

А. А. Жук

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ДООПЕРАЦИОННОЙ ДИАГНОСТИКИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

Научный руководитель: канд. биол. наук, доц. Т. Ю. Принькова

Кафедра биологической химии,

Белорусский государственный медицинский университет, г. Минск

Резюме. В данной работе предложена система предиктивных факторов, включающих опухолеассоциированные белки, показатели свертывания крови и эндогенной интоксикации для дооперационной диагностики рака тела матки. Сущность предложенной системы заключается в создании алгоритмов прогнозирования распространенности опухоли и риска прогрессирования заболевания с использованием искусственной нейронной сети. Проведена оценка диагностической эффективности созданной модели.

Ключевые слова: искусственная нейронная сеть, рак тела матки, онкомаркеры, дооперационная диагностика.

Resume. The system of predictive factors, including tumor-associated proteins, parameters of hemostasis and endogenous intoxication for preoperative diagnosis of uterine body cancer was created. The proposed system is the algorithms for predicting the tumor stage and the risk of tumor progression which were created using an artificial neural network. The diagnostic efficiency of the proposed model was estimated.

Keywords: artificial neural network, uterine body cancer, tumor markers, preoperative diagnosis.

Актуальность. Рак тела матки (РТМ) – одно из самых распространенных заболеваний, в Беларуси отмечается неуклонный рост заболеваемости и смертности от этой патологии [5]. Прогноз при РТМ и выживаемость во многом зависят от стадии заболевания на момент постановки диагноза и от возможности раннего выявления опухолевой прогрессии. Такая информация является определяющей при выборе индивидуальной тактики лечения (объема предстоящей операции, комплекса адъювантных и неoadъювантных терапевтических мероприятий). Несмотря на широкий спектр методов ранней диагностики клинической стадии РТМ, существующие на сегодняшний день критерии, не позволяют с уверенностью судить о распространенности опухоли до начала хирургического лечения. Поэтому актуальным является поиск информативных критериев неблагоприятного течения РТМ, которые позволили бы до начала хирургического лечения судить о распространении опухоли и выявить пациенток с высоким риском опухолевой прогрессии.

Искусственная нейронная сеть (ИНС) – модель, а также её программное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма. Главное преимущество такой системы – возможность обучения, в процессе которого нейронная сеть способна выявлять сложные зависимости между входными и выходными данными, которые невозможно запрограммировать в какой-либо вычислительный алгоритм. Такой «искусственный мозг» способен к анализу и переработке данных в режиме «человеческого мозга». Это открывает совершенно новые перспективы в диагностических исследованиях, в том числе с применением биохимических показателей [1].

Цель: разработать алгоритмы оценки распространенности опухоли и риска прогрессирования рака тела матки на дооперационном этапе лечения с использованием искусственной нейронной сети.

Задачи:

1. Оценить диагностическую эффективность исследуемых параметров крови в качестве предикторов распространенности и риска прогрессирования опухоли у пациенток, страдающих раком тела матки.

2. Разработать алгоритмы определения диагностически значимых параметров крови для дооперационной диагностики распространенности рака тела матки риска опухолевой прогрессии у пациенток, страдающих раком тела матки I стадии на основе искусственной нейронной сети.

3. Сопоставить результаты определения диагностической эффективности созданных на основе искусственной нейронной сети моделей и регрессионных математических моделей для дооперационной диагностики рака тела матки.

Материал и методы. В рамках исследования был проведён ретроспективный анализ ряда биохимических лабораторных показателей крови (опухолеассоциированных белков, показателей системы гемостаза и эндогенной интоксикации) у 107 пациенток с диагнозом РТМ, проходивших лечение на базе ГУ «РНПЦ онкологии и медицинской радиологии им. Н.Н. Александрова» в период с 2010 по 2013 годы [3]. Во всех случаях верифицирована гистологическим методом эндометриоидная аденокарцинома различной степени дифференцировки.

