

УДК 616.8-053.13/31:004.032.26

*Л.И. Гриненко, Ю.Е.Лях, В.Г.Гурьянов***ВЫБОР НАИБОЛЕЕ ЗНАЧИМЫХ НЕБЛАГОПРИЯТНЫХ ФАКТОРОВ, ВЛИЯЮЩИХ НА ВОЗНИКНОВЕНИЕ ПЕРИНАТАЛЬНОЙ ПАТОЛОГИИ ЦЕНТРАЛЬНОЙ НЕРВНОЙ СИСТЕМЫ МЕТОДОМ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ**

Военно-врачебная комиссия УМВД Украины в Донецкой области, Донецкий медицинский университет им. Горького

Ключевые слова: дети, перинатальное поражение ЦНС, нейросетевое моделирование.

Несмотря на то, что изучению перинатальной патологии ЦНС уделялось и уделяется огромное внимание [1- 6, 8, 12] эта проблема не потеряла своей актуальности. Это объясняется прежде всего тем, что в большинстве случаев отклонения в нервно-психическом развитии у детей своими корнями уходят в перинатальный период; а так же тем, что заболевания нервной системы чаще, чем другая патология, становятся причиной детской инвалидности. Перинатальное поражение ЦНС – общая группа состояний, синдромов и заболеваний нервной системы плода и новорожденного, развивающихся вследствие действия повреждающих факторов в перинатальном периоде. Существуют различные классификации факторов риска перинатальной патологии. Число их колеблется от 40 до 100 (Б.В.Лебедев, 1995). В то же время значимость отдельных факторов пре- и перинатального периода недостаточно определены. С этой целью и было проведено данное исследование.

Для прогнозирования вероятности развития перинатальной патологии ЦНС был использован метод нейросетевого моделирования [10]. В последние десятилетия в мире бурно развивается новая прикладная область математики, специализирующаяся на искусственных нейронных сетях (НС). Актуальность исследований в этом направлении подтверждается массой различных применений НС. Это автоматизация процессов распознавания образов, адаптивное управление, организация ассоциативной памяти и многие другие приложения. Модели НС могут быть программного и аппаратного исполнения. В дальнейшем речь пойдет в основном о первом типе. Несмотря на существенные различия, отдельные типы НС обладают несколькими общими чертами.

Во-первых, основу каждой НС составляют относительно простые, в большинстве случаев

– однотипные, элементы (ячейки), имитирующие работу нейронов головного мозга. Далее под нейроном будет подразумеваться искусственный нейрон, то есть ячейка НС. Каждый нейрон характеризуется своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов.

Каждый синапс характеризуется величиной синаптической связи или ее весом w_i , который по физическому смыслу эквивалентен электрической проводимости. Текущее состояние нейрона определяется, как взвешенная сумма его входов: $s = \sum x_i * w_i$. Выход нейрона есть функция его состояния: $y = f(s)$.

Возвращаясь к общим чертам, присущим всем НС, отметим, во-вторых, принцип параллельной обработки сигналов, который достигается путем объединения большого числа нейронов различных слоев, а также, в некоторых конфигурациях, и нейронов одного слоя между собой, причем обработка взаимодействия всех нейронов ведется послойно.

Очевидно, что процесс функционирования НС, то есть сущность действий, которые она способна выполнять, зависит от величин синаптических связей, поэтому, задавшись определенной структурой НС, отвечающей какой-либо задаче, разработчик сети должен найти оптимальные значения всех переменных весовых коэффициентов. Этот этап называется обучением НС. Великое множество различных алгоритмов обучения делятся на два больших класса: детерминистские и стохастические. В первом из них подстройка весов представляет собой жесткую

последовательность действий, во втором – она производится на основе действий, подчиняющихся некоторому случайному процессу. Важно отметить существование бинарных и аналоговых сетей. Первые из них оперируют с двоичными сигналами, и выход каждого нейрона может принимать только два значения: логический ноль («заторможенное» состояние) и логическая единица («возбужденное» состояние). В аналоговых сетях выходные значения нейронов способны принимать непрерывные значения. Еще одна классификация делит НС на синхронные и асинхронные. В первом случае в каждый момент

времени свое состояние меняет лишь один нейрон. Во втором – состояние меняется сразу у целой группы нейронов, как правило, у всего слоя. Однако описание всего многообразия существующих НС выходит за рамки данной статьи.

