиоэлектроники, Украина

ПРИМЕНЕНИЕ ДИСКРИМИНАНТНОГО АНАЛИЗА ДЛЯ ДИАГНОСТИКИ СТАДИЙ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ

В статье авторами рассмотрены существующие методы синтеза математической модели на основе теории вероятности и полностью обосновано применение метода дискриминантных функций. Предлагается диагностическая модель дифференциации стадий артериальной гипертензии, полученная в результате дискриминантного анализа данных клинических, инструментальных и лабораторных исследований. Полученные дискриминантные функции используются при построении территориальной карты диагностирования артериальной гипертензии, а статистические оценки синтезированной модели показывают ее высокую значимость и адекватность.

Ключевые слова: дискриминантный анализ, территориальная карта, артериальная гипертензия, стадия, дифференциация признаков.

Введение

Среди хронических неинфекционных заболеваний одно из первых мест принадлежит артериальной гипертензии (АГ), ею страдают от 10 до 30% взрослого населения индустриальных стран [1]. АГ является также одним из главных факторов риска других широко распространенных сердечнососудистых заболеваний [2]. Так, существует тесная положительная корреляция между уровнем артериального давления (АД), заболеваемостью и смертностью от ишемической болезни сердца, инфаркта миокарда, инсульта, что приводит к ранней инвалидизации и смертности [3].

Очевидно, что успех профилактических и лечебных мероприятий в значительной степени зависит от своевременного выявления заболевания, в связи с чем важным является ранняя диагностика и профилактика АГ [4]. Эффективные меры раннего выявления и предупреждения развития АГ и ее осложнений позволят предупредить до 50% осложнений.

Так, в течение последнего десятилетия эффективные профилактические и медикаментозные мероприятия в отношении АГ в США позволили снизить смертность от мозгового инсульта на 70% и от инфаркта миокарда на 50%. Данную ситуацию медики расценивают важнейшим достижением в решении проблемы предупреждения развития и прогрессирования сердечнососудистой заболеваемости и смертности [5].

АГ чрезвычайно редко удается диагностировать на ранних стадиях, поскольку пациенты не обращаются своевременно за медицинской помощью.

Более 50% больных АГ не знают о своем диагнозе, из этого числа только 25% проводит медикаментозную терапию АГ и только у 12,5% терапия является адекватной и эффективной. Поэтому ранняя диагностика АГ является важной задачей современной медицины.

1. Анализ литературы и постановка задачи исследования

Артериальная гипертензия (АГ) относится к наиболее распространенным заболеваниям человека. По данным официальной статистики в Украине 30% взрослого населения страдает этим недугом. Среди всех людей с повышенным артериальным давлением знают о наличии заболевания 67,8% сельских и 80,8% городских жителей, а лечатся соответственно 38,3% и 48,4%. Эффективность лечения составляет 8,1% среди жителей сельской местности и 18,7% – среди жителей города.

АГ является основной причиной инвалидности и смертности в Украине. Так по данным эпидемиологических исследований в 2010 году в Украине зарегистрировано 25,9 млн. человек с заболеваниями сердечнососудистой системы, что составляет 57% населения, из них лица трудоспособного возраста 9,46 млн. человек. Наиболее распространенной является АГ – 46,5% (12,1 млн. чел.) из них трудоспособного возраста – 5,2 млн. человек [6].

Одним из приоритетных направлений в борьбе с распространенностью, заболеваемостью и смертностью от сердечнососудистой патологии и ее осложнений является разработка новых методов диагностики и прогнозирования развития АГ.

В клинической практике диагностика ранних стадий АГ находится на достаточно низком уровне, что в основном связано, с тем, что эпизодическое и лабильное повышение АД больные зачастую игнорируют и соответственно редко обращаются за квалифицированной медицинской помощью на ранних этапах развития болезни. Диагностика АГ, основанная на поражении органов-мишеней является достаточно трудоемким диагностическим процессом с вовлечением целого ряда инструментальных методов диагностики и консультаций специалистов (офтальмолога, невропатолога). Поэтому, разработка математических моделей и использование современных методов статистического анализа при установлении стадии АГ является актуальным направлением современной науки.

