

ИСПОЛЬЗОВАНИЕ МЕТОДА ДЕРЕВЬЕВ КЛАССИФИКАЦИИ ДЛЯ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ ЗАБОЛЕВАНИЙ СИСТЕМЫ КРОВООБРАЩЕНИЯ У ПАЦИЕНТОВ

В.С.Глушанко, А.В.Шульмин, Л.И.Орехова

Витебский государственный ордена Дружбы народов медицинский университет,
пр. Фрунзе, 27, 210023, г. Витебск, Республика Беларусь

В статье изучены возможности использования метода деревьев классификации в прогнозировании развития патологии системы кровообращения у пациентов. На основании результатов социологического опроса построены таблицы и диаграммы, которые показывают иерархические зависимости, имеющиеся в системе медико-биологических индикаторов, отображающих разнообразные аспекты состояния пациента с заболеваниями системы кровообращения. Выявлены факторы, наиболее значимо способствующие развитию данной патологии: наличие недвижимости у пациента, неудовлетворительные условия труда при неизменности характера труда, изменение характера труда вследствие заболевания, удовлетворенность пациента взаимоотношениями в семье и наличие тяжелых периодов в семье, отсутствие определенного распорядка дня, размер одежды свыше 44, предположение болезни от сильных переживаний, окружность груди до 96 см включительно. Представлены диапазоны вариации индикаторов, соответствующие наиболее высокой степени развития кардионеврологических нарушений. Для обработки данных применялся пакет статистических программ SPSS Statistics 22 Trial.

Ключевые слова: болезни системы кровообращения; фактор риска; опросник; обработка многомерных статистических данных; метод деревьев классификации.

Введение

Хронические неинфекционные заболевания, среди которых ведущую позицию занимают болезни системы кровообращения (БСК), определяют на сегодняшний день уровень заболеваемости, трудовых потерь по болезни и смертности населения в развитых и многих развивающихся странах и «представляют собой главную проблему здравоохранения Европейского региона» [1]. Смертность от БСК в Республике Беларусь составляет более 65 тыс. человек (694 на 100 тыс.) ежегодно, из них 13 тыс. – в трудоспособном возрасте (136,3 на 100 тыс. трудоспособного возраста) [12], что обуславливает выявление и коррекцию причин и факторов риска (ФР) развития данной патологии.

Как известно, БСК представляют собой мультифакториальное заболевание [5], что подразумевает решение вопросов управления применительно к медико-биологическим системам в контексте данной патологии с необходимостью принятия решений в условиях многофакторности. Для этого требуется многомерный анализ связей и взаимозависимостей между факторами риска вышеуказанной патологии [14].

В связи с этим, представляется актуальным применение математических методов к анализу медико-биологических данных с целью восстановления взаимозависимостей между факторами БСК, характеризующими состояние здоровья пациента [14].

Возможности применения математических методов анализа в различных областях медицины исследуются в работах, посвященных поиску функциональных и логических закономерностей в кардиологии [2, 13], пульмонологии [3, 4, 8], неврологии [7, 10], хирургии [6].

Цель настоящего исследования – изучение возможностей применения метода деревьев решений в прогнозировании развития кардионеврологической патологии у пациентов с учетом влияния социально-дистрессовых ФР БСК по результатам социологического опроса.

Данная работа была проведена для оптимизации авторской комплексной методики опроса, по которой проведено анкетирование пациентов [11].

Материалы и методы

Метод деревьев решений является одним из методов многомерного анализа, позволяющим наглядно и удобно демонстрировать закономерности

сти [9]. Дерево решений – это определенным образом структурированные данные, включающие узлы принятия решений, соединенные друг с другом ребрами. Дерево состоит из корневого узла и листовых узлов, располагающихся в конце любой цепочки идущих подряд ребер. Оно конструируется на основе обучающей выборки, содержащей значения входных переменных и соответствующие значения прогнозируемого индикатора [14].

Узел принятия решений выполняет проверку условия на значение входной переменной, ребра отображают каждую из вероятных вариаций. В узлах дерево решений представляет значение прогнозируемого показателя в зависимости от значений входных переменных. Построение дерева решений направлено на расположение подмножества с уникальными свойствами прогнозируемого показателя на конце каждого из исходящих из узла принятия решений ребер [14]. Для наглядного представления возможностей выявления скрытых иерархических закономерностей применительно к данным медико-биологической природы было исследовано влияние переменных, характеризующих социально-дистрессовые детерминанты (ФР), на кардионеврологическую патологию.

