

Прогнозирование особенностей течения хронического гепатита С с использованием байесовских сетей

Л.М. Самоходская¹, Е.Е. Старостина¹, А.В. Сулимов¹, Т.Н. Краснова¹, Т.П. Розина^{1,2}, В.Г. Авдеев¹, И.А. Савкин¹, В.Б. Сулимов¹, Н.А. Мухин^{1,2}, В.А. Ткачук¹, В.А. Садовничий¹

¹ФГБОУ ВО «Московский государственный университет им. М.В. Ломоносова», Москва, Россия;

²ФГАОУ ВО «Первый Московский государственный медицинский университет им. И.М. Сеченова» Минздрава России (Сеченовский Университет), Москва, Россия

Резюме

Материалы и методы. В исследование включено 253 больных хроническим гепатитом С (ХГС) и циррозом печени (ЦП), у которых исследовались точечные мутации генов, участвующих в воспалительных реакциях и противовирусном иммунитете (*IL-1β-511C/T*, *IL-10-1082G/A*, *IL28B C/T*, *IL28B T/G*, *TNF-α-238G/A*, *TGF-β-915G/C*, *IL-6-174G/C*), активаторов локального печеночного фиброза (*AGT G-6A*, *AGT 235 M/T*, *ATRI 1166 A/C*), гемохроматоза (*HFE C282Y*, *HFE H63D*), тромбоцитарных рецепторов (*ITGA2 807 C/T*, *ITGB3 1565 T/C*), белков свертывающей системы и эндотелиальной дисфункции (*FII 20210 G/A*, *FV 1691G/A*, *FVII 10976 G/A*, *FXIII 103 G/T*, *eNOS 894 G/T*, *CYBA 242 C/T*, *FBG -455 G/A*, *PAI -675 5G/4G*, *MTHFR 677 C/T*). С использованием байесовских сетей (БС) изучалось предикторное значение клинико-лабораторных факторов для следующих состояний – конечных точек (КТ): развитие цирроза печени (КТ1), скорость фиброза (КТ2), наличие портальной гипертензии (КТ3) и наличие криоглобулинов (КТ4).

Результаты и обсуждение. Кроме традиционных факторов нами был показан вклад следующих мутаций. При прогнозировании КТ1 – ЦП – значимыми оказались наличие мутации гена *HFE H63D*, *C282Y*, *CYBA 242 C/T*, *AGT G-6G*, *ITGB31565 T/C*. Нами также выявлена связь между темпом прогрессирования фиброза печени и наличием полиморфизма генов *AGT G-6G*, *AGT M235T*, *FV 1691G/A*, *ITGB31565 T/C*. Среди генетических факторов, связанных с портальной гипертензией, оказались полиморфизм генов *PAI-I-675 5G/4G*, *FII 20210 G/A*, *CYBA 242 C/T*, *HFE H63D* и *IL-6 174G/C*. Наличие криоглобулинов и криоглобулинемический васкулит (КТ4) ассоциировано с мутациями в гене *MTHFR C677T*, генах *ATRI1166C* и *HFE H63D*.

Заключение. Полученные результаты позволили уточнить вклад основных патофизиологических и генетических факторов, определяющих статус больного и исход заболевания, а также выявить значимость точечных мутаций генов, контролируемых основные пути формирования и прогрессирования ХГС.

Ключевые слова: хронический гепатит С, полиморфизм генов, тромбофилия, тромбоцитарные рецепторы, байесовские сети.

Для цитирования: Самоходская Л.М., Старостина Е.Е., Сулимов А.В. и др. Прогнозирование особенностей течения хронического гепатита С с использованием байесовских сетей. *Терапевтический архив.* 2019; 91 (2): 32–39. DOI: 10.26442/00403660.2019.02.000076

Prediction of features of the course of chronic hepatitis C using Bayesian networks

L.M. Samokhodskaya¹, E.E. Starostina¹, A.V. Sulimov¹, T.N. Krasnova¹, T.P. Rosina^{1,2}, V.G. Avdeev¹, I.A. Savkin¹, V.B. Sulimov¹, N.A. Mukhin^{1,2}, V.A. Tkachuk¹, V.A. Sadovnichii¹

¹M.V. Lomonosov Moscow State University, Moscow, Russia;

²I.M. Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation (Sechenov University), Moscow, Russia

Materials and methods. 253 patients with chronic hepatitis C (CHC) and liver cirrhosis were included in the study. Assessment of gene polymorphisms of genes involved in inflammatory reactions and antiviral immunity (*IL-1β-511C/T*, *IL-10-1082G/A*, *IL28B C/T*, *IL28B T/G*, *TNF-α-238G/A*, *TGF-β-915G/C*, *IL-6-174G/C*), activators of local hepatic fibrosis (*AGT G-6A*, *AGT 235 M/T*, *ATRI 1166 A/C*), hemochromatosis (*HFE C282Y*, *HFE H63D*), platelet receptors (*ITGA2 807 C/T*, *ITGB3 1565 T/C*), coagulation proteins and endothelial dysfunction (*FII 20210 G/A*, *FV 1691G/A*, *FVII 10976 G/A*, *FXIII 103 G/T*, *eNOS 894 G/T*, *CYBA 242 C/T*, *FBG -455 G/A*, *PAI-675 5G/4G*, *MTHFR 677 C/T*) was carried. Using Bayesian networks we studied the predictor value of clinical and laboratory factors for the following conditions – end points (EP): development of cirrhosis (EP1), fibrosis rate (EP2), presence of portal hypertension (EP3) and cryoglobulins (EP4).

Results and discussion. In addition to traditional factors we have shown the contribution of the following mutations. Predicting EP1- liver cirrhosis – *HFE H63D*, *C282Y*, *CYBA 242 C/T*, *AGT G-6G*, *ITGB31565 T/C* gene mutations were significant. We also found a link between the rate of progression of liver fibrosis and gene polymorphisms of *AGT G-6G*, *AGT M235T*, *FV 1691G/A*, *ITGB31565 T/C*. Among the genetic factors associated with portal hypertension there are gene polymorphisms of *PAI-I-675 5G/4G*, *FII 20210 G/A*, *CYBA 242 C/T*, *HFE H63D* and *IL-6 174G/C*. Cryoglobulins and cryoglobulinemic vasculitis (EP4) are associated with gene mutations *MTHFR C677T*, *ATRI1166C* and *HFE H63D*.

Conclusion. The results obtained allow to detect the major pathophysiological and genetic factors which determine the status of the patient and the outcome of the disease, to clarify their contribution, and to reveal the significance of point mutations of genes that control the main routes of HCV course and progression.

Keywords: chronic hepatitis C, gene polymorphism, thrombophilia, platelet receptors, Bayesian networks.