Этиология эссенциальной АГ до настоящего времени остается недостаточно ясной [7], однако в отношении проявлений болезни ситуация значительно более определенная – доминируют изменения гемодинамики, результирующей которых является транзиторное или постоянное повышение системного АД [8]. Поэтому представляется обоснованным использовать для целей диагностики именно показатели центральной гемодинамики.

Согласно одной из концепции в функционировании сердечнососудистой системы выделяют стабильный (постоянный) и пульсирующий (изменяющийся) компоненты. Систолическое (САД) и диастолическое артериальное давление (ДАД) рассматривают как экстремумы колебаний АД вокруг среднего уровня [9]. По мнению большинства исследователей, увеличению периферического сопротивления сосудов принадлежит ведущая роль в гемодинамическом механизме повышения АД при АГ [7]. Приблизительно у трети молодых пациентов с мягкой АГ наблюдается увеличение минутного объема сердца [10].

Для разработки критериев индивидуальной диагностики АГ по совокупности нескольких признаков применяли логистическую линейную регрессию [11 – 13]. Строили модели для различных сочетаний показателей с целью эффективно предсказывать наличие или отсутствие события, т.е. принадлежность индивида к больным АГ или здоровым соответственно.

Используя методы регрессионного анализа, можно достаточно точно выявить факторы, оказывающие существенное влияние на результирующий показатель, а также определить силу и направление этого влияния. Однако синтезированные модели могут оказаться неточными или, в случае сложных моделей, переобученными, что снижает эффективность их использования [14, 15].

Применение методов кластерного анализа позволяет сопоставить объекты моделирования по их качественным характеристикам, агрегировать экспертные оценки текущего и прогнозируемого уровней развития объектов и т.д. Несмотря на свои широкие возможности, использование данного метода анализа требует значительной затраты времени и сил для проведения соответствующих расчетов, а полученные результаты группирования не всегда адекватно интерпретируются, особенно если информативные признаки, характеризующие объекты исследования, были достаточно разнородны [16].

Указанные выше недостатки позволяет учесть дискриминантный анализ, дающий возможность быстро и качественно классифицировать исследуемые объекты и синтезировать адекватную линейную математическую модель. Данный метод математического анализа имеет ряд преимуществ, таких как: учитывается вариабельность параметра, рассматривается совокупность всех клинических и параклинических показателей, взятых со своими коэффициентами, которые отражают удельный вес влияния каждого показателя на постановку диагноза [17, 18].

Целью работы является применения многофакторного дискриминантного анализа для неинвазивной диагностики и прогнозирования стадии АГ.

2. Основной материал

При проведении дискриминантного анализа нами были проанализированы данные 238 пациентов с различными стадиями АГ. Все пациенты были разделены на 3 группы:

- 1-я группа – 74 пациентов с АГ I стадии;
- 2-я группа – 74 пациентов с АГ II стадии;
- 3-я группа – 72 пациентов с АГ III стадии.

С математической точки зрения все пациенты рассматривались как совокупность объектов с варьирующимися количественными характеристиками. На основании этих характеристик определялась группа, к которой относится объект. Это позволило нам спрогнозировать группы (наличие АГ) для новых объектов из той же совокупности.

Для определения значимых для формирования диагноза признаков были проанализированы: антропометрические показатели (рост, вес, индекс массы тела, окружность талии, окружность бедер, индекс талия/бедро); гемодинамические (уровни систолического, диастолического и пульсового давления, частота сердечных сокращений); клиниколабораторные показатели углеводного обмена (уровень глюкозы, инсулина, гликированного гемоглобина с математическим расчетом индексов инсулинорезистентности

НОМА, Саго, Duncan); показатели углеводного спектра (общий холестерин, триглицериды, холестерин липопротеидов высокой, низкой, очень низкой плотности, коэффициент атерогенности, аполипротеин В100); показатели активности адипокиновой системы (интерлейкин-6, фактор некроза опухоли-α, адипонектин, ингибитор активатора плазминогена-1).

Математическая обработка результатов проводилась на персональном компьютере с использованием SPSS 17.0.