Деревья классификации строились на основании данных, собранных при анкетировании пациентов в ГУЗ «Витебский областной клинический кардиологический центр», ГУЗ «Витебский областной клинический специализированный центр», ГУЗ «Витебская областная клиническая больница». Респонденты были проанкетированы по авторской комплексной методике опроса «Опросник для исследования социально-дистрессовых детерминант болезней системы кровообращения», разработанной аспирантом Л.И.Ореховой и профессором В.С.Глушанко и одобренной комитетом по этике клинических испытаний Витебского государственного медицинского университета (протокол от 14.11.2019 г. №9) [11]. Применительно к данному исследованию были сформированы исследуемая и контрольная группы пациентов с разным основным заболеванием (типом патологии) в возрасте от 18 до 72 лет. Для сопоставимости групп наблюдения по полу, возрасту в пакете статистических программ “IBM SPSS Statistics 22 Trial” использована процедура «Взвесить наблюдения» с целью подвести выборочные статистики к параметрам генеральной совокупности.

Основным критерием включения для пациентов исследуемой группы наблюдения явился установленный в «Медицинской карте стационарного пациента» (форма №003/у-07) основной диагноз «Ишемическая болезнь сердца» (ИБС) (па-

циенты кардиологических отделений), для контрольной группы – установленный основной диагноз «Хронический холецистит», «Язвенная болезнь (желудка и 12-перстной кишки)», «Хронический энтерит», «Гастроэзофагеальная рефлюксная болезнь» и т.д. (пациенты отделения гастроэнтерологии). Исследуемая группа наблюдения включала 93 кардиологических пациента, контрольная группа – 40 гастроэнтерологических пациентов.

Во время социологического опроса у каждого опрошенного были отмечены значения показателей, характеризующих комплексное состояние пациента с БСК. В число таких параметров включены показатели, описывающие возраст, пол, образование, социальную группу, профессию, семейное положение и количество детей у пациента на момент анкетирования (блок №1 опросника «Паспортная часть»); качество жизни пациента (блок №2 «Качество жизни»); факторы или ситуации сфер жизнедеятельности анкетированного из определенного перечня, имеющие место в его жизни и связанные с процессом труда/условиями проживания/жизнедеятельностью семьи/особенностями быта и вызывающие у него отрицательные эмоции, снижающие настроение и создающие нервно-психическую напряженность (блок №3 «Сферы жизнедеятельности»); антропометрические и лабораторные показатели, а также клинические проявления (симптомы) заболеваний (блок №4 «Социально-медицинский портрет респондента»); ФР патологии у пациента: злоупотребление алкоголем, курение, уровень физической активности, избыточная масса тела, питание, прогулки на свежем воздухе (блок №5 «Факторы, детерминирующие профилактические меры») [11].

Процедура построения деревьев классификации для прогнозирования развития БСК выполнялась средствами пакета статистических программ “IBM SPSS Statistics 22 Trial” на основании данных пилотной выборки численностью 133 наблюдения (93 кардиологических пациента + 40 гастроэнтерологических пациентов). В соответствии с требованиями алгоритма CHAID (Chi Squared Automatic Interaction Detection) был зафиксирован уровень значимости $\alpha=0,05$, также автоматически задано максимальное количество уровней $n=3$. В качестве критерия остановки алгоритма было установлено минимальное количество наблюдений в родительском узле $N_{\min}=20$ и минимальное количество наблюдений в дочернем узле $N_{\min}=10$. Критерий связи между категориальными переменными – хи-квадрат (χ^2).

Результаты и обсуждение

Первоначально в качестве независимых переменных влияния были выбраны вопросы из блока №2 «Качество жизни» (табл. 1), так как блок №1 содержит паспортные показатели, используемые в качестве переменных влияния.

Построенное дерево классификации выглядит следующим образом (рис. 1).

Четко выделяется следующая тенденция: с вероятностью 75,5% у пациента имеется кардиологическая патология при наличии у него недвижимого имущества ($p=0,016$, $\chi^2=5,793$, ст.св.=1).

В табл. 2 отражены доли верно предсказанных значений по данному дереву классификации, в котором общая процентная доля верно предсказанных результатов относительно высокая – 70,7%.