Для оценки взаимосвязи показателей крови со степенью распространенности РТМ были сформированы 2 группы пациенток. Первую группу составили 87 пациенток с нераспространенным РТМ (I-II стадии), во вторую группу вошли 20 пациенток с распространенным РТМ (III-IV стадии). Для определения степени риска опухолевой прогрессии пациентки с I стадией РТМ, были разделены на две подгруппы: в первую подгруппу вошли 55 пациенток с низким риском развития опухолевой прогрессии (IA стадия, G1-2); вторую подгруппу составили 20 пациенток с промежуточным (IA G3, IB G1-2) и высоким риском опухолевой прогрессии (IB G3).

Обработку данных производили с помощью программ Microsoft Excel, Statistica 8.0, MedCalc и использованием непараметрических методов вариационной статистики. Для изучения связей между показателями вычисляли коэффициент корреляции Спирмена (r). Построение прогностических моделей осуществляли с помощью метода логистической регрессии. Для создания алгоритма на основе ИНС использовали программу Neural Network Tools для Excel.

Результаты и их обсуждение. На первом этапе исследования был проведен анализ исследуемых показателей крови с распространенностью и риском прогрессирования РТМ. Наибольшая взаимосвязь по результатам корреляционного анализа была выявлена для следующих параметров: sICAM-1, VEGF, p185, ТБКРП, КПБ, ЭПР- α ($r > 0,25$; $p < 0,05$).

Обнаруженная взаимосвязь дала основание включить эти параметры в ROC-анализ, по результатам которого были рассчитаны значения диагностической эффективности (ДЭ) каждого из показателей в отдельности для оценки распространенности РТМ и риска опухолевой прогрессии соответственно (таблица 1).

Таблица 1. Диагностическая эффективность параметров крови пациенток, страдающих раком тела матки для оценки распространенности рака и риска опухолевой прогрессии

Маркер	ДЭ, % (для оценки распространенности)	ДЭ, % (для оценки риска опухолевой прогрессии)
sICAM-1 (фактор межклеточной адгезии)	84,1	60
VEGF (сосудисто-эндотелиальный фактор роста)	74,8	70,1
p185 (рецептор эпидермального фактора роста)	72,9	70,1
ТБКРП (продукты, реагирующие с тиобарбитуровой кислотой)	83,2	74,7
КПБ (карбонильные производные белков)	54,2	-
ЭПР-α (параметр, отражающий степень «загруженности» центров связывания альбумина)	81,3	49,3

С целью повысить диагностическую эффективность был проведен анализ наиболее информативных предикторов в совокупности. Для этого были применены метод логистической регрессии и метод построения ИНС.

При регрессионном анализе в уравнение вошли четыре наиболее информативных предиктора (sICAM-1, p185, КПБ, а также величина параметра ЭПР-α):

$$P1 = \frac{\exp(-5,96 + 2,29 \times \text{ЭПР}\alpha + 0,008 \times \text{sICAM1} + 0,091 \times \text{p185} - 4,13 \times \text{КПБ})}{1 + \exp(-5,96 + 2,29 \times \text{ЭПР}\alpha + 0,008 \times \text{sICAM1} + 0,09 \times \text{p185} - 4,13 \times \text{КПБ})}$$

где P1 – вероятность прогнозирования распространенного РТМ; ЭПР-α, sICAM-1, p185, КПБ – значения соответствующих показателей сыворотки крови пациенток, страдающих РТМ. Общий процент правильно классифицированных случаев нераспространенного РТМ (I-II стадии) и распространенного РТМ (III-IV стадии) с использованием данной модели составил 86,3% [4].

Для построения нейросетевой модели была выбрана простая искусственная нейронная сеть с обратной связью. На входные нейроны подавалось численное значение определенного предиктора, для каждого маркера был запрограммирован отдельный нейрон (рисунок 1). В качестве обучающей выборки для ИНС были взяты параметры крови 107 пациенток с верифицированным РТМ. На ее основе была построена обучающая база данных.

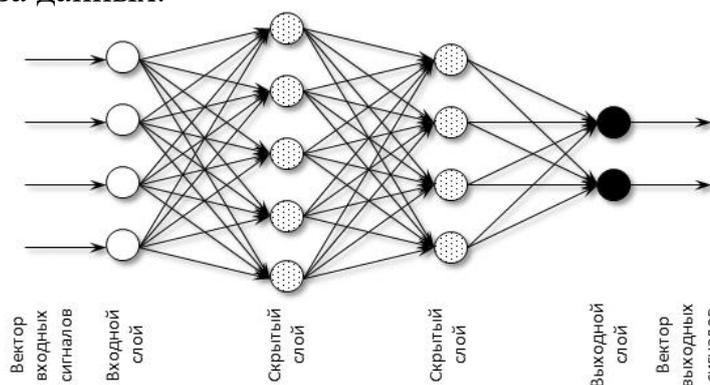


Рисунок 1 – Схема простой искусственной нейронной сети

ИНС обучали с использованием параметров крови, вошедших в регрессионную модель ((sICAM-1, p185, КПБ, ЭПР- α). По результатам анализа было установлено, что ИНС, обученная по тем же четырем показателям, правильно распределяет пациенток с нераспространенным и распространенным РТМ в 91,8% случаях (рисунок 2).

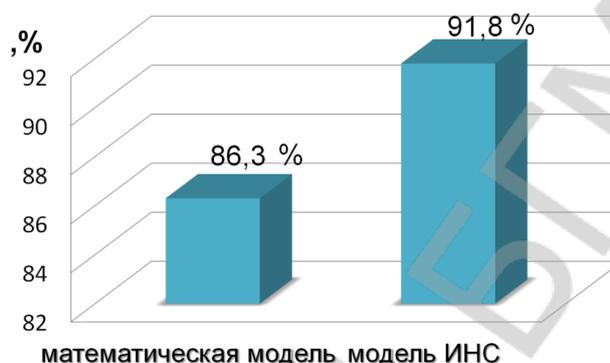


Рисунок 2 – Результаты сравнения диагностической эффективности регрессионной модели и модели ИНС для оценки распространенности РТМ по 4-м предикторам (sICAM-1, p185, КПБ, ЭПР- α)

При увеличении количества предикторов до шести процент правильно классифицированных случаев увеличивается до 94,6% (рисунок 3).

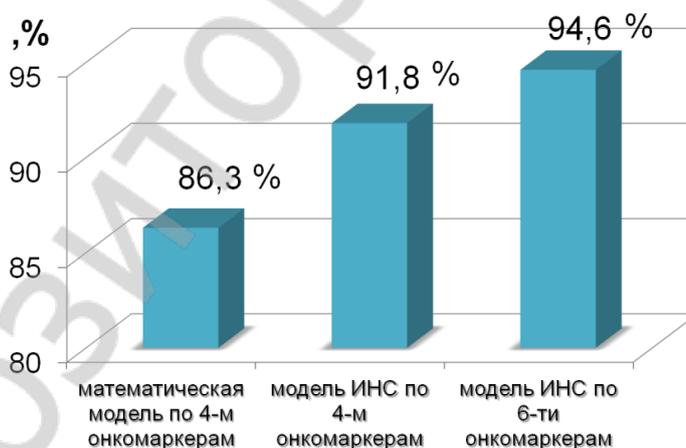


Рисунок 3 – Результаты сравнения диагностической эффективности регрессионной модели и модели ИНС для оценки распространенности РТМ по 6-м предикторам (sICAM-1, p185, КПБ, ЭПР- α , ТБК, VEGF)

При сопоставлении результатов анализа риска опухолевой прогрессии установлено, что математическая регрессионная модель оценки риска опухолевой прогрессии при раке тела матки I стадии, учитывающая определение в крови двух показателей (p185 и sICAM-1), обладает диагностической эффективностью 78,7% [5].

$$P2 = \frac{\exp(-4,73 + 0,0082 \times p185 + 0,0076 \times sICAM1)}{1 + \exp(-4,73 + 0,0082 \times p185 + 0,0076 \times sICAM1)}$$

где P2 – вероятность прогнозирования промежуточного или высокого риска опухолевой прогрессии у пациенток, страдающих РТМ I стадии; sICAM-1, p185 – значения соответствующих показателей сыворотки крови пациенток.