For citation: Samokhodskaya L.M., Starostina E.E., Sulimov A.V., et al. Prediction of features of the course of chronic hepatitis C using Bayesian networks. *Therapeutic Archive.* 2019; 91 (2): 32–39. DOI: 10.26442/00403660.2019.02.000076

АСТ – аспаргатаминотрансфераза
АФП – альфа-фетопротейн
БС – байесовские сети
ВИЧ – вирус иммунодефицита человека
ВРВП – варикозное расширение вен пищевода
ИМТ – индекс массы тела
КТ – конечные точки

ПДРФ – полиморфизм длин рестриктных фрагментов
ПЦР – полимеразная цепная реакция
РФ – ревматоидный фактор
ХГВ – хронический гепатит В
ХГС – хронический гепатит С
ЦП – цирроз печени

Введение

Течение хронического гепатита С (ХГС) подвержено значительным индивидуальным различиям, что обусловлено, помимо прочего, генетическими факторами. Стратегия лечения ХГС опирается на индивидуальный подход к пациенту. Современные противовирусные препараты позволяют добиться впечатляющих результатов. Вместе с тем, в первую очередь из-за длительного латентного течения заболевания, приводящего к поздней диагностике в период уже развившихся осложнений, а также из-за нежелательных эффектов препаратов требуется дальнейший поиск предикторов неблагоприятных исходов заболевания. Процесс ремоделирования печеночной ткани очень сложен, хронический гепатит прогрессирует от мягкого воспаления к более тяжелому, а затем к фиброзу и циррозу. Участвуют как иммуновоспалительные, так и неиммунные реакции. В этот процесс вовлечено множество медиаторов. Индивидуальность каждого клинического случая, несомненно, зависит от точечных полиморфизмов генов, контролирующих эти механизмы. Однако вклад их оценить сложно, данные исследований противоречивы. Ранее мы изучали полиморфизмы генов, влияющие на выраженность воспалительной реакции, эндотелиальной дисфункции и ангиотензинпревращающего фермента, свертывающей системы, а также гемохроматоза [1]. С использованием методов логистической регрессии удалось уточнить влияние клинико-генетических факторов на скорость развития фиброза, портальной гипертензии и исходы терапии [1]. Показана связь полиморфизма генов и прогрессирования фиброза, портальной гипертензии, развития криоглобулинемического васкулита. Однако индивидуальный прогноз на основе всех ранее накопленных нами данных не проводился.

Прогностические вероятностные модели на основе байесовских сетей (БС) используются в кардиологии и других областях медицины [2], в том числе в онкологии, где многофакторный характер заболевания создает трудности прогнозирования рецидива заболевания, но именно инди-

видуальный прогноз во многом определяет стратегию лечения [3]. Модели на основе БС [2–5] обладают гибкостью и универсальностью, способны к самообучению и интеграции разнородных данных, а также они нечувствительны к наличию ошибочных или неполных данных. Попытка использования БС для определения прогностической значимости полиморфизма генов эндотелиальной дисфункции в течении ХГС предпринималась нами ранее. Представляется актуальным построение вероятностных моделей на основе БС с их оптимизацией для определения особо важных прогностических параметров, влияющих на течение ХГС.

Материалы и методы

В исследование включено 253 больных ХГС и циррозом печени (ЦП), наблюдавшихся в Клинике нефрологии, внутренних, профессиональных заболеваний и пульмонологии им. Е.М. Тареева (директор – академик РАН Н.А. Мухин) в период с ноября 2009 г. по июнь 2014 г. Создана база данных, которая включала основные клинические, биохимические, инструментальные параметры, характеризующие факторы вируса и хозяина, позволяющие оценить клиническое течение заболевания и исходы. Критериями включения служили наличие ХГС или ЦП в его исходе, принадлежность к европеоидной расе, а также подписанное информированное согласие на участие в исследовании.

Критериями исключения больных являлись: 1) употребление алкоголя (>20 мл/сут для женщин и >40 мл/сут для мужчин), 2) наличие дополнительных этиологических факторов поражения печени (коинфекция ХГВ, ВИЧ, болезнь Вильсона–Коновалова, аутоиммунный гепатит, наследственный гемохроматоз), 3) сахарный диабет 1-го и 2-го типа.

Для определения стадии фиброза больным выполняли биопсию печени (119 больных) с последующим гистологическим исследованием; 52 больным выполнена эластометрия печени с помощью аппарата FibroScan (EchoSens, Франция). У 52 больных стадия фиброза расценена как F4 на основании данных клинико-лабораторного и инструментального обследования без проведения эластометрии и биопсии печени. Скорость прогрессирования фиброза оценивали по формуле, предложенной Т. Pounard и соавт. [6]:

$$\text{Скорость прогрессирования фиброза печени (ед. фиброза/год)} = \frac{\text{стадия фиброза по METAVIR (ед. фиброза)} / \text{длительность инфицирования (год)}}{1}$$

Исследовались точечные мутации следующих генов:

1. Участвующих в воспалительных реакциях и противовирусном иммунитете – *IL-1 β -511C/T*, *IL-10 -1082G/A*, *IL28B C/T*, *IL28B T/G*, *TNF- α -238G/A*, *TGF- β -915G/C*, *IL-6 -174G/C*.
2. Активаторов локального печеночного фиброза – *AGT G-6A*, *AGT 235 M/T*, *ATRI 1166 A/C*.
3. Гемохроматоза – *HFE C282Y*, *HFE H63D*.
4. Тромбоцитарных рецепторов – *ITGA2 807 C/T*, *ITGB3 1565 T/C*.
5. Белков свертывающей системы и эндотелиальной дисфункции – *FII 20210 G/A*, *FV 1691G/A*, *FVII 10976 G/A*,

Контактная информация:

Старостина Екатерина Евгеньевна – врач-гастроэнтеролог, медицинский научно-образовательный центр ФГБОУ ВО «МГУ им. М.В. Ломоносова»; тел.: +7(926)719-23-10; e-mail: starostinaee@gmail.com

Сведения об авторах:

Самоходская Лариса Михайловна – к.м.н., доц., Медицинский научно-образовательный центр ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Сулимов Алексей Владимирович – ведущий программист, Научно-исследовательский вычислительный центр ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Краснова Татьяна Николаевна – к.м.н., доц., факультет фундаментальной медицины ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Розина Тэона Павловна – к.м.н., доц., факультет фундаментальной медицины ФГБОУ ВО «МГУ им. М.В. Ломоносова», ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова»

Авдеев Владимир Георгиевич – к.м.н., доц., факультет фундаментальной медицины ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Савкин Игорь Алексеевич – программист, Научно-исследовательский вычислительный центр ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Сулимов Владимир Борисович – д.ф.-м.н., зав. лаб., Научно-исследовательский вычислительный центр ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Мухин Николай Алексеевич – д.м.н., проф., академик РАН, факультет фундаментальной медицины ФГБОУ ВО «МГУ им. М.В. Ломоносова», ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова»

Ткачук Всеволод Арсеньевич – д.б.н., проф., академик РАН, декан факультета фундаментальной медицины ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Садовничий Виктор Антонович – д.ф.-м.н., проф., академик РАН, ректор ФГБОУ ВО «МГУ им. М.В. Ломоносова»

Таблица 1. Формулировка КТ

КТ	Целевые показатели/Общее кол-во	КТ = 0/Кол-во		КТ = 1/Кол-во	
		Нет	Есть	Нет	Есть
КТ1	Наличие ЦП	253	164	Нет	89
КТ2	Скорость фиброза	187	89	Медленная	98
КТ3	Наличие портальной гипертензии	89	13	Нет	76
КТ4	Наличие криоглобулинов	220	138	Нет	82

FXIII 103 G/T, eNOS 894 G/T, CYBA 242 C/T, FBG -455 G/A, PAI -675 5G/4G, MTHFR 677 C/T.