К числу особенностей рассматриваемого метода относится то, что одно дерево классификации не описывает всех существующих зависимостей. Оно отражает только наиболее значимые зависимости. При этом существенный эффект на формирование дерева решений оказывают выбор независимых и влияющих переменных, значения параметров метода, таких как уровень значимо-

Таблица 1

Свод переменных для модели дерева классификации по блоку №2 «Качество жизни»

Спецификации	Метод построения	CNAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Характеристика жилья, уровень дохода, наличие автомобиля, наличие недвижимости, причина изменения места работы, чередование труд/отдых, характеристика общественной жизни, наличие друзей на работе, наличие друзей среди родственников, наличие друзей среди других людей, участие в общественной жизни, участие в политической жизни, оценка экологической обстановки, характеристика здоровья в целом, удовлетворенность жизнью, ощущение безопасности
	Проверка	Нет
	Максимальное количество уровней	3
	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
Результаты	Включены независимые переменные	Наличие недвижимости
	Число узлов	3
	Число конечных узлов	2
	Глубина	1

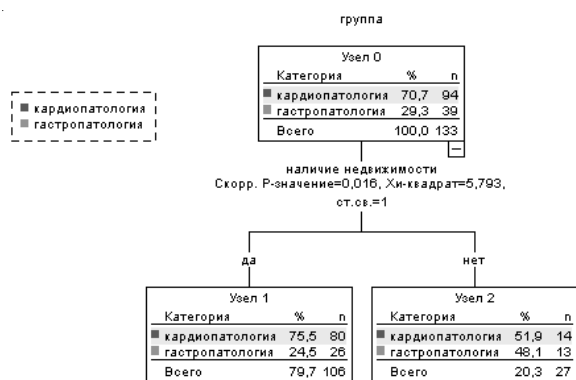


Рис. 1. Дерево классификации по блоку №2 «Качество жизни»

сти, критерий останковки процедуры роста дерева, изначальное количество категорий для непрерывных переменных.

Для демонстрации данного эффекта выберем другие независимые переменные – социально-дистрессовые детерминанты и факторы, связанные с процессом труда (блок №3 «Сферы жизнедеятельности») (табл. 2).

В результате, в число независимых переменных вошли новые факторы, и, как следствие, были получены новые классифицирующие правила (рис. 2).

В данном случае рост вероятности развития кардиологической патологии до 79,2% связан с

Свод переменных для модели дерева классификации по социально-дистрессовым факторам, связанным с процессом труда

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Причина изменения места работы, личная зарплата, изменение характера труда вследствие заболевания, проявления изменения характера труда, переутомление, напряженный темп работы, неудовлетворительные условия труда, недостаточное стимулирование труда, отсутствие работы по специальности, монотонный труд, недостаточные возможности повышения квалификации и продвижения по службе, конфликтные отношения с отдельными сотрудниками, конфликтные отношения с руководством, недостаточное количество социальных льгот, недостаточная самостоятельность в выполнении трудовых операций, несоответствие труда индивидуальным способностям и интересам, недостаточная престижность специальности, частые внеурочные работы, неудобный транспорт до предприятия, неудобный график работы, работа не по специальности, понижение по должности с более низкой зарплатой, дисциплинарное взыскание, осознание своей профессиональной непригодности
	Проверка	Нет
	Максимальное количество уровней	3
Результаты	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
	Включены независимые переменные	Изменение характера труда вследствие заболевания, неудовлетворительные условия труда
	Число узлов	5
	Число конечных узлов	3
	Глубина	2