Вначале был рассчитан оптимальный объем выборки. Так как из предварительно проведенных исследований стало известно, что среди всех пациентов больные с третьей стадией АГ составляют примерно 34,5% ($\omega = 76/220 = 0,345454$), при уровне надежности $\gamma=0,95$ и предельной погрешности выборки $\Delta_\omega=5-6\%$ необходимый объем выборки составляет:

$$n_\omega = \frac{t^2 \cdot \omega \cdot (1-\omega)}{\Delta_\omega} = 174 \div 238 \text{ человек,}$$

где t - стандартизированный коэффициент Стьюдента.

При статистическом анализе не учитывались данные пациентов, у которых определение стадия АГ было затруднено или не полностью обусловлено. С математической точки зрения все пациенты рассматривались как совокупность объектов с переменными количественными и качественными показателями.

При определении существенных для диагностики стадии АГ признаков использовали карту кодировки для 37 диагностических признаков. Этим признакам, в различной степени характерным для первой, второй и третьей стадии АГ, поставили в соответствие 37-мерный вектор, учитывающий наличие и величину каждого признака. При диагностике 3-х стадий АГ получили 4 области: 1 – точки, присущие только первой стадии АГ, 2 – только второй, 3 – только третьей и 4 – промежуточная область.

В случае дифференциации 3-х стадий заболевания необходимы две дискриминантные функции. Результаты определения канонических коэффициентов дискриминантных функций приведены в табл. 1.

Таким образом, получены следующие дискриминантные функции:

$$\begin{aligned} DF_1(X) &= 0,945 \cdot x_1 - 0,199 \cdot x_2 + 0,874 \cdot x_3 + \\ &+ 0,361 \cdot x_4 - 0,395 \cdot x_5 + 0,073 \cdot x_6 - 0,385 \cdot x_7 + \\ &+ 0,009 \cdot x_8 + 0,070 \cdot x_9 - 2,028, \\ DF_2(X) &= 0,408 \cdot x_1 + 0,011 \cdot x_2 + 0,859 \cdot x_3 - \\ &- 0,331 \cdot x_4 + 0,115 \cdot x_5 - 0,008 \cdot x_6 + 0,203 \cdot x_7 - \\ &- 0,008 \cdot x_8 + 0,195 \cdot x_9 - 3,328. \end{aligned} \quad (1)$$

На основе значений обеих дискриминантных функций построена территориальная карта (рис. 1).

Таблица 1

Канонические коэффициенты дискриминантных функций

Показатель	Переменная	Дискриминантная функция	
		1	2
стадия СН	x_1	0,945	0,408
Ожирение, степень	x_2	-0,199	0,011
индекс САГО	x_3	0,874	0,859
ХС ЛПВП, ммоль/л	x_4	0,361	-0,331
Апо В, г/л	x_5	-0,395	0,115
ИЛ-6, пг/мл	x_6	0,073	-0,008
наличие ИБС	x_7	-0,385	0,203
ФНО-альфа, пг/мл	x_8	0,009	-0,008
адипонектин, мкг/мл	x_9	0,070	0,195
Константа		-2,028	-3,328

