

**Рус.:**УДК: 004.048

*Оптимизация информационных данных, анализируемых в процессе нейросетевого прогнозирования исхода патологий гемостаза*

Сидорова Маргарита Александровна, Сержантова Наталья Александровна

*Аннотация:*

Нейросетевые диагностические и прогностические системы в последние пять лет активно внедряются в клиническую практику. Подобные системы обладают большей чувствительностью и специфичностью по сравнению с традиционными шкалами оценки тяжести состояний, однако в описании большинства таких систем отсутствует информация о принципах и алгоритмах выбора параметров, подаваемых на вход нейронной сети. Зачастую это приводит к избыточности исследуемых параметров, что, в свою очередь, негативно влияет на результат работы нейросетевой системы. Предложен алгоритм выбора информативных параметров для оптимизации информационных данных, анализируемых в процессе нейросетевого прогнозирования исхода патологий гемостаза.

*Ключевые слова:* гемостаз, прогноз, исход, нейронная сеть, алгоритм, информативный параметр.

**Eng.:** *Optimization of information data analyzed in the process of neural network prediction outcome pathology of hemostasis*

Sidorova Margarita Alexandrovna, Serzhantova Natalia Alexandrovna

*Abstract:*

Neural network systems for diagnostics and prognosis are actively introduced in clinical practices in last five years. These systems have higher detection sensitivity and specificity compared to traditional scales for valuation status severity. However the description most of these systems are not include information of principles and algorithms for selection neural network input parameters. Often this leads to redundancy of researched parameters, which in turn have negative impact upon work result of neural network systems. In this article proposed to the algorithm of informative parameters selection for information data optimization, analyzed in process neural network prognosis of outcome pathology hemostasis.

*Keywords:* hemostasis, prediction, the outcome, a neural network, algorithm, informative parameter.

## **Введение**

Система гемостаза – это биологическая система, обеспечивающая, с одной стороны, сохранение жидкого состояния циркулирующей крови, а с другой – предупреждение и прекращение кровотечений.

Расстройства гемостаза – один из самых частых и потенциально опасных патологических состояний, встречающихся в практической медицине.

Нарушения гемостаза могут проявляться и как самостоятельные синдромы и как вторичные геморрагические и тромботические осложнения других заболеваний: инфекционно-септических, сердечно-сосудистых, иммунных, акушерской патологии, болезней крови, печени, почек и т.д.[2].

Сидорова М. А., Сержантова Н. А., Оптимизация информационных данных, анализируемых в процессе нейросетевого прогнозирования исхода патологий гемостаза // «Живые и биокосные системы». – 2015. – № 14; URL: <http://www.jbks.ru/archive/issue-14/article-10>

В последние годы интерес к лабораторной диагностике системы гемостаза стремительно растет. В точной и оперативной информации о состоянии гемокоагуляции нуждаются врачи: кардиологи, хирурги, реаниматологи, акушеры-гинекологи, нейрофизиологи и многие другие специалисты. Параметры системы гемостаза являются важными для диагностики многих заболеваний [5].

Особая значимость контроля свертывающей системы крови в последнее время связана с активным применением в клинической практике современных высокоэффективных антикоагулянтов прямого и непрямого действия, антиагрегантов и тромболитиков. Узкий терапевтический коридор и избирательная индивидуальная чувствительность к препаратам разных групп требует точного индивидуального подбора дозировки и продолжительности лечения, а главное целесообразности и эффективности проводимой терапии. Более того, большое количество людей имеют определенные факторы риска, обусловленные генетическими дефектами отдельных звеньев гемостаза, применением некоторых фармакологических препаратов, в том числе антибиотиков, наличием сопутствующих заболеваний. В обычных условиях дисбаланс компонентов свертывающей системы может не проявиться, но в условиях оперативного вмешательства или бесконтрольно назначенной терапии риск тяжелых осложнений у таких пациентов многократно возрастает. В тоже время, выявление этих факторов риска у разных групп пациентов и принятие соответствующих профилактических мероприятий позволяет избежать тяжелых осложнений. Поэтому своевременная и точная диагностика и качественное прогнозирование исхода патологий гемостаза на основе анализа параметров свертывания крови имеет принципиальное значение для клинической практики.