Применение НС в неврологии, на наш взгляд, поможет более реально оценить влияние тех или иных причин на патологию ЦНС. С целью выявления значимых факторов пре- и перинатального периода на последующее развитие ребенка в возрасте до 1 года и состояние его нервной системы нами и было проведено данное исследование.

Материал и методы исследования

Для выявления и изучения основных негативных факторов, влияющих на возникновение перинатальной патологии ЦНС нами было обследовано по специально составленной схеме 102 ребёнка до 1 года (из них 50 мальчиков и 52 девочки), которые прошли обследование и лечение в центре реабилитации больных спинномозговой травмой и церебральным параличом г.Макеевки.

В качестве входных параметров для построения модели использовались результаты стандартного анкетирования. В анкету входило 173 пункта. Все вопросы, входящие в анкету были разделены на 3 группы (I, II, III):

I) вопросы, относящиеся к родовому периоду, которые характеризовали, в основном, состояние здоровья родителей и протекание беременности;

II) вопросы, относящиеся к протеканию родов и состоянию ребенка в первые дни его жизни;

III) вопросы, относящиеся к развитию ребенка в первые месяцы (до 1 года).

Для построения модели было проведено предварительное шкалирование вариантов ответов. При этом для большей части вопросов ответы были сведены к вариантам: нет - 0, да - 1. В случае, когда вариантов ответов было более 2, они были упорядочены по степени отклонения от ответа нет (0); каждому варианту ответа было поставлено в соответствие число, равное его рангу. Так, например, на вопрос анкеты: «Курит ли мать?», - были следующие варианты ответов:

1) нет - 0; 2) до беременности - 1; 3) во время беременности - 2; 4) и до и во время беременности - 3;

Ответы на 170 пунктов вопросов были выбраны в качестве входных переменных – прогнозирующих признаков. В качестве выходных пе-

ременных – прогнозируемых признаков было выбрано 3 пункта (которые далее будем именовать Res1, Res2, Res3):

Res1) Какой диагноз был выставлен неврологом в родильном доме?

А) Здоров - 0. Б) Перинатальное поражение ЦНС - 1.

Res2) Какой диагноз выставлен в возрасте 1 года?

А) Здоров - 0. Б) Болен - 1.

Res3) Ведущий синдром.

А) Здоров - 0. Б) Перинатальная энцефалопатия - 1. В) Минимальная мозговая дисфункция - 2. Г) Детский церебральный паралич - 3.

При этом при прогнозировании Res1 были использованы только входные переменные I и II группы, при прогнозировании Res2 и Res3 – переменные I, II, и III групп и результаты Res1.

Независимо от выбора метода, применяемого для прогнозирования выходных параметров, одним из важных шагов при построении модели является отбор входных переменных, наиболее значимо, влияющих на выходную переменную. Кроме того, что это улучшит прогностические способности модели, отбор значимых входных признаков, позволит лучше понять суть моделируемого явления и эффективно управлять прогнозируемой величиной.

В работе Айвазян С.А. и соавторов [9] показано, что включение в прогностическое правило малоинформативных переменных может заметно ухудшить его качество. Каждый признак наряду с положительным вкладом в разделение несет в себе в силу ограниченности выборки и шумовую (случайную) составляющую. Если много малоинформативных признаков, то отношение сигнал/шум для небольшой группы высокоинформативных признаков может быть значительно выше, чем для всей группы признаков.

Отсюда авторами сделан вывод, что при относительно небольшом объеме выборки малоинформативные признаки в прогностическое правило лучше не включать [9]. В Кратком руководстве по SNN [7] сделан парадоксальный на первый взгляд вывод о том, что при построении модели иногда бывает лучше с целью понижения размерности отбросить даже признаки, несущие полезную информацию.