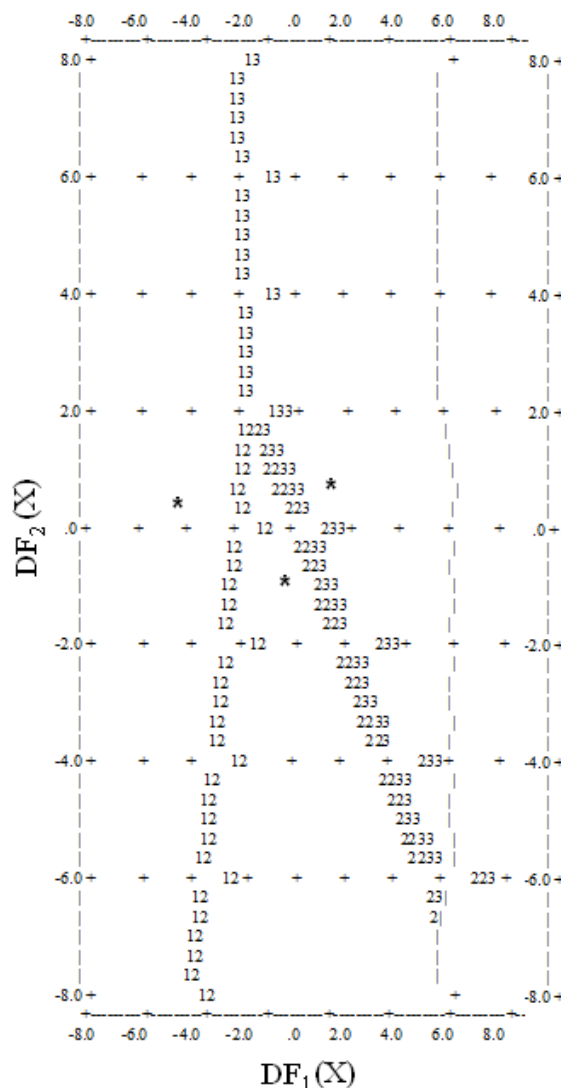


Рис. 1. Территориальная карта: 1 – 1-я стадия АГ, 2 – 2-я стадия АГ, 3 – 3-я стадия АГ, * - центр тяжести групп

На территориальной карте 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.

Таким образом, если отобразить состояние пациентов с различной стадией АГ на координатной плоскости ($DF_1(X)$, $DF_2(X)$) согласно рассчитанным значениям дискриминантных функций они образуют три класса (рис. 2).

Точность прогнозирования показана в табл. 2.

В колонке «Всего» указано общее количество случаев, которые фактически относятся к соответствующим группам: 74 пациентов с первой стадией АГ, 74 – со второй и 72 – с третьей. В колонках «Прогнозируемая принадлежность к группе» указано фактическое количество случаев, относящихся к каждой группе.

Так ко второй группе было ошибочно отнесено 7 пациентов первой группы и 3 пациента третьей группы; к третьей группе было ошибочно отнесено 8 пациентов второй группы, а все пациенты первой группы были определены безошибочно.

Анализ результатов классификации показал, что прогноз для групп 1 (первая стадия), 2 (вторая стадия) и 3 (третья стадия) проведен верно и корректно классифицированы, в общей сложности, 91,8% всех больных АГ. Соответственно, точность предсказания стадии АГ для обследуемых пациен-

тов может достигать 91,8% в том случае, если выборка пациентов будет идентична тем пациентам, данные которых послужили основой для прогноза.

Мерой удачного разделения на группы служат корреляционные коэффициенты между рассчитанными значениями дискриминантных функций и показателями принадлежности к группе, приведенные в табл. 3.

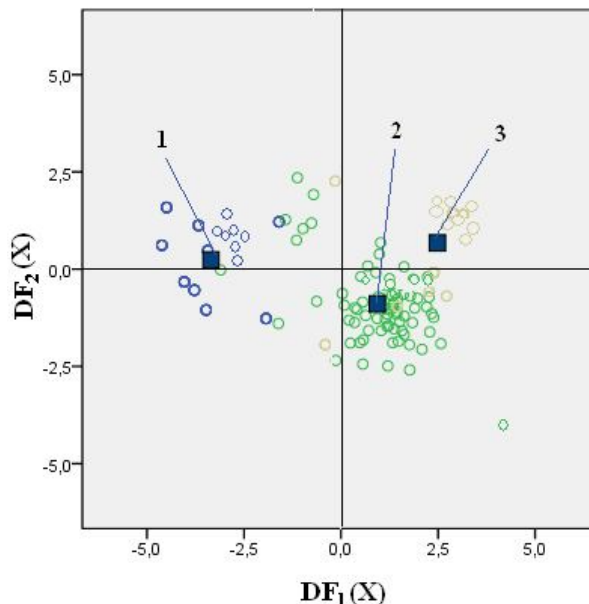


Рис. 2. Распределение пациентов с различной стадией АГ: 1 – 1-я стадия АГ, 2 – 2-я стадия АГ, 3 – 3-я стадия АГ

Таблица 2

Результаты классификации объектов

Первоначально	Тяжесть течения АД	Прогнозируемая принадлежность к группе			Всего
		Группа 1	Группа 2	Группа 3	
Количество	1 стадия	74	0	0	74
	2 стадия	7	64	3	74
	3 стадия	0	8	64	72
%	1 стадия	100,0	0,0	0,0	100,0
	2 стадия	9,5	86,5	4,1	100,0
	3 стадия	0,0	11,1	88,9	100,0

Таблица 3

Собственные значения

Функция	Собственные значения	% дисперсии	Совокупный %	Каноническая корреляция
1	6,172	93,3	93,3	0,928
2	0,442	6,7	100,0	0,554

Полученные собственные значения приведены в порядке убывания их величин. Величина собственного значения связана с дискриминантными возможностями функции: чем больше собственное значение, тем лучше различие. Таким образом, первая функция имеет наибольшие дискриминантные возможности, вторая – обеспечивает максимальную разницу после первой. Фактические числа, равные 6,172 и 0,442, указывают на то, что дискриминантные возможности первой дискриминантной функции в 14 раз больше другой.