Выделение геномной ДНК проводили из ЭДТА-стабилизированной периферической венозной крови согласно протоколу с помощью коммерческого набора QIAmp DNA Blood Mini Kit и автоматической станции QIAcube™ (QIAGEN, Германия).

Полиморфные аллели генов цитокинов определяли методом полиморфизма длин рестриктных фрагментов (ПДРФ), а гена *HFE* – аллельспецифичной полимеразной цепной реакции (ПЦР) в термоциклере MasterCycler фирмы Eppendorf (Германия). Полиморфизм генов *MTHFR 677 C/T, FII 20210 G/A, FV 1691G/A, FVII 10976 G/A, FXIII 103 G/T, ITGA2 807 C/T, ITGB3 1565 T/C, FBG -455 G/A, PAI -675 5G/4G* определяли с помощью коммерческих наборов «ДНК-технология», термоциклера DTrime и программного обеспечения для приборов ДТ-96 («ДНК-технология», Россия).

Предикторное значение клинико-лабораторных, в том числе и генетических, факторов исследовалось для следующих четырех прогнозируемых состояний (табл. 1) – конечных точек (КТ): развитие цирроза печени (КТ1), скорость фиброза (КТ2), наличие портальной гипертензии (КТ3) и наличие криоглобулинов (КТ4). Целевые числа в колонках представляют собой общее количество пациентов из числа пациентов с соответствующим состоянием конечной точки (целевого показателя).

Статистический анализ. Прогнозирование проводилось с использованием БС. Каждая БС – это направленный

ациклический граф, каждой вершине (узлу) которого сопоставлена одна переменная и таблица условных вероятностей [6]. Переменные, соответствующие узлам БС, могут представлять собой параметры пациентов, КТ и другую информацию. Каждый узел БС может находиться в некотором количестве состояний, соответствующих дискретным значениям переменной, которую этот узел представляет. Вероятность состояния данного узла зависит только от текущего состояния его «родителей» – узлов, из которых исходят стрелки, идущие к данному узлу. Таблица условных вероятностей данного узла содержит вероятности наблюдений состояний этого узла при различных состояниях узлов-родителей. Таблицы получаются в результате выполнения процедуры обучения БС, а процедура опроса обученной БС позволяет определить вероятности состояний узлов в зависимости от имеющейся информации о значениях других узлов.

Использовалась наивная топология сети, в которой у всех узлов есть один общий родитель, называемый корневым узлом, в качестве которого выбиралась одна из КТ. Для определения надежности (качества) предсказания с помощью БС и сравнения различных БС использовали ROC-кривую [7–9] и численное значение площади под этой кривой AUC (Area Under the Curve): чем ближе величина AUC к единице, тем выше надежность предсказания сети. В данной работе для построения ROC-кривой для заданной БС применялся наиболее точный метод «исключения по одному».

С целью повышения надежности предсказаний БС проводилась ее оптимизация с целевой функцией – величиной

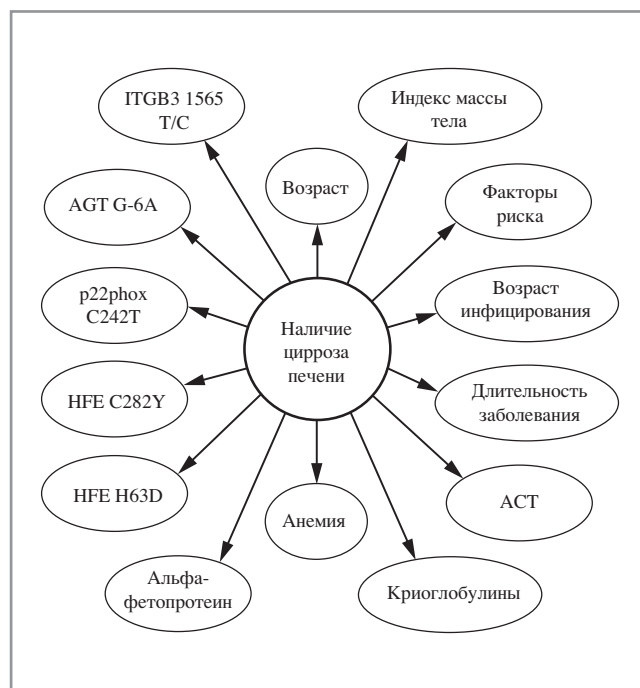


Рис. 1. БС 14+1 для КТ1.

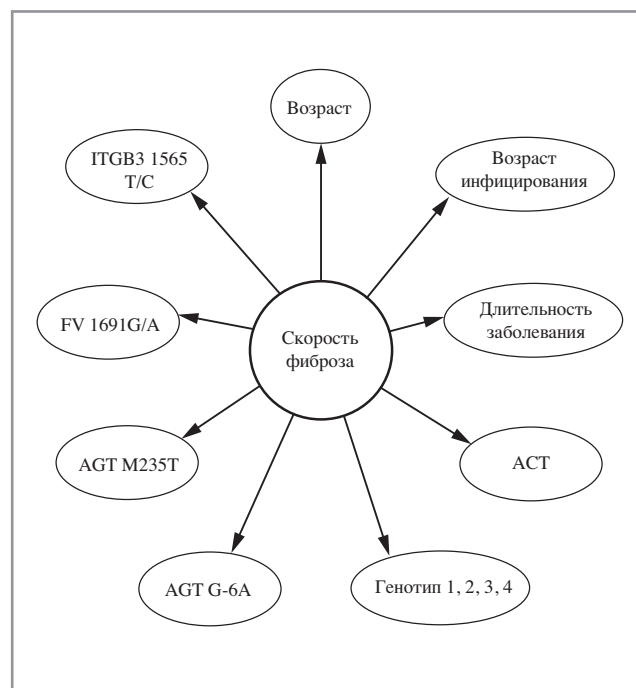


Рис. 2. БС 9+1 для КТ2.