Для объективной оценки тяжести состояния пациента при диагностике и прогнозировании исхода патологий гемостаза разработано несколько методов, которые объединяет общий принцип — подсчет количественного критерия (суммы баллов) по той или иной шкале [1]. К этому методу можно отнести интегральные шкалы оценки тяжести состояния (APACHE, APACHE II, APACHE III, SAPS, SAPS II, SOFA, MODS и ряд других).

Несмотря на более чем двадцатилетнюю историю существования интегральных систем диагностики, оценки тяжести состояния и прогнозирования исхода с постоянно увеличивающимся количеством этих систем, диагностика и прогнозирование исхода патологий гемостаза в отделениях реанимации и интенсивной терапии остается несовершенной,

Все используемые шкалы имеют общие недостатки: низкая точность диагностики и прогнозирования исхода для отдельного пациента. Ни одна из существующих шкал не позволяет получить прогноз исхода с высокой чувствительностью и высокой специфичностью одновременно.

Решение проблемы видится в применении интеллектуальных информационных систем на базе искусственных нейронных сетей для

диагностики и прогнозирования исхода различных заболеваний. Подобные системы в последние 5 лет активно внедряются в клиническую практику и позволяют решать узкие задачи диагностики и прогнозирования, обеспечивая чувствительность и специфичность диагностики и прогнозирования на уровне 80- 85 %. Однако проведенные исследования показали, что при проектировании таких систем недостаточно внимания уделяется выбору информативных параметров, что влияет на качество результатов диагностики и прогнозирования [3].

### **Цель исследования**

Целью исследования является разработка алгоритма, обеспечивающего оптимизацию информационных данных, подаваемых на вход нейросетевой системы для диагностики и прогнозирования исхода патологий гемостаза.

### **Материал и методы исследования**

Современная медицина оперирует множеством различных биохимических показателей, характеризующих коагуляционный гемостаз. Ниже перечислены основные параметры гемостаза (вместе с общепринятыми сокращениями), наиболее часто используемые при диагностике нарушений свертывания крови в силу своей информативности и невысокой стоимости реагентов для лабораторной диагностики.

1. Вс – время свертывания (мин.)
2. Эх – эхитоксовое время (сек.)
3. АтЗ – антитромбин III (мг/л)
4. АПТВ - активированное парциальное (частичное) тромбопластиновое время (сек.)
5. АПТВк - активированное парциальное (частичное) тромбопластиновое время (сек.) в контрольном образце
6. Тв – тромбиновое время (сек.)
7. Оф – ортофенантролин (мг/дл)
8. Кл – клампинг-тест (тест склеивания стафилококков) (г/л)
9. ААТ2 – показатель агрегационной активности одного тромбоцита при индукции агрегации разведением гемолизата  $10^{-2}$  (сек.)
10. ААТ6 – показатель агрегационной активности одного тромбоцита при индукции агрегации разведением гемолизата  $10^{-6}$  (сек.)
11. ИАТ – индекс активации тромбоцитов (ААТ6/ААТ2)
12. Хзф – хагеман-зависимый фибринолиз (мин.)
13. Иэф – индуцированный стрептокиназой эуглобулиновый фибринолиз (сек.)
14. Пти – протромбиновый индекс (%)
15. Фг – фибриноген (г/л)
16. Тромб – количество тромбоцитов (тыс./мкл).

Для исследования диагностических и прогностических способностей нейросетевой системы были использованы значения приведенных выше 16 параметров гемостаза 150 пациентов (75 здоровы, 75 – больны перитонитом с различными исходами). Проведенные исследования показали, что для

процесса предварительной диагностики такое количество входных параметров является избыточным, а некоторые из параметров свертывания не только не увеличивают вероятность правильной классификации состояния пациента, но и во многих случаях приводят, к снижению качества постановки диагноза [4]. В связи с этим, выполнялся процесс исключения неинформативных параметров из симптомокомплексов предполагаемой патологии с целью сокращения трудоемкости исследования.