Для того, чтобы определить, представляет ли интерес вторая функция, необходимо определить канонический коэффициент корреляции функций.

Судя по значениям коэффициентов, равные 0,928 и 0,554, существует достаточно высокая связь между артериальной гипертензией и значением первой дискриминантной функции, что и подтверждается процентом дисперсии этой функции (93,3 %). Также выявлена умеренная связь между стадиями артериальной гипертензии и значением второй дискриминантной функции.

Качественная оценка тесноты связи r_{xy} величин X и Y была обнаружена на основе шкалы Чеддока. Судя по значениям коэффициентов корреляции обе дискриминантные функции дают значимые результаты для дифференциации стадий АГ и разделения пациентов между группами и, следовательно, могут быть использованы соответствующим образом.

Тест, проведенный с помощью критерия «Лямбда Уилкса» (табл. 4), на предмет, значимо ли различаются между собой средние значения дискриминантных функций во всех группах, показал значимый результат (значение $p < 0,0001$). Чем ближе значение λ к 0, тем лучше различие классов, а чем ближе к 1, тем различия хуже (классы совпадают).

Таблица 4

Лямбда Уилкса

Тест функций	Лямбда Уилкса	df	Значимость
1 к 2	0,097	18	0,0001
2	0,693	8	0,0001

Таким образом, полученные дискриминантные функции позволяют провести корректную дифференциацию стадии артериальной гипертензии по 9 информативным признакам.

Заключение

Таким образом, применение дискриминантного анализа и метода дискриминантных функций позволило синтезировать математическую модель дифференциации стадии артериальной гипертензии, что

дало возможность классифицировать пациентов с разными стадиями артериальной гипертензии и повысить качество лечебно-диагностических мероприятий.

Впервые были использованы количественные характеристики результатов клинико-лабораторной диагностики изменений обмена углеводов, липидов, адипокинов, которые позволили объективизировать исследования и получить диагностическую информацию прогностически значимых показателей.

Литература

1. Landry, F. Exercise hypertension in the perspective of systemic arterial hypertension [Text] / F. Landry, M. Jette, G. Blumchen // *Herz*. – 1987. – Vol. 12. – P. 75–82.
2. Небиеридзе, Д.В. Мягкая артериальная гипертензия [Текст] / Д.В. Небиеридзе // *Русский медицинский журнал*. – 1997. – Т.5, №9. – С. 566-570.
3. Rakugi, H. Links between hypertension and myocardial infarction [Text] / H. Rakugi., H. Yu, A. Kamitani // *Am. Heart J.* – 1996. – Vol. 132, №1. – P. 213–221.
4. Nicholls, M. Hypertrophy, heart failure [Text] / M. Nicholls // *Heart*. – 1996. – Vol. 76. – P. 92–97.
5. Бритов, А.Н. Профилактика артериальной гипертензии на популяционном уровне: возможности и актуальные задачи [Текст] / А.Н Бритов // *Русский медицинский журнал*. – 1997. – Т.5, №9. – С. 571-576
6. Коваленко, В.М. Демографія и стан здоров'я народу України [Текст] / В.М.Коваленко В.М Корнацький. – К.: Медицина, 2010. – С. 31-50.
7. Кушаковский, М.С. Гипертоническая болезнь (эссенциальная гипертензия): Причины, механизмы, клиника, лечение [Текст] / М.С Кушаковский. – СПб.: СОТИС, 1995. – 309 с.
8. World Health Organization: Hypertension control [Text] // *Report of a WHO Expert Committee*. – Geneva: World Health Organization, 1996. – 104 p.
9. Safar, M.E. Arterial alterations in hypertension with a disproportionate increase in systolic over diastolic blood pressure [Text] / M.E. Safar, L. Cloarec-Blanchard, G.M. London // *J. Hypertens.* – 1996. – Vol. 14, №2. – P. 103–110.
10. Julius, S. Hyperkinetic borderline hypertension in Tecumseh [Text] / S. Julius, L. Kraus, N.J. Schork // *J. Hypertens.* – 1991. – Vol. 9. – P. 77–84.
11. SAS Institute Inc.: SAS/STAT user's guide, version 6.04 ed. [Text]. – Cary: SAS Institute Inc., 1993. – 464 p.
12. SAS Institute Inc.: SAS Technical Report SAS/STAT Software: CALIS and LOGISTIC procedures, release 6.04. [Text]. – Cary: SAS Institute Inc., 1990. – 236 p.
13. Андерсон, Т. Введение в многомерный статистический анализ [Текст] / Т. Андерсон. – М.: Феникс, 1963. – 516 с.