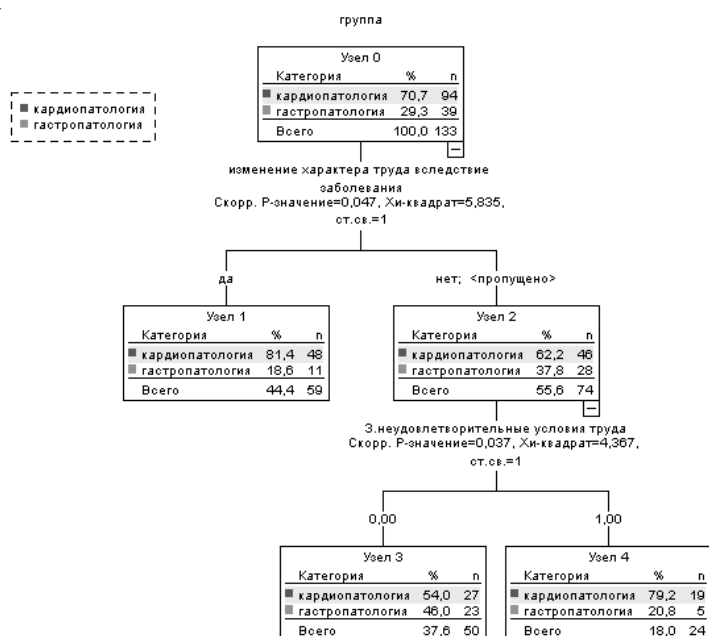


Рис. 2. Дерево классификации по социально-дистрессовым факторам, связанным с процессом труда*

* (здесь и далее): пропущенные значения – это неизвестные значения переменной или значения, помеченные респондентом как пропущенные.

тем, что пациент считает неудовлетворительные условия труда при неизменности характера труда фактором, имеющим место в его жизни и вызывающим у него отрицательные эмоции, снижающим настроение и создающим нервно-психическую напряженность ($p=0,037$, $\chi^2=4,367$, ст.св.=1). Если же у пациента имеется изменение характера труда вследствие заболевания, то с вероятностью 81,4% у него – кардиопатология ($p=0,047$, $\chi^2=5,835$, ст.св.=1).

При выборе следующей группы независимых переменных – социально-дистрессовые детерминанты и факторы, связанные с условиями проживания (блок №3 «Сферы жизнедеятельности»), были получены следующие результаты (табл. 3).

В данном случае дерево классификации не построено (рис. 3), что может означать отсутствие статистической значимости вопросов в анкете, зависящей

Таблица 3

Свод переменных для модели дерева классификации по социально-дистрессовым факторам, связанным с условиями проживания

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Переезд в город из сельской местности, проживание в стесненных жилищных условиях, частичное отсутствие квартирных удобств, частая смена места жительства, переезд в сельскую местность из города, вынужденное изменение места жительства, причины смены места жительства, удовлетворенность местом жительства, вид жилья, жилая площадь, какие жилищные условия
	Проверка	Нет
	Максимальное количество уровней	3
	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
Результаты	Включены независимые переменные	Независимые переменные не включены
	Число узлов	1
	Число конечных узлов	1
	Глубина	0



Рис. 3. Дерево классификации по социально-дистрессовым факторам, связанным с условиями проживания

логии связана с отсутствием удовлетворенности пациента от взаимоотношений в семье ($p=0,009$, $\chi^2=8,865$, ст.св.=1) и наличием тяжелых периодов в семье ($p=0,001$, $\chi^2=11,801$, ст.св.=1).

от нескольких причин: непонятная формулировка вопроса, пациенты при ответе на вопрос поняли его неверно, вопрос действительно статистически не значим.

Следующая выбранная нами группировка независимых переменных – социально-дистрессовые детерминанты и факторы, связанные с жизнедеятельностью семьи (блок №3 «Сферы жизнедеятельности»), представлена в табл. 4.

По результатам включения статистически значимых независимых переменных было построено следующее дерево классификации (рис. 4).

В данной ситуации отмечена связь увеличения вероятности развития кардиологической патологии до 84,8% с удовлетворенностью пациента взаимоотношениями в семье и наличием тяжелых периодов в семье ($p=0,009$, $\chi^2=8,865$, ст.св.=1). 50%-я вероятность развития кардиологической или гастроэнтерологической пато-

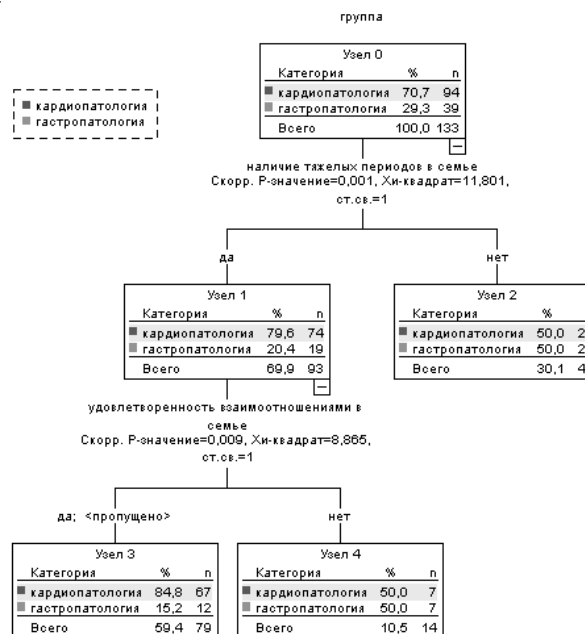


Рис. 4. Дерево классификации по социально-дистрессовым факторам, связанным с жизнедеятельностью семьи