При сокращении количества параметров в первую очередь рассчитывается коэффициент корреляции для определения наличия взаимосвязи между параметрами. Анализ коэффициентов корреляции показал, что большинство параметров у здоровых пациентов слабо коррелированы, а исключение параметров на основе исследований выборки патологических значений нецелесообразно, так как может привести к ошибкам распознавания сетью нормальных значений. В связи с этим потребовалось оценить «вклад» каждого параметра в работу нейронной сети. Матрица весовых коэффициентов для такой оценки не подходит, так как определить степень влияния остальных параметров на исследуемый, не представляется возможным.

Логичным решением для выявления наиболее информативных параметров является применение статистических методов исследования выборки (как стандартных, так и специализированных). При этом исследование параметров гемостаза проводится согласно алгоритму, представленному на рисунке 1.



*Рисунок 1 - Схема алгоритма выбора информативных параметров с использованием методов статистического анализа*

Предварительная обработка (блок 03) данных подразумевает применение метода описательной статистики, с помощью которого производится расчет робастных характеристик (моды, медианы, перцентиля, асимметрии, эксцесса), далее следует определения необходимого объема выборки, согласно правилам планирования эксперимента. Подробный анализ (блоки 04-08) предусматривает применение различных критериев (D-критерий Дарбина–Уотсона, RS-критерий, поворотных точек, диагностический коэффициент, коэффициент по Шеннону, коэффициент по Кульбаку) с

помощью которых определяются адекватные, точные и наиболее информативные параметры.

При этом использовался алгоритм «отсева» малоинформативных симптомов (рисунок 2), заключающийся в том, что производится вычислительная диагностика с использованием всех симптомов; определяется эффективность диагностики по формуле  $\eta = P_{np} + Ш_1 P_1 + Ш_2 P_2$ , где  $\eta$  - эффективность диагностики,  $P_{np}$  - процент правильных диагнозов,  $Ш_1$  - штраф за ошибку первого рода,  $P_1$  - процент ошибок первого рода,  $Ш_2$  - штраф за ошибку второго рода,  $P_2$  - процент ошибок второго рода; затем исключается один из параметров и по тем же исходным данным вновь ставится диагноз и определяется эффективность  $\eta$ . Если новое значение эффективности больше предыдущего, то исключается другой параметр и опять определяется эффективность. Так продолжается до тех пор, пока не сформируется набор параметров, при котором исключение из набора любого параметра или добавление параметра в набор приведет к уменьшению значения  $\eta$ .