14. Дрейпер, Н. Прикладной регрессионный анализ [Текст] / Н. Дрейпер, Г. Смит. – М.: Издательский дом «Вильямс», 2007. – 118 с.

15. 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.

16. Дюк, В.А. Компьютерная психодиагностика [Текст] / В.А. Дюк. – СПб.: Питер, 1994. – 318 с.

17. Каримов, Р.Н. Обработка экспериментальной информации [Текст]: Уч. пособие в трех частях. – Ч.3. – Многомерный анализ. / Р.Н. Каримов. – Саратов: СГТУ, 2000. – 108 с.

18. Справочник по прикладной статистике [Текст] / под ред. Э. Ллойда. – М.: Финансы и статистика, 1990. – 526 с.

Поступила в редакцию 6.03.2012

Рецензент: д-р физ.-мат. наук, проф., заведующий кафедрой биомедицинских электронных устройств и систем А.И. Бых, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

ЗАСТОСУВАННЯ ДИСКРИМІНАНТНОГО АНАЛІЗУ ДЛЯ ДІАГНОСТИКИ СТАДІЙ АРТЕРІАЛЬНОЇ ГІПЕРТЕНЗІЇ

О.В. Висоцька, А.П. Порван, Д.А. Амбросов, Т.М. Амбросова

У статті авторами розглянуті існуючі методи синтезу математичної моделі на основі теорії ймовірності та цілком обгрунтоване застосування методу дискримінантних функцій. Запропонована діагностична модель диференціації стадій артеріальної гіпертензії, яка отримана в результаті дискримінантного аналізу даних клінічних, інструментальних та лабораторних досліджень. Отримані дискримінантні функції використовуються при побудові територіальної карти діагностування артеріальної гіпертензії, а статистичні оцінки синтезованої моделі показують її високу значущість і адекватність.

Ключові слова: дискримінантний аналіз, територіальна карта, артеріальна гіпертензія, стадія, диференціація ознак.

USING DISCRIMINANT ANALYSIS FOR DIAGNOSTIC OF ARTERIAL HYPERTENSION STAGES

E.V. Visotskaja, A.P. Porvan, D.A. Ambrosov, T.N. Ambrosova

In this article the authors review existing methods for the synthesis of mathematical models based on probability theory and fully justify the use of discriminant functions. Diagnostic model is proposed to differentiate the stages of hypertension resulting from discriminant analysis of clinical, instrumental and laboratory tests. The resulting discriminant functions are used to construct maps of the territorial diagnosis of hypertension and statistical evaluation of synthetic models show its great significance and value.

Key words: discriminant analysis, the territorial map, arterial hypertension, stage, signs of differentiation.

Высоцкая Елена Владимировна – канд. техн. наук, доцент, доцент каф. биомедицинских электронных устройств и систем, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Порван Андрей Павлович - канд. техн. наук, научный сотрудник каф. биомедицинских электронных устройств и систем, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

Амбросов Дмитрий Антонович – студент каф. биомедицинских электронных устройств и систем, Харьковский национальный университет радиоэлектроники, Харьков, Украина. e-mail: eambrosov@mail.ru

Амбросова Татьяна Николаевна – канд. мед. наук, доцент кафедры пропедевтики внутренней медицины №1, Харьковский национальный медицинский университет, Харьков Украина.