Таблица 2. Различие в параметрах БС для исследуемых конечных точек (КТ1 – наличие ЦП, КТ2 – скорость фиброза, КТ3 – наличие портальной гипертензии, КТ4 – наличие криоглобулинов)

Включение параметров в начальные БС/ Количество узлов-листьев + корневой узел	КТ1/ 40+1	КТ2/ 42+1	КТ3/ 39+1	КТ4/ 45+1
Гепатоцеллюлярная карцинома	Вкл.	Вкл.		Вкл.
Класс по Чайлду				Вкл.
Васкулит	Вкл.	Вкл.	Вкл.	
Печеночно-клеточная недостаточность				Вкл.
Портальная гипертензия		Вкл.		Вкл.
Асцит брюшной полости				Вкл.
Варикозно расширенные вены пищевода				Вкл.
Анемия	Вкл.	Вкл.		Вкл.
Тромбоцитопения		Вкл.		Вкл.
Криоглобулины	Вкл.	Вкл.	Вкл.	
Фиброз			Вкл.	Вкл.

AUC, при варьировании числа узлов-листьев. Эта оптимизация проводилась с помощью разработанной нами программы SiLVIA (Simple Learn Variable Influence Analyzer) тем же способом, как это делалось в работах [10–13].

Более детальное, но все же краткое изложение этой методологии применения БС для прогностических моделей в медицине представлено в работе [10].

Подготовка данных и построение БС. На основании базы данных пациентов, представленной в виде Excel таблицы, для каждого исхода КТ строились начальные БС с наивной топологией (к одному исходу привязывались

Таблица 3. Качество предсказания AUC для начальных БС

База данных отбора (90%)	КТ1	КТ2	КТ3	КТ4
Начальные БС	40+1	42+1	39+1	45+1
Количество пациентов	228	166	81	201
Значение AUC	0,86	0,60	0,48	0,64

предполагаемые факторы риска). Корневым узлам БС соответствовали бинарные параметры КТ. В качестве узлов-листьев БС использовались все имеющиеся параметры базы данных, кроме параметров, значения которых напрямую зависят от КТ. В табл. 2 приведено различие во включенных параметрах в начальные сети для различных КТ, а также общее число узлов для всех начальных БС.

Поскольку БС работают с дискретными переменными, для всех непрерывных узлов осуществлена дискретизация – распределение значений соответствующих непрерывных параметров на несколько интервалов на основе экспертного мнения врачей-специалистов. Для проверки результатов предсказания на основе БС мы разбили полученную базу данных случайным образом на две части в соотношении 90% для базы отбора наиболее значимых прогностических параметров (основная группа) и 10% для базы контроля (контрольная группа). Дальнейшая работа и анализ осуществлялись с использованием построенных БС, а также двух баз данных пациентов – отбора и контроля.

Оценка надежности предсказания сетей осуществлялась программой ANN, использующей метод ROC-кривых. Эта оценка и выявление наиболее значимых прогностических параметров осуществлялись на базе данных отбора, содержащей 90% пациентов исследуемой базы. Для всех четырех КТ проведена валидация начальных БС: построены ROC-кривые и осуществлен расчет численного значения AUC (табл. 3).

Далее для улучшения качества предсказания всех КТ и выявления значимых прогностических параметров

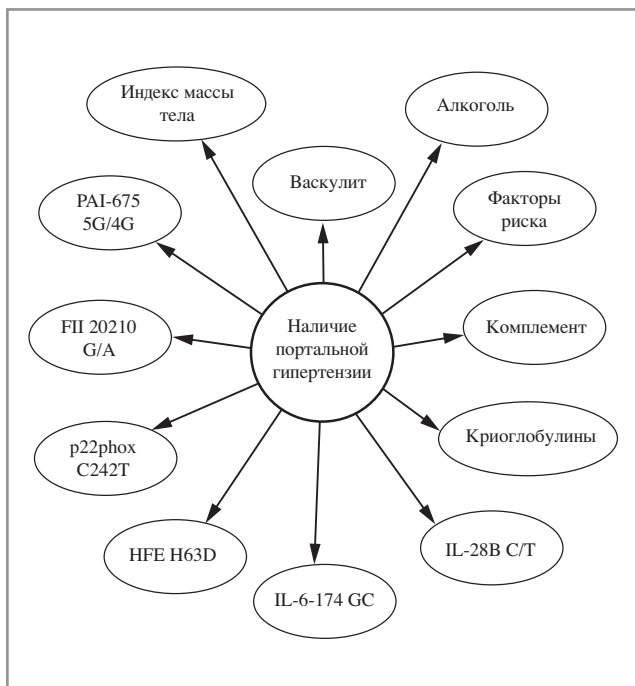


Рис. 3. БС 12+1 для КТ3.

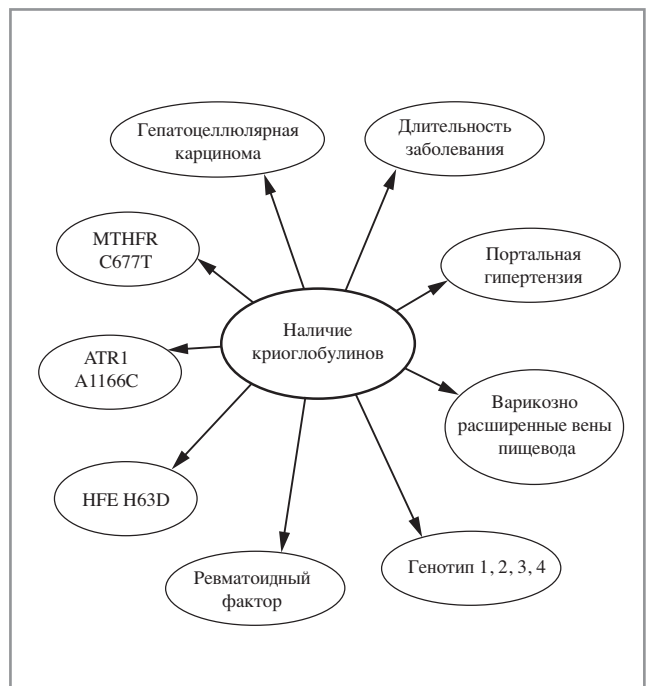


Рис. 4. БС 9+1 для КТ4.

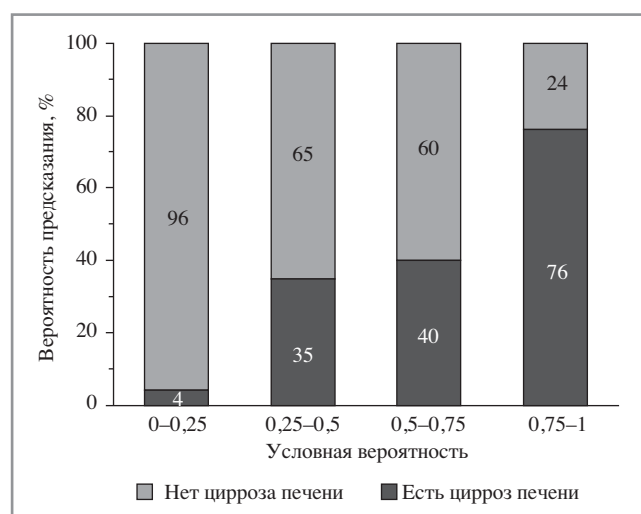


Рис. 5. Гистограмма риска предсказания ЦП (КТ1).