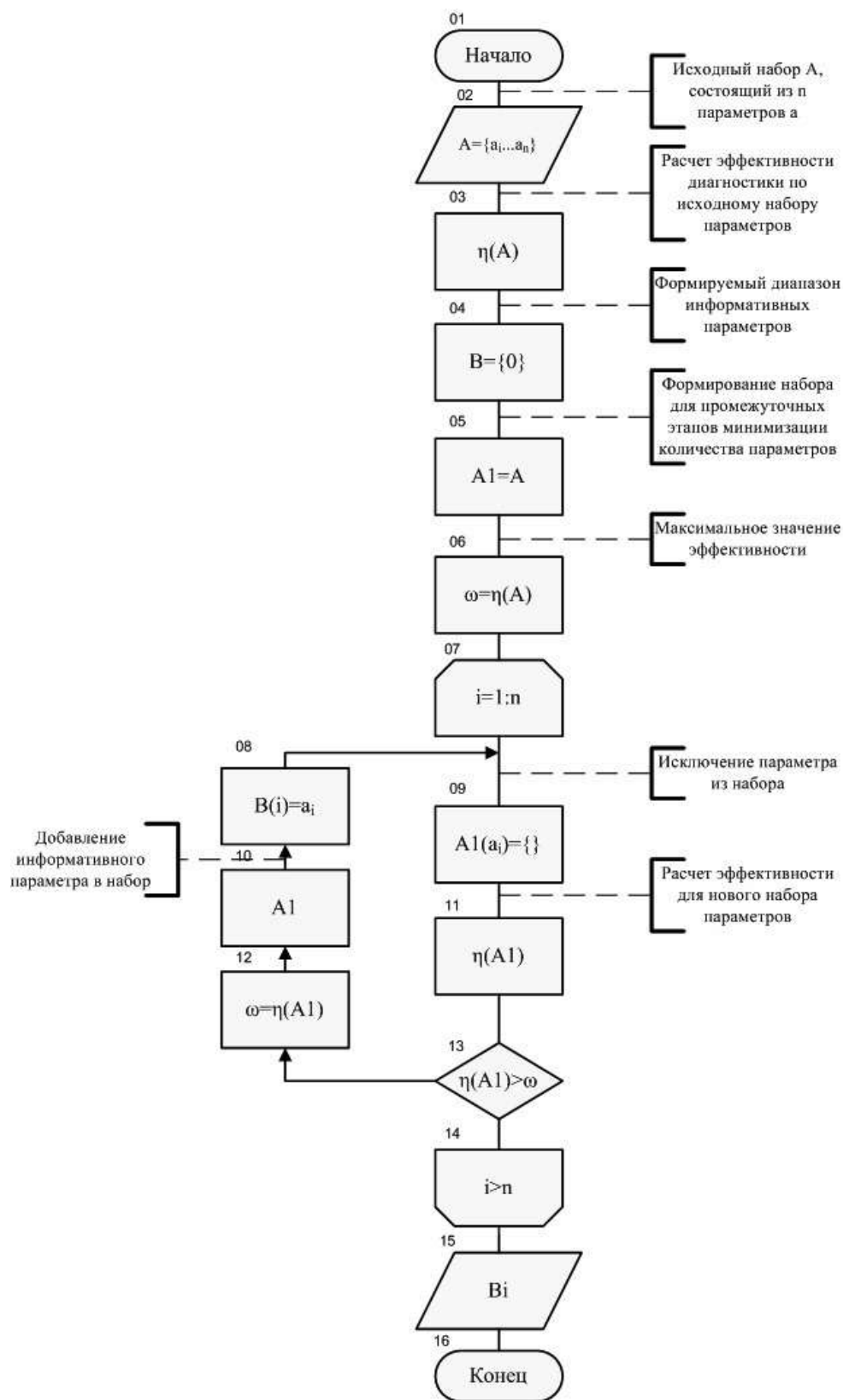


Рисунок 2 - Схема алгоритма «отсева» малоинформативных симптомов

Тестирование разработанных алгоритмов проводилось на реальных клинических данных.

Для оценки значимости каждого из рассматриваемых параметров рассчитали  $\eta$  - эффективность диагностики,  $P_{np}$  – процент правильных

диагнозов как отношение количества верных срабатываний нейронной сети к объему выборки, выраженное в процентах,  $P_1$  – процент ошибок первого рода (ложного срабатывания нейронной сети),  $P_2$  – процент ошибок второго рода (ложных пропусков); при исключении того или иного параметра (таблица 1). Штраф за ошибку первого рода  $Ш_1$  и штраф за ошибку второго рода  $Ш_2$  принимаем равными 0. В этом случае ложноположительные и ложноотрицательные результаты исключаются из массива данных, используемых для расчета эффективности диагностики. В частных случаях решения задачи оценки эффективности нейросетевой диагностики значения штрафов можно выбрать из диапазона  $[-1;1]$ . Использование отрицательных значений штрафов  $Ш_1$  и  $Ш_2$  будет свидетельствовать о более высоких требованиях, предъявляемых к нейронным сетям при оценке качества их работы. Положительные значения штрафов рекомендуется использовать с осторожностью (например, в случаях, если в проектируемых системах необходимо обеспечить только высокую чувствительность или высокую специфичность результата), так как в этом случае значение эффективности диагностики  $\eta$  становится менее информативным.