Свод переменных для модели по социально-дистрессовым факторам, связанным с жизнедеятельностью семьи

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Наличие заболевания, измена, физическая травма, смерть близкого родственника, пьянство супруга, снижение полового влечения, уход из семьи супруга(и), развод, грубость и бестактность супруга (и), угроза распада семьи, подозрение в измене, вмешательство близких родственников в отношения между супругами, неравномерное распределение обязанностей в семье, угроза благополучию семьи, детей, невнимание к детям, частые конфликты, бесплодие супруга(и), наличие в семье часто болеющих, ревность, невнимание супруга(и), появление психологического отчуждения, тяжелое заболевание близкого родственника, наличие члена семьи, требующего постоянного ухода, отсутствие детей в семье, лишение свободы супруга(и), отсутствие социальной поддержки, отсутствие эмоциональной поддержки, различие во взглядах и интересах, нежелание супруга(и) иметь детей, разногласия по характеру проведения свободного времени, неудовлетворенность своей ролью как члена семьи, наличие психических заболеваний у членов семьи, кто воспитывал, наличие тяжелых периодов в семье, причина тяжелых периодов, количество лет в браке, причина развода, удовлетворенность взаимоотношениями в семье
	Проверка	Нет
	Максимальное количество уровней	3
	Минимальное количество наблюдений в родительском узле	20
Результаты	Минимальное количество наблюдений в дочернем узле	10
	Включены независимые переменные	Наличие тяжелых периодов в семье, удовлетворенность взаимоотношениями в семье
	Число узлов	5
	Число конечных узлов	3
	Глубина	2

Результаты анализа социально-дистрессовых детерминант и факторов, связанных с особенностями быта в блоке «Сферы жизнедеятельности», представлены в табл. 5.

На рис. 5 отражено дерево классификации по социально-дистрессовым факторам, связанным с особенностями быта, согласно которому предиктором развития кардиопатологии в 74,6% случаях можно считать фактор отсутствия определенного распорядка дня у пациента, вызывающий у него отрицательные эмоции и создающий нервно-психическую напряженность ($p=0,016$, $\chi^2=5,811$, ст.св.=1).

Следующую категорию независимых переменных, позволяющих прогнозировать развитие кардиологической патологии, составляет блок воп-

росов №4 «Социально-медицинский портрет респондента» (табл. 6).

При построении данного дерева классификации (рис. 6) переменными, значительно влияющими на высокую вероятность развития кардиопатологии, являлись: размер одежды свыше 44 в 77,5% случаев ($p<0,01$, $\chi^2=20,906$, ст.св.=1), предположение болезни от сильных переживаний в 90,3% случаев ($p<0,01$, $\chi^2=12,424$, ст.св.=1), окружность груди до 96 см включительно ($p=0,016$, $\chi^2=10,165$, ст.св.=1).

Последний анализируемый блок №5 независимых переменных составили вопросы о факторах, детерминирующих профилактические меры (табл. 7).

Таблица 5

Свод переменных для модели по социально-дистрессовым факторам, связанным с особенностями быта

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Ухудшение сна, самолечение, появление чувства одиночества, неудачная любовь, неуверенность в будущем, прерванное образование, однообразие форм досуга, появление чувства отчужденности к окружающим, уменьшение культурных мероприятий, принижение значения Вас как личности, учащение курения, недостаточная физическая активность, срыв жизненных планов, неудача в достижении цели, отсутствие регулярного режима питания, недостаточно хорошие условия для питания, уменьшение времени на досуг, ограничение межличностных контактов, отсутствие определенного распорядка дня, неприятности в быту, увеличение времени на домашнюю работу, сокращение прогулок, сокращение хобби, учащение потребления алкоголя, совершение правонарушений, появление конфликтов вне работы, лишение свободы, финансовые проблемы в быту, следование распорядку дня, характер сна, длительность сна
	Проверка	Нет
	Максимальное количество уровней	3
	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
Результаты	Включены независимые переменные	Отсутствие определенного распорядка дня
	Число узлов	3
	Число конечных узлов	2
	Глубина	1



Рис. 5. Дерево классификации по социально-дистрессовым факторам, связанным с особенностями быта

Как и в ситуации с факторами, связанными с условиями проживания, непостроенное дерево классификации (рис. 7) может предполагать отсутствие статистической значимости переменных (вопросов в блоке опросника), так же обусловленное нечетким изложением или неправильным пониманием сущности вопроса,

или действительно его статистической недостоверностью.