осуществлена оптимизация БС по числу узлов с использованием величины AUC в качестве целевой функции. При этом количество прогностических параметров для всех исследуемых КТ уменьшилось примерно в четыре раза.

Результаты и обсуждение

1. Оценка надежности предсказания БС. Самое лучшее качество предсказания получилось для КТ1 (предсказание ЦП), значение AUC равнялось 0,9, что является отличным показателем в экспертных системах [14]. Наибольшее увеличение качества предсказания достигнуто для КТ3 (предсказание портальной гипертензии), значение AUC увеличилось с 0,48 до 0,81. Качество предсказания оптимизированных БС для всех КТ приведено в табл. 4.

В оптимизированные БС вошли как клинические, так и генетические параметры пациентов (рис. 1-4).

При прогнозировании КТ1 – ЦП – значимыми оказались следующие параметры: возраст пациента, возраст инфицирования, длительность заболевания, факторы риска, индекс массы тела (ИМТ), концентрация аспаратамиотрансферазы (АСТ), наличие анемии, криоглобулинов, повышенного уровня альфа-фетопротеина (АФП), а также наличие мутации гена гемохроматоза (*HFE H63D, C282Y*), гена, кодирующего р22phox-субъединицу NADPH оксидазы (*C242T*), ангиотензина (*AGT G-6G*), гена тромбоцитарных рецепторов (*ITGB31565 T/C*).

При прогнозировании КТ2 – скорость прогрессирования фиброза – остались следующие параметры: возраст, возраст инфицирования, длительность заболевания, активность АСТ, наличие определенных генотипов вируса гепатита С. Выявлена связь между темпом прогрессирования фиброза печени и наличием полиморфизма генов ангиотензина (*AGT G-6G, AGT M235T*), фибриногена – Лейденской мутации (*FV I691G/A*), тромбоцитарных рецепторов (*ITGB31565 T/C*).

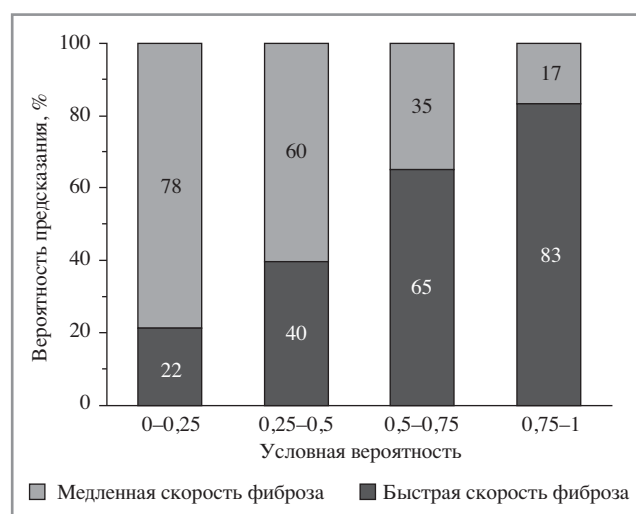


Рис. 6. Гистограмма риска предсказания скорости фиброза (КТ2).

При прогнозировании КТ3 – портальная гипертензия – в качестве узлов остались наличие васкулита и криоглобулинов, снижение гемолитической активности комплемента (т. е. наличие криоглобулинемического васкулита). Повышенный ИМТ и злоупотребление алкоголем способствуют более быстрому течению болезни и формированию портальной гипертензии. Среди генетических факторов, связанных с портальной гипертензией, оказались полиморфизм генов ингибитора активатора плазминогена (*PAI-1 -675 5G/4G*), фибриногена (*FII 20210 G/A*), гена, кодирующего р22phox-субъединицу NADPH-оксидазы (*C242T*), гена гемохроматоза *HFE H63D* и *IL-6 174GC*, отвечающего за повышенную секрецию провоспалительного цитокина IL-6 и гена *IL28B C/T*.

Наличие криоглобулинов и криоглобулинемический васкулит (КТ4), по полученным нами данным, ассоциированы с длительностью заболевания, портальной гипертензией, варикозным расширением вен пищевода (ВРВП), а также гепатоцеллюлярной карциномой, повышением титра ревматоидного фактора (РФ), мутациями в гене *MTHFR C677T*, генах *ATR A1166C* и *HFE H63D*.

2. Применение БС для предсказания целевых показателей. На основе проведенного анализа созданы гистограммы риска, позволяющие предсказывать вероятность наступления одного из вышеперечисленных исходов. Гистограммы риска связывают условную вероятность предсказания с априорной вероятностью предсказания для одной из групп риска пациентов. Для каждого исхода (КТ) рассчитаны четыре группы риска по значениям условных вероятностей (0-0,25; 0,25-0,5; 0,5-0,75; 0,75-1). В дальнейшем, используя информацию об имеющихся в базе данных пациентов с их известными исходами, определяли частоту наступления целевого показателя в каждой из этих групп риска (рис. 5-8).

Выводы о вероятности попадания пациента в ту или иную группу риска существенно зависят не только от са-

Таблица 4. Качество предсказания AUC после оптимизации БС

База данных отбора (90%)	ЦП КТ1	Скорость фиброза КТ2	Портальная гипертензия КТ3	Криоглобулинемия КТ4
Оптимизированные БС	14 + 1	9 + 1	12 + 1	9 + 1
Количество пациентов	228	166	81	201
Значение AUC	0,90	0,76	0,81	0,76

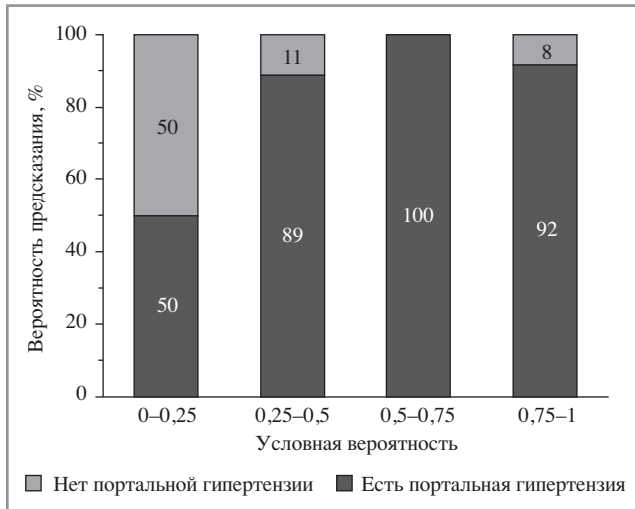


Рис. 7. Гистограмма риска предсказания портальной гипертензии.