*Таблица 1 – Значение вероятности правильной классификации при исключении параметра из анализируемого множества*

Исключаемый параметр	$P_{np}$ – процент правильных диагнозов при исключении параметра	$P_1$ – процент ошибок первого рода	$P_2$ – процент ошибок второго рода
Нет	85	0	15
вс	85	4	11
эх	81	0	19
атЗ	88	0	12
апТВ	92	0	8
апТВК	85	0	15
тв	73	0	27
оф	77	0	23
кл	81	0	19
ААТ2	88	0	12
ААТ6	85	0	15
ИАТ	88	0	12
хзф	77	0	23
изф	73	0	27
пти	88	0	12
фг	88	0	12
тромб	85	0	15

В таблице 1 цветом выделены параметры, исключение которых приводит к увеличению процента правильной классификации состояния пациента. Для оптимизации информационных данных исключили эти параметры из обучающих и тестовых множеств. На основании проведенных исследований в качестве входных параметров выделили следующие: вс, эх, апТВК, тв, оф, кл, ААТ6, хзф, изф, тромб. Таким образом, получили нейронную сеть, на выход которой подаются 10 параметров свертывания крови. Для данной



нейронной сети была рассчитана эффективность диагностики состояния пациента, которая составила 92%.

### **Результаты исследования и их обсуждение**

В результате проведения исследований были разработаны алгоритмы выбора информативных параметров и «отсева» малоинформативных симптомов. Данные алгоритмы прошли апробацию на реальных клинических данных. Использование при нейросетевой диагностике всего массива входных параметров позволяет получить верный диагноз или прогноз исхода заболевания с эффективностью 85%. После исключения малоинформативных признаков (АтЗ, АПТВ, ААТ2, ИАТ, ПТИ, Фг) из обучающих и тестовых множеств, эффективность диагноза или прогноза для тех же пациентов возросла до 92%. Таким образом, оптимизация информационных данных с помощью разработанных алгоритмов позволила повысить значение вероятности постановки верного диагноза (или получения качественного прогноза) на 7%.

### **Заключение**

Результаты исследований являются важными для обеспечения своевременной и качественной диагностики нарушений в системе гемостаза на ранних стадиях развития заболеваний. Исследования являются основой создания нейросетевой системы диагностики и прогнозирования исхода патологий гемостаза, что, в конечном итоге, опосредовано определяет перспективы сохранения здоровья населения.

### **Список литературы**

1. Кишкун, А.А. Руководство по лабораторным методам диагностики. – М.: Издательская группа «ГЕОТАР-Медиа», 2007.
2. Момот А.П. Патология гемостаза принципы и алгоритмы клинко-лабораторной диагностики. СПб. - 2006.
3. Осовский, С. Нейронные сети для обработки информации / Пер. с польского И. Д. Рудинского. – М.: Финансы и статистика, 2002.
4. Сидорова М.А., Ерушова Н.А. Обоснование выбора параметров нейросетевой скрининговой диагностики нарушений системы гемостаза // Мехатроника, автоматизация, управление: Теоретический и прикладной научно-технический журнал. - 2009. - №8. – М., 2009. - С. 59 – 66.
5. Сидорова М.А., Сержантова Н.А., Филиппова Л.А. Диагностика и прогнозирование исходов перитонита с помощью нейросетевой системы исследования параметров гемостаза // Медицинская техника: Научно-технический журнал. - 2011. - № 2. - М., 2011. - С. 42 – 74.

### **Spisok literatury**

1. Kishkun, A.A. Rukovodstvo po laboratornym metodam diagnostiki. – М.: Izdatel'skaya gruppy «GEOTAR-Media», 2007.
2. Momot A.P. Patologiya gemostaza principy i algoritmy kliniko-laboratornoj diagnostiki. SPb. - 2006.

3. Osovskij, S. Neironnyye seti dlya obrabotki informacii / Per. s pol'skogo I. D. Rudinskogo. – M.: Finansy i statistika, 2002.
4. Sidorova M.A., Erushova N.A. Obosnovanie vybora parametrov nejrosetevoj skringovoj diagnostiki narushenij sistemy gemostaza // Mexatronika, avtomatizaciya, upravlenie: Teoreticheskij i prikladnoj nauchno-texnicheskij zhurnal. - 2009. - №8. – M., 2009. - S. 59 – 66.
5. Sidorova M.A., Serzhantova N.A., Filippova L.A. Diagnostika i prognozirovaniye isxodov peritonita s pomoshh'yu nejrosetevoj sistemy issledovaniya parametrov gemostaza // Medicinskaya texnika: Nauchno-texnicheskij zhurnal. - 2011. - № 2. - M., 2011. - S. 42 – 74.