Таким образом, в ходе исследования был подтвержден тезис о высоком уровне влияния на развитие БСК вышеуказанных индикаторов, сформулированных в виде вопросов в авторском опроснике. В последующих работах предполагается проведение более детального анализа и выбор как зависимых и независимых переменных, так и переменных влияния, которые могут несколько модулировать полученные результаты.

Заключение

Метод деревьев решений представляет собой эффективный инструмент обработки медико-биологических данных, что было подтверждено его применением при прогнозировании развития кардионеврологической патологии у пациентов с учетом влияния социально-дистрессовых ФР БСК по данным социологического опроса.

В результате исследования выявлены отдельные факторы и их совокупности (сочетания), оказывающие наиболее значимое влияние на процесс

Таблица 6

Свод переменных для модели дерева классификации по социально-медицинскому портрету респондента

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Раннее начало ишемической болезни сердца у близких родственников, рост, масса тела, размер бюста, окружность груди, окружность талии (примерная), окружность бедер (примерная), размер одежды, систолическое и диастолическое артериальное давление, прием гипотензивных лекарственных средств, контроль артериального давления и пульса, наличие атерогенной дислипидемии, наличие сопутствующих заболеваний, наличие сахарного диабета, контроль и уровень глюкозы, контроль и уровень холестерина, наличие атеросклероза сосудов, прием оральных контрацептивов (жен.), наличие мигрени, наличие синдрома обструктивного апноэ, наличие транзиторных ишемических атак, наличие кардиологической или неврологической патологии у родственников, болезнь от сильных переживаний, отношение к здоровью, меры для сохранения здоровья
	Проверка	Нет
	Максимальное количество уровней	3
Результаты	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
	Включены независимые переменные	Размер одежды, болезнь от сильных переживаний, окружность груди (жен.)
	Число узлов	7
	Число конечных узлов	4
	Глубина	3

развития БСК у пациентов. В частности, обнаружено, что значительное влияние на развитие БСК оказывают: наличие недвижимости у пациента, неудовлетворительные условия труда как негативный фактор, вызывающий отрицательные эмоции и создающий нервно-психическую напряженность при неизменности характера труда, изменение характера труда вследствие заболевания, отсутствие удовлетворенности пациента взаимоотношениями в семье и наличие тяжелых периодов в семье, отсутствие определенного распорядка дня как негативный эмоциональный фактор, создающий нервно-психическую напряженность, размер одежды свыше 44, предположение болезни от сильных переживаний, окружность груди до 96 см включительно.

Алгоритм поиска значимых факторов метода деревьев решений позволил значительно сократить число входных анализируемых параметров, что является важным в процессе принятия решений при выборе профилактической стратегии и разработке концепции ФР. Со-