В связи с вышеизложенным: вопросу выбора значимых входных переменных мы посветили следующий раздел работы.

2. Отбор переменных. Понижение размерности.

2.1. Схема последовательного испытания наборов признаков.

Следуя принципам изложенным выше [9] изложим традиционную схему последовательного испытания набора признаков, которая была использована нами при отборе значимых переменных из некоторого их набора:

1. Была выбрана функция потерь, характеризующая ошибку построенного прогноза $Q(\Omega)$, где Ω - набор прогнозирующих признаков.

2. Для каждого набора переменных строился наилучший (в смысле Q) критерий классификации.

3. Среди всех построенных наборов Ω отбирался тот (те), в который входило наименьшее число переменных и при котором Q минимально.

Схемы генерации наборов переменных были аналогичны схемам, которые используются в

регрессионном анализе и опираются на эвристическое предположение, что наилучший набор из $k+1$ переменных часто содержит в себе наилучший набор из k переменных. Одним из стандартных методов в этом случае является корреляционный анализ [11], результаты применения которого будут приведены ниже. Этот метод используется для количественной оценки взаимосвязи двух наборов данных, представленных в безразмерном виде. Коэффициент корреляции выборки представляет собой ковариацию двух наборов данных, деленную на произведение их стандартных отклонений.

Однако в общем случае пошаговые процедуры не гарантируют получения оптимального набора переменных, и тогда полный перебор неизбежен. В случае полного перебора всех возможных наборов переменных приходится сталкиваться с так называемой проблемой «проклятия размерностей», так, например, в нашем случае полный перебор включает в себя исследование $2^{170} \approx 1.5 \cdot 10^{51}$ моделей, что неосуществимо даже с технической точки зрения. В последнее время при решении подобных задач обращаются к генетическим алгоритмам, которые, сокращая количество переборов до вполне разумных значений, позволяют решить задачу с приемлемым уровнем точности. Эти методы «генетического алгоритма отбора значимых переменных» изложены, например, в [7] будут также использованы далее при решении поставленной задачи.

Результаты исследования и их обсуждение

Результаты применения корреляционного анализа. С целью понижения размерности модели прогнозирования нами был проведен корреляционный анализ переменных.

Предварительный анализ позволил выделить признаки, имеющие нулевую выборочную дисперсию. Таких признаков оказалось 47. Они не были включены в модель и не несли в себе никакой информации для исследуемой выборки. После отбрасывания этих переменных в модели осталось 123 входных переменных.

В качестве критерия отбора значимых переменных были выбраны значения коэффициента корреляции данной переменной с признаками Res_2 и Res_3 .

В таблице 1 приведен список переменных, наиболее тесно связанных с прогнозируемыми (здесь R – коэффициент корреляции).

Следует обратить внимание на тот факт, что

наиболее значимыми оказались только переменные, относящиеся ко II и III группе. Единственная переменная, относящаяся к группе I – V1, имеющая сильную отрицательную связь, (вопрос: «Удовлетворены ли Вы развитием своего ребенка?») в дальнейшем из-за сильной субъективности оценки была исключена нами из прогнозирующих признаков.

Что касается переменных, коэффициент корреляции которых с прогнозируемыми признаками оказался $R > 0.7$ (коэффициент корреляции отличен от 0 на уровне значимости $p < 0.001$) то они таковы:

V81 – «Как оценен новорожденный по шкале Апгар? »;

V104 – «В каком возрасте ребенок стал держать голову? »;

V106 – «Когда стал сидеть? »;

V107 – «Самостоятельно садиться?»;

- V108 – «Самостоятельно вставать?»;
 V109 – «Самостоятельно ходить?»;
 V113 – «Как спал в течении 1-го года?».