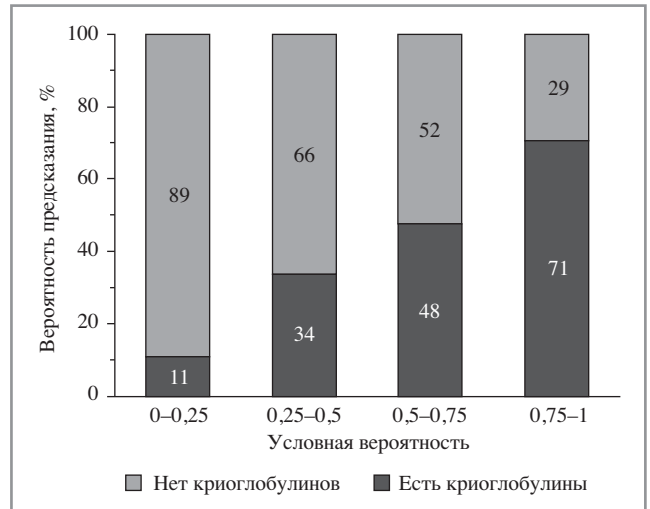


Рис. 8. Гистограмма риска предсказания наличия криоглобулинов.

мой группы риска, но и от рассматриваемой КТ: если для КТ1 в первой группе риска вероятность развития ЦП действительно мала – 4% (см. рис. 5), то вероятность развития портальной гипертензии для КТ3 (см. рис. 7) в первой группе риска весьма велика – 50%.

3. Проверка качества предсказания на контрольных группах. После оптимизации БС качество предсказания для всех КТ увеличилось. Для проверки полученных результатов мы обучили оптимизированные БС на базе данных отбора и далее осуществили предсказание для пациентов, входящих в базу данных контроля. Так как не для всех пациентов в исходной базе данных известны значения конечных точек (см. табл. 1), то контрольные группы для разных КТ содержали различное количество пациентов. Предсказательная способность для контрольной группы в нашем исследовании оказалась самой лучшей для развития ЦП (КТ1), значение AUC на 25 пациентов контрольной группы равнялось 0,8 (табл. 5). Подробное распределение пациентов контрольной группы КТ1 на группы риска продемонстрировано на рис. 9. Полученный результат подтверждает возможность применения экспертной системы предсказания ЦП с использованием оптимизированных БС. Результаты проверки контрольных групп для остальных КТ оказались не убедительны. Возможно, плохое предсказание связано с малым количеством пациентов в основной и контрольной группах.

Заключение

В проведенном исследовании представлен метод прогнозирования развития ЦП, скорости прогрессирования фиброза, портальной гипертензии и наличия криоглобулинов и криоглобулинемического васкулита у больных хроническим гепатитом С с использованием БС. Полученные результаты позволили нам уточнить вклад не только тра-

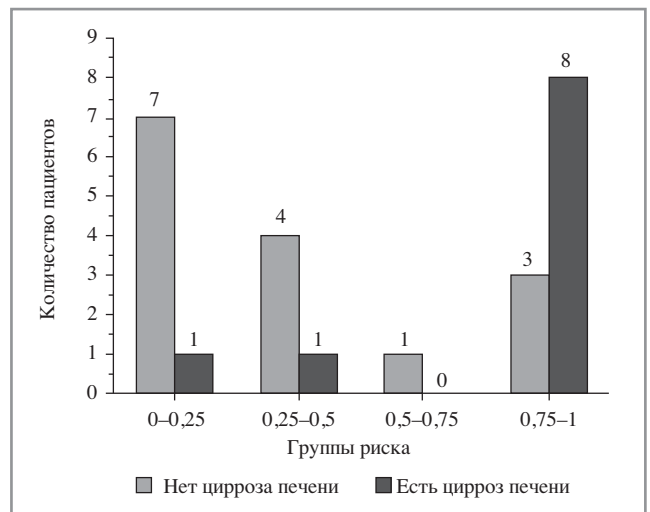


Рис. 9. Предсказание для пациентов контрольной группы КТ1.

диционно исследуемых клинических факторов вируса и хозяина на развитие четырех исходов, определяющих статус больного, но и прогнозировать течение болезни в зависимости от индивидуальных реакций, зависящих от точечных мутаций генов, контролирующих основные патогенетические пути формирования и прогрессирования ХГС.

Кроме того, проведенное исследование продемонстрировало возможные направления дальнейшего поиска механизмов прогрессирования ХГС, поскольку не только подтверждена связь изучаемых исходов с факторами, которые хорошо известны из данных литературы, но и выявлена связь с молекулярными механизмами, которые могут быть уточнены при расширении базы данных.

Таблица 5. Проверка качества предсказания для контрольной группы

База данных отбора (10%)	ЦП КТ1	Прогрессирование КТ2	Портальная гипертензия КТ3	Криоглобулинемический васкулит КТ4
Оптимизированные БС	14 + 1	9 + 1	12 + 1	9 + 1
Количество пациентов	25	21	8	19
Значение AUC	0,80	0,57	0,21	0,44

Связь данных клинико-демографических параметров и наличия ЦП у пациента понятна, однако выявленные генетические связи требуют дальнейшего тщательного изучения. Так, ассоциация между полиморфизмом в гене *HFE* и ЦП, по-видимому, связана с перегрузкой железом, что способствует прогрессированию болезни печени. Того же мнения придерживаются В.У. Tung и соавт., по данным которых полиморфизм в гене *HFE* ассоциирован с перегрузкой железом у больных на стадии гепатита и компенсированного ЦП [15–17].

Наше исследование впервые продемонстрировало роль полиморфизма гена, кодирующего р22phox-субъединицу NADPH оксидазы (*C242T*), на течение и прогноз хронических заболеваний печени. Известно, что наличие этого полиморфизма ассоциируется с артериальной гипертензией [18] и позволяет предположить, что данная мутация приводит к эндотелиальной дисфункции, тем самым способствуя прогрессированию фиброза через портальную гипертензию.

Полиморфизм гена *AGT* обуславливает не только активацию ангиотензина в плазме, но и активацию локально тканевой его формы, что является стимулом к формированию фиброза органов [19].

Нами впервые продемонстрировано влияние полиморфизма гена тромбоцитарных рецепторов *ITGB31565 T/C* на прогрессирование фиброза печени. Связь между полиморфизмом тромбоцитарного рецептора *ITGB31565 T/C* и наличием цирроза печени, возможно, объясняется дисфункцией тромбоцитов при наличии мутации, что приводит к активации сосудистого компонента (в том числе и развитие лейкокластического васкулита).

Известно, что носительство мутации в гене *FV 1691G/A* связано с формированием микротромбозов в ткани печени [20–22], несомненно способствующих прогрессированию фиброза.