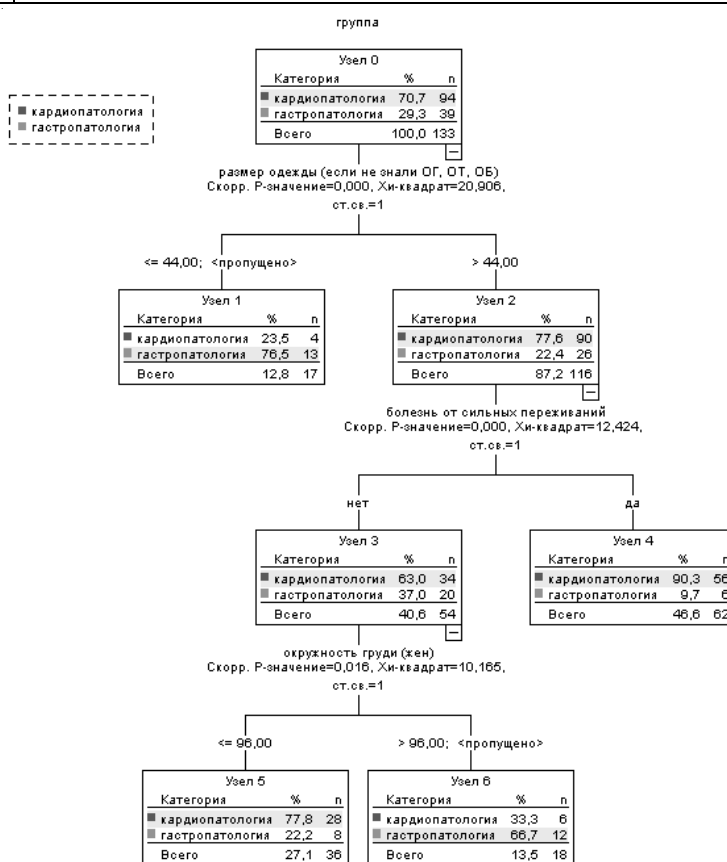


Рис. 6. Дерево классификации по социально-медицинскому портрету респондента

Таблица 7

Свод переменных для модели дерева классификации по факторам, детерминирующим профилактические меры

Спецификации	Метод построения	CHAID
	Зависимая переменная	Группа (тип патологии)
	Независимые переменные	Употребление алкоголя, дозы алкоголя за сутки, 6 и более доз за 1 случай выпивки, невозможность прекратить пить, невозможность выполнить то, что необходимо из-за выпивки, необходимость выпить, чтобы начать день после выпивки, чувство вины после выпивки, травмы в результате выпивки, чувство тревоги из-за выпивки у людей, курение, количество сигарет в день, стаж курения, отказ от курения, интенсивность физической активности, физические нагрузки 30 минут большая часть дней в неделю, длительность физической активности в неделю, кратность физической активности в неделю, заедание стресса, приведение индекса массы тела (ИМТ) в норму, наличие разнообразного питания, употребление фруктов, употребление овощей, употребление хлеба грубого помола, употребление нежирной рыбы, употребление постного мяса, употребление морепродуктов, употребление обезжиренных молочных продуктов, ограничение соли, ограничение сахара, наличие прогулок, длительность прогулок, кратность прогулок, ИМТ
	Проверка	Нет
	Максимальное количество уровней	3
	Минимальное количество наблюдений в родительском узле	20
	Минимальное количество наблюдений в дочернем узле	10
Результаты	Включены независимые переменные	Независимые переменные не включены
	Число узлов	1
	Число конечных узлов	1
	Глубина	0



Рис. 7. Дерево классификации по факторам, детерминирующим профилактические меры

гласно полученным данным, проведена оптимизация авторской комплексной методики опроса «Опросник для исследования социально-дистрессовых детерминант болезней системы кровообращения» Л.И.Ореховой и В.С.Глушанко путем сокращения количества и модификации сущности вопросов для последующего анкетирования.

Применение полученных результатов для скринингового отбора группы риска БСК позволит начать профилактическое воздействие на ранних этапах.

В перспективе планируется использование метода деревьев классификации в рамках создания информационно-аналитической системы прогнозирования БСК у населения. Кроме того, представляет интерес дальнейшее изучение медико-биологической природы выявленных зависимостей текущего уровня развития БСК от факторов, провоцирующих появление и характеризующих развитие данной патологии.

Литература

1. Cardiovascular diseases [Electronic resource] / World Health Organization, Regional office for Europe, 2020. – Mode of access: <http://www.euro.who.int/ru/home>. – Date of access: 04.06.2020.
2. Chapter 11. Classification Algorithms and Regression Trees [Electronic resource] – 2020. – Mode of access: <https://rafalab.github.io/pages/649/section-11.pdf>. – Date of access: 04.06.2020.
3. Безруков, Н.С. Автоматизированная система диагностики заболеваний легких / Н.С.Безруков,

- Е.Л.Еремин, Ю.М.Перельман // Проблемы управления. – 2007. – №5. – С.75–80.
4. Безруков, Н.С. Построение и моделирование адаптивной нейро-нечеткой системы в задаче медицинской диагностики / Н.С.Безруков, Е.Л.Еремин // Информатика и системы управления. – 2005. – №2 (10). – С.36–46.
 5. Глушанко, В.С. Анализ проблемы распространенности модифицируемых факторов риска развития болезней системы кровообращения (обзорная статья) / В.С.Глушанко, Л.И.Орехова // Современные проблемы здравоохранения и медицинской статистики: электронный научно-практический рецензируемый журнал. – 2019. – №2. – С.363–380.
 6. Драгун, И.А. Автоматизированная система количественной оценки операционного риска: автореф. дис. ... канд. тех. наук / И.А.Драгун. – Барнаул, 2006. – 22 с.
 7. Карасева, Т.С. Решение задач медицинской диагностики методами интеллектуального анализа данных / Т.С.Карасева // Решетневские чтения. – 2015. – Т.2. – С.46–47.
 8. Колосов, В.П. Пути построения прогнозных моделей в пульмонологии / В.П.Колосов, Ю.М.Перельман, Н.В.Ульянычев // Информатика и системы управления. – 2005. – №2 (10). – С.64–71.
 9. Методы и модели анализа данных: OLAP и Data Mining / А.А.Барсегян [и др.]. – СПб.: БХВПетербург, 2004.
 10. Нафтулин, И.С. Применение алгоритмов Data Mining для решения задачи диагностики типа инсульта / И.С.Нафтулин, О.Ю.Реброва // Научная сессия МИФИ. – 2008. – Т.10. – С.73.
 11. Орехова, Л.И. Опросник для исследования социально-дистрессовых детерминант болезней системы кровообращения: метод. рекомендации / Л.И.Орехова, В.С.Глушанко / под ред. проф. В.С.Глушанко. – Витебск: ВГМУ, 2019. – 21 с.
 12. Официальная статистика. Здравоохранение [Электронный ресурс] / Национальный статистический комитет Республики Беларусь. – Минск, 2020. – Режим доступа: https://www.belstat.gov.by/ofitsialnaya-statistika/solialnaya-sfera/zdravoohranenie_2/. – Дата доступа: 04.06.2020.
 13. Суфиянов, В.Г. Разработка адаптивных статистических моделей классификации и прогнозирования: автореф. дис. ... канд. физ.-мат. наук / В.Г.Суфиянов. – Пермь, 2004.
 14. Ташкинов, А.А. Применение метода деревьев классификации к прогнозированию уровня развития моторики у больных с нарушениями двигательных функций / А.А.Ташкинов, А.В.Вильдеман, В.А.Бронников // Российский журнал биомеханики. – 2008. – Т.12, №4 (42). – С.84–95.

METHOD OF CLASSIFICATION TREES USE TO PREDICT DEVELOPMENT OF CIRCULATORY SYSTEM DISEASES IN PATIENTS

V.S.Glushanko, A.V.Shulmin, L.I.Orekhova
Vitebsk State Order of Peoples' Friendship Medical University, 27, Frunze Ave., 210023, Vitebsk, Republic of Belarus

Opportunities for method of classification trees use in predicting development of circulatory system diseases (CSD) in patients are studied in the article. Tables and diagrams, basing on results of sociological survey, show hierarchical dependencies available in system of biomedical indicators, reflecting various aspects of condition of patient with CSD. Factors, most significantly contributing in CSD development, were identified: real estate existence, unsatisfactory working conditions with unchanged nature of work, changes in nature of work due to illness, patient's satisfaction with family relationships and presence of difficult periods in patient's family, specific daily routine lack, patient's dress size over 44, assumption of illness from intense experiences, patient's chest circumference up to and including 96 cm. Ranges of indicators variation, corresponding to the highest degree of development of cardio-neurological disorders, are presented. SPSS Statistics 22 Trial used to process the data.

Keywords: circulatory system diseases; risk factor; questionnaire; processing of multivariate statistical data; method of classification trees.

Сведения об авторах:

Глушанко Василий Семёнович, д-р мед. наук, профессор, УО «Витебский государственный ордена Дружбы народов медицинский университет», зав. кафедрой общественного здоровья и здравоохранения с курсом ФПК и ПК; тел. (+37529) 7104955; e-mail: glushanko@mail.ru.

Шульмин Андрей Владимирович, д-р мед. наук, профессор; УО «Витебский государственный ордена Дружбы народов медицинский университет», кафедра общественного здоровья и здравоохранения с курсом ФПК и ПК; тел.: (+37533) 3128140; e-mail: andrey.shulmin@yandex.ru.

Орехова Любовь Игоревна, магистр медицинских наук; УО «Витебский государственный ордена Дружбы народов медицинский университет», кафедра общественного здоровья и здравоохранения с курсом ФПК и ПК, ассистент, аспирант; тел.: (+37529) 2972964; e-mail: orexoval@list.ru.

Поступила 08.06.2020 г.