Таким образом, в результате проведенного отбора значимых переменных стандартными методами корреляционного анализа нам удалось сформировать группу наиболее значимых пере-

менных только из III группы признаков, которые являются с одной стороны достаточно сильными, но с другой стороны достаточно очевидными факторами. Поэтому, для построения прогноза на более раннем этапе был проведен анализ входных данных с привлечением генетических алгоритмов.

Таблица 1

Список переменных, наиболее тесно связанных с прогнозируемыми

Входная	R	Выходная											
Res2	R>0.3 (p<0,01)	V69	V69b	V69f	V74	V81	V82	V83	V85	V86	V87	V91	V93
Res2	R>0.3 (p<0,01)	V94	V96	V104	V105	V106	V107	V108	V109	V110	V111	V112	V113
Res2	R<-0.3 (p<0,01)	V1	V75										
Res2	R>0,7 (p<0,01)	V104	V106	V107	V108	V109							
Res3	R>0.7 (p<0,01)	V81	V106	V107	V108	V109	V113						

2.3. Результаты применения генетического алгоритма отбора значимых переменных.

Для применения генетического алгоритма отбора значимых переменных необходимо иметь возможность построения простых моделей прогнозирования входных переменных. Эта возможность была реализована нами с помощью пакета анализа «Statistica Neural Networks», где существует возможность проведения такого отбора. Для этого были построены достаточно простые в обучении «Вероятностные нейронные сети» [7] с различными наборами входных переменных, для каждого набора устанавливался штраф за каждую входную переменную, который прибавлялся к ошибке прогнозирования модели. После проведения отбора выделялся набор переменных, имеющий наименьшую ошибку прогнозирования. Так как отбор наиболее значимых переменных носил в этом случае не абсолютный, а статистический характер, то отбор был проведен многократно, с различным значением штрафа за входную переменную.

По I группе признаков анализировались модели, содержащие до 65 входных переменных. В качестве прогнозируемой переменной была выбрана переменная Res1. Ниже приведен набор из 12 переменных, которые оказались наиболее значимыми в этой группе.

V5 – «Страдает ли мать ребенка соматичес-

кими заболеваниями?».

V19B – «Инфекции матери, вызванные простейшими?».

V20 – «В каком триместре беременности протекал инфекционный процесс?».

V32 – «Являлась ли мать ребенка во время беременности пассивным курильщиком?».

V41 – «Какая по счету данная беременность?».

V44B – «Медицинские аборт до настоящей беременности?».

V48 – «Выраженный ранний токсикоз беременной?».

V49 – «Поздний токсикоз беременной?».

V50 – «Анемия матери во время беременности?».

V51 – «Угроза прерывания в I-II половине беременности?».

V61 – «Принимала ли беременная лекарственные препараты?».

V62 – «В каком триместре беременности принимала лекарственные препараты?».

По II группе признаков анализировались модели, содержащие до 38 входных переменных. В качестве прогнозируемых переменных были выбраны переменные Res2, Res3. Ниже приведен набор из 10 переменных, которые были отобраны по этой группе, как наиболее весомые.

V69B – «Преждевременные роды».

V69E – «Раннее излитие околоплодных вод?».

V74 – «В каком сроке гестации родился ребенок?».

V81 – «Как оценен ребенок по шкале Апгар?».

V83 – «Крик?» (громкий, тихий).

V85 – «Наблюдался ли в течении первых часов жизни повторный цианоз?».

V91 – «Как протекал период послеродовой адаптации?» (гладко, с патологией).

V93 – «На какие сутки выписан новорожденный?».

V94 – «Был ли выставлен новорожденному соматический диагноз?».

V96 – «Проводилась ли медикаментозная терапия в остром периоде?».

Можно отметить, что этот набор признаков, в основном совпадает с набором, полученных методом корреляционного анализа ($R > 0.3$, коэффициент корреляции отличен от 0 на уровне значимости $p < 0.01$).

По III группе признаков нами анализировались модели, содержащие до 19 входных переменных. В качестве прогнозируемых переменных были выбраны переменные Res2, Res3. Ниже приведен набор из 7 наиболее значимых переменных, которые были отобраны по этой группе.