Обученная по тем же двум показателям ИНС правильно прогнозирует риск опухолевой прогрессии у пациенток с РТМ 1 стадии в 90,4% случаев (рисунок 4).

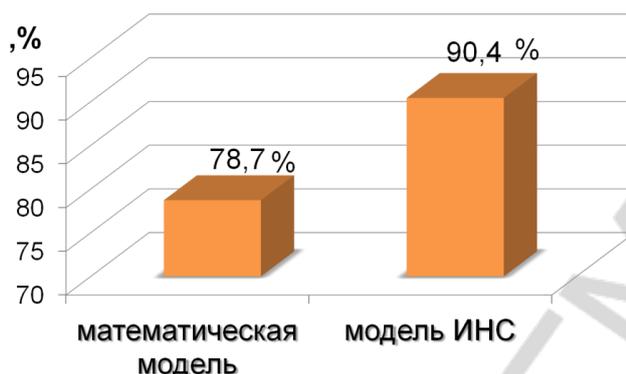


Рисунок 4 – Результаты сравнения диагностической эффективности регрессионной модели и модели ИНС для оценки опухолевой прогрессии при РТМ I стадии по 2-м предикторам (sICAM-1, p185)

Выводы:

1. Использование обученной нейронной сети позволяет правильно диагностировать распространенность опухоли при раке тела матки 94,6% случаях, а риск опухолевой прогрессии при раке тела матки I стадии в 90,4%.

2. Нейронные сети обладают преимуществом перед математическими регрессионными моделями и позволяют повысить диагностическую эффективность дооперационной диагностики рака тела матки.

A.A. Zhuk

PROSPECTS FOR THE USE OF ARTIFICIAL NEURAL NETWORK IN PREOPERATIVE DIAGNOSIS OF ONCOLOGICAL DISEASES

Tutor: assistant professor T. Y. Prinkova

Department of Biological Chemistry,

Belarusian State Medical University, Minsk

Литература

1. Anil K., Jain, Jianchang Mao, K.M. Mohiuddin. Artificial Neural Networks: A Tutorial, IEEE Computer, Vol.29, No.3, March/1996, p. 44.

2. Головкин В. А. Нейросетевые технологии обработки данных: учеб. пособие / В. А. Головкин, В. В. Краснопрошин. – Минск: БГУ, 2017 – 252 с.

3. Опухолеассоциированные белки, показатели свертывания крови и эндогенной интоксикации в оценке опухолевой прогрессии у женщин, страдающих раком тела матки / Т.Ю. Принькова, В.И. Прохорова, Т.П. Цырусь, Л.М. Шишло, Н.Н. Колядко, А.Д. Таганович // Инновации в медицине и фармации 2013: материалы науч.-практич. конф. молодых ученых / под ред. А.В. Сикорского, О.К. Кулаги, А.В. Стахейко, Т.В. Тереховой – г. Минск: БГМУ, 2013. – С. 58–67.

4. Принькова, Т.Ю. Использование лабораторных показателей в дооперационной оценке распространенности рака тела матки / Т.Ю. Принькова // Новые исследования молодых ученых 2013: сб. науч. работ / Белорусский гос. мед. ун-т; под общ. ред. А.В. Сикорского, О.К. Кулаги. – Минск, 2013. – С. 65–69.

5. Статистика онкологических заболеваний в Республике Беларусь (2007-2017) [Текст*]: [аналит. обзор по данным Белорус. канцер-регистра / А. Е. Океанов, П. И. Моисеев, Л. Ф. Левин; под ред. О. Г. Суконого]; Министерство здравоохранения РБ, ГУ «РНПЦ онкологии и медицинской радиологии им. Н. Н. Александрова». - Минск: РНПЦ ОМР им. Н. Н. Александрова, 2013. - 373 с.