КТ2 – скорость прогрессирования фиброза – зависела от уже известных параметров, таких как возраст, возраст инфицирования, длительность заболевания, концентрация АСТ, наличие определенных генотипов вируса гепатита С [5, 23]. Представляет интерес выявленная нами связь между темпом прогрессирования фиброза печени при ХГС и наличием полиморфизма генов ангиотензина (*AGT G-6G*, *AGT M235T*), фибриногена, Лейденской мутации (*FV 1691G/A*), тромбоцитарных рецепторов (*ITGB31565 T/C*).

Таким образом, одни и те же полиморфизмы оказались тесно связаны с темпом развития ЦП и с его выраженностью, что, очевидно, позволяет относить больных с этими особенностями к наиболее прогностически неблагоприятной группе, требующей лечения.

Ремоделирование портальной системы кровотока с развитием портальной гипертензии (КТ3) оказалось связанным с наличием повышенного ИМТ и злоупотреблением алкоголем, что в качестве неблагоприятных ко-факторов

развития портальной гипертензии хорошо известно. А взаимосвязь портальной гипертензии с наличием криоглобулинов и криоглобулинемического васкулита, вероятно, связано не только с тем, что наиболее часто развитие криоглобулинемического васкулита ассоциировано с более поздней стадией болезни [24, 25]. Наше исследование выявило ряд генетических факторов, связанных с портальной гипертензией. Ими оказались гены, участвующие через активацию эндотелия и формирование микроангиопатий в повреждении сосудистой стенки, такие как полиморфизм генов ингибитора активатора пламиногена (*PAI-1 -675 5G/4G*), фибриногена (*FII 20210 G/A*), гена, кодирующего р22phox-субъединицу NADPH оксидазы (*p22phox C242T*).

Взаимосвязь с портальной гипертензией установлена с полиморфизмом гена гемохроматоза *HFE H63D*, связанного с синдромом перегрузки железом. Перегрузка железом может приводить к агрессивным локальным окислительным процессам. Ген *IL-6 174GC*, отвечающий за повышенную секрецию провоспалительного цитокина IL-6, вероятно, способствует более быстрым фибротическим процессам в ткани печени через активную воспалительную реакцию, тем самым ускоряя развитие портальной гипертензии. Влияние мутации в гене *IL28B C/T* на наличие портальной гипертензии трудно объяснимо, однако можно предположить, что данная мутация приводит к ослаблению естественной противовирусной защиты, а следовательно, способствует персистенции вируса. Ранее показана связь гена *IL28B C/T* с ответом на противовирусную терапию.

Наличие криоглобулинов (КТ4), по полученным нами данным, ассоциировано с длительностью заболевания, портальной гипертензией, ВРВП, а также гепатоцеллюлярной карциномой, что легко объяснимо, так как все перечисленные факторы определяются стажем болезни и длительной персистенцией в организме вируса гепатита С, приводящих к изменению иммунного ответа. Более высокий уровень гомоцистеина при наличии мутации в гене *MTHFR C677T* может способствовать формированию криоглобулинемического васкулита, что также показано и в работе М. Casato и соавт. [26]. Нами не найдено данных, объясняющих влияние полиморфизма генов *ATR A1166C* и *HFE H63D* на образование криоглобулинов, но возможно, что они играют роль в повреждении сосудистой стенки через оксидативный стресс (*HFE H63D*) и индукцию (*ATR A1166C*) лейкокластического васкулита, где повреждающая роль тромбоцитов хорошо известна.

Полученные в данной работе оптимизированные облученные БС могут быть использованы для уточнения прогноза у каждого нового пациента на основе его персональных параметров при расчете гистограммы риска.

Исследование выполнено в рамках проекта № 14-50-00029 Российского научного фонда. Конфликт интересов отсутствует.

ЛИТЕРАТУРА/REFERENCES

1. Самоходская Л.М., Старостина Е.Е., Яровая Е.Б., Краснова Т.Н., Мухин Н.А., Ткачук В.А., Садовничий В.А. Математическая модель прогноза скорости фиброза печени у больных с хроническим гепатитом С на основе комбинаций геномных маркеров. *Вестник Российской академии медицинских наук*. 2015;70(5):651-61 [Samokhodskaya LM, Starostina EE, Yarovaya EB, Krasnova TM, Mukhin NA, Tkachuk VA, Sadovnichiy VA. Mathematic model of prognosis of the liver fibrosis progression rate in patients with chronic hepatitis C based on the combination of genomic markers. *Vestnik Rossiyskoy Akademii Medicinskikh Nauk*. 2015;70(5):651-61 (In Russ.)].
2. Сулимов А.В., Втюрина Д.Н., Романов А.Н., Масленников Е.Д., Сулимов В.Б., Курочкин И.Н., Упоров И.В., Затеищikov Д.А., Носиков В.В., Варфоломеев С.Д. Экспертные системы персонализированной медицины: применение байесовских сетей для предсказания состояния пациентов. В кн.: Пост-геномные исследования и технологии (под ред. чл.-корр. РАН С.Д. Варфоломеева). Москва: Изд-во МГУ; 2011. С. 641-702 [Sulimov AV, Vtyurina DN, Romanov AN, Maslennikov ED, Sulimov VB, Kurochkin IN, Uporov IV, Zateystchikov DA, Nosikov VV, Varfolomeev SD. Expert systems of personalized medicine: using Bayesian networks for prediction of patients condition. In: *Post-genomnye issledovaniya i tekhnologii* [Post-

- genomic investigations and technologies] (edited by a corresponding member of the Russian Academy of Sciences SD Varfolomeev). Moscow: MSU publishing house; 2011. P. 641-702 (In Russ.).
3. Генс Г.П., Сулимов А.В., Моисеева Н.И., Каткова Е.В., Вельшер Л.З., Коробкова Л.И., Савкин И.А., Сулимов В.Б. Применение генных сигнатур и медицинских экспертных систем для прогнозирования клинических исходов рака молочной железы. *Вестник РОНЦ им. Н.Н. Блохина РАМН*. 2015;26(4):16-33 [Gens GP, Sulimov AV, Moiseeva NI, Katkova EV, Velsher LZ, Korobkova LI, Savkin IA, Sulimov VB. Application of gene signatures and medical expert systems for prediction of the clinical outcomes of breast cancer. *Vestnic RONC im. N.N. Blochina RAMN*. 2015;26(4):16-33 (In Russ.).]
 4. Lucas PJ, van der Gaag LC, Abu-Hanna A. Bayesian network sinbio-medicine and health-care. *Artif Intell Med*. 2004;30(3):201-14. doi: 10.1016/j.artmed.2003.11.001
 5. Gevaert O, De Smet F, Timmerman D, Moreau Y, De Moor B. Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks. *Bioinformatics*. 2006;22(14):184-90. doi: 10.1093/bioinformatics/btl230
 6. Poynard T, Ratziu V, Charlotte F, Goodman Z, Mchutchison J, Albrecht J. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C. *J Hepatol*. 2001;34(5):730-9.
 7. Jensen FV, Nielsen TD. Bayesian Networks and Decision Graphs. New York: Springer Verlag; 2007: 193 p.
 8. Obuchowski NA. ROC analysis. *Am J Roentgenol*. 2005;184(2):364-72. doi: 10.2214/ajr.184.2.01840364
 9. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36. doi: 10.1148/radiology.143.1.7063747
 10. Maslennikov ED, Sulimov AV, Savkin IA, Evdokimova MA, Zateyshchikov DA, Nosikov VV, Sulimov VB. An intuitive risk factors search algorithm: usage of the Bayesian network technique in personalized medicine. *J Applied Statistics*. 2015;42(1):71-87. doi: 10.1080/02664763.2014.934664
 11. Генс Г.П., Сулимов А.В., Моисеева Н.И., Овсий О.Г., Вельшер Л.З., Рыбалкина Е.Ю., Селезнева И.И., Савкин И.А., Сулимов В.Б. Поиск подходов к прогнозированию исходов рака молочной железы с помощью байесовских сетей. *Онкология*. 2014;3(5):37-46 [Gens GP, Sulimov AV, Moiseeva NI, Ovsy O, Velsher LZ, Rybalkina EYu, Seleznyova II, Savkin IA, Sulimov VB. Search for approaches for prediction of breast cancer using Bayesian networks. *Onkologiya*. 2014;3(5):37-46 (In Russ.).]
 12. Топтыгина А.П., Азиатцева В.В., Савкин И.А., Кислицин А.А., Семикина Е.Л., Гребенников Д.С., Алешкин А.В., Сулимов А.В., Сулимов В.Б., Бочаров Г.А. Прогнозирование специфического гуморального иммунного ответа на основании исходных параметров иммунного статуса детей, привитых против кори, краснухи и эпидемического паротита. *Иммунология*. 2015;36(1):22-30 [Topytgina AP, Aziatceva VV, Savkin IA, Kislitsyn AA, Semikina EL, Grebennikov DS, Aleshkin AV, Sulimov AV, Sulimov VB. Prediction of a specific humoral immune response based on the initial parameters of the immune status of children vaccinated against measles, rubella and mumps. *Immunologiya*. 2015;36(1):22-30 (In Russ.).]
 13. Sulimov AV, Meshkov AN, Savkin IA, Katkova EV, Kutov DC, Hasanova ZB, Konovalova NV, Kukharchuk VV, Sulimov VB. Genome-wide analysis of genetic associations for prediction of polygenic hypercholesterolemia with bayesian networks. *J Comp Eng Math*. 2015;2(4):11-26. doi: 10.14529/jcem150402
 14. Vanagas G. Receiver operating characteristic curves and comparison of cardiac surgery risk stratification systems. *Interact Cardiovasc Thorac Surg*. 2004;3(2):319-22. doi: 10.1016/j.icvts.2004.01.008
 15. Tung BY, Emond MJ, Bronner MP, Raaka SD, Cotler SJ, Kowdley KV. Hepatitis C, iron status, and disease severity: relationship with HFE mutations. *Gastroenterology*. 2003;124(2):318-26. doi: 10.1053/gast.2003.50046
 16. Barbaro G, Di Lorenzo G, Ribersani M, Soldini M, Giancaspro G, Bellomo G, Belloni G, Grisorio B, Barbarini G. Serum ferritin and hepatic glutathione concentrations in chronic hepatitis C patients related to the hepatitis C virus genotype. *J Hepatol*. 1999;30(5):774-82.
 17. Erhardt A, Maschner-Olberg A, Mellenthin C, Kappert G, Adams O, Donner A, Willers R, Niederau C, Haussinger D. HFE mutations and chronic hepatitis C: H63D and C282Y heterozygosity are independent risk factors for liver fibrosis and cirrhosis. *J Hepatol*. 2003;38(3):335-42.
 18. Moreno MU, José GS, Fortuño A, Belouqui O, Díez J, Zalba G. The C242T CYBA polymorphism of NADPH oxidase is associated with essential hypertension. *J Hypertens*. 2006;24(7):1299-306. doi: 10.1097/01.hjh.0000234110.54110.56
 19. Altarescu G, Haim S, Elstein D. Angiotensinogen promoter and angiotensinogen II receptor type 1 gene polymorphisms and incidence of ischemic stroke and neurologic phenotype in Fabry disease. *Biomarkers*. 2013;18(7):595-600. doi: 10.3109/1354750X.2013.836244
 20. Anstee QM, Dhar A, Thursz MR. The role of hypercoagulability in liver fibrogenesis. *Clin Res Hepatol Gastroenterol*. 2011;35(8-9):526-33. doi: 10.1016/j.cline.2011.03.011
 21. Plompen EP, Darwish Murad S, Hansen BE, Loth DW, Schouten JN, Taimr P, Hofman A, Uitterlinden AG, Stricker BH, Janssen HL, Leebek FW. Prothrombotic Genetic Risk Factors are associated with an Increased Risk of Liver Fibrosis in the General Population: The Rotterdam Study. *J Hepatol*. 2015;63(6):1459-65. doi: 10.1016/j.jhep.2015.07.026
 22. Wright M, Goldin R, Hellier S, Knapp S, Frodsham A, Hennig B, Hill A, Apple R, Cheng S, Thomas H, Thursz M. Factor V Leiden polymorphism and the rate of fibrosis development in chronic hepatitis C virus infection. *Gut*. 2003;52(8):1206-10.
 23. Bochud PY, Cai T, Overbeck K, Bochud M, Dufour JF, Mollhaupt B, Borovicka J, Heim M, Moradpour D, Cerny A, Malinverni R, Francioli P, Negro F. Genotype 3 is associated with accelerated fibrosis progression in chronic hepatitis C. *J Hepatol*. 2009;51(4):655-66. doi: 10.1016/j.jhep.2009.05.016
 24. Dammacco F, Sansonno D. Therapy for hepatitis C virus-related cryoglobulinemic vasculitis. *N Engl J Med*. 2013;369(11):1035-45. doi: 10.1056/NEJMra1208642
 25. Saadoun D, Asselah T, Resche-Rigon M, Charlotte F, Bedossa P, Valla D, Piette JC, Marcellin P, Cacoub P. Cryoglobulinemia is associated with steatosis and fibrosis in chronic hepatitis C. *Hepatology*. 2006;43:1337-45. doi: 10.1002/hep.21190
 26. Casato M, Carlesimo M, Francia A, Timarco C, Antenucci A, Bove M, Martini H, Visentini M, Fiorilli M, Conti L. Influence of inherited and acquired thrombophilic defects on the clinical manifestations of mixed cryoglobulinaemia. *Rheumatology (Oxford)*. 2008;47:1659-63. doi: 10.1093/rheumatology/ken303

Поступила 16.09.2018