V97 – «Болел ли ребенок соматическими заболеваниями в течении 1-го месяца жизни?».

V106 – «Когда стал сидеть?».

V107 – «Самостоятельно садиться?».

V108 – «Самостоятельно встать?».

V109 – «Самостоятельно ходить?».

V112 – «Как спал в течении 1-го месяца?».

V113 – «Как спал в течении 1-го года?».

В этом случае также можно отметить совпадение результата отбора с методом корреляционного анализа ($R > 0.7$, коэффициент корреляции отличен от 0 на уровне значимости $p < 0.001$).

Таким образом, в результате проведенного отбора значимых переменных с применением генетического алгоритма, с помощью методов нейросетевого моделирования, сформированы группы наиболее значимых переменных для I, II и III группы признаков. При этом результаты отбора по II и III группам переменных совпадают с таковыми же для стандартных методов что свидетельствует о надежности результата отбора. Кроме того мы попытались выявить из отобранных переменных те, которые наиболее часто встречаются и сочетаются между собой и в конечном итоге приводят к возникновению того или иного патологического состояния. В каче-

стве прогнозируемых диагнозов были выбраны - Перинатальная энцефалопатия и Детский церебральный паралич, т.е. переменная R3. Кроме того мы попытались проанализировать наличие и сочетание этих факторов у здоровых детей контрольной группы. У детей с грубой задержкой психо-моторного развития наиболее часто встречаемыми факторами были анемия матери во время беременности (у 90% обследуемых) и наличие раннего токсикоза во время беременности (76%). Сочетание этих факторов было отмечено нами в 66% случаев. Интересно, что в 71% случаев такие дети рождались у матерей, имеющих повторную беременность и наличие хотя бы одного мед. аборта в анамнезе (66% случаев). Что касается II группы вопросов, то наибольшее количество раз встречалось следующее сочетание факторов – низкая оценка по шкале Апгар, тихий крик ребенка при рождении и получение интенсивной медикаментозной терапии в остром периоде. Поскольку у всех детей отмечалась грубая задержка развития, то 6 показателей из 7 имели место у 100% обследуемых детей в III группе вопросов. Необходимо отметить, что в 85% случаев отмечалось сочетание выше указанных факторов из I, II и III групп вопросов.

Анализируя группу детей, которым был выставлен диагноз - Перинатальная энцефалопатия в I группе фигурировали те же факторы, что и в группе детей с ДЦП. Однако по второй группе признаков в большинстве случаев имело место сочетание следующих факторов – низкая оценка по шкале Апгар, тяжело протекающий период адаптации и наличие соматического диагноза у ребенка. Обращает на себя внимание, что если у детей с ДЦП оценка по шкале Апгар 0-3 б. имела место в 50% случаев, то у детей с ПЭ только в 20% случаев. Относительно III группы нужно сказать, что отставание в психо-моторном развитии было менее выражено. Если сравнивать детей (ДЦП и ПЭ), то их четкое разделение на две группы – дети с выраженными нарушениями и дети с относительной нормализацией изменений со стороны нервной системы, отмечалось нами с 6 по 8 месяц жизни ребенка, что совпадает с данными, представленными Ю.И. Барашневым, 1999г. (четыре фазы развития патологического процесса на протяжении первого года жизни). У здоровых детей контрольной группы наиболее часто встречалась в первой группе вопросов анемия матери во время беременности. Это говорит о том что наличие какого либо одного патологического фактора оказывает существен-

ное значение только в сочетании с другими факторами. Что касается II и III группы вопросов, то у здоровых детей нами не отмечено выраженных особенностей.

В заключении, хотелось бы сказать, что в ко-

нечном итоге предполагается построение математической модели прогнозирования развития перинатальной патологии ЦНС, и данная работа является лишь начальным этапом осуществления этой идеи.

Л.І. Гриненко, Ю.Є. Лях, В.Г. Гур'янов

ВИБІР НАЙБІЛЬШ ЗНАЧУЩИХ НЕСПРИЯТЛИВИХ ФАКТОРІВ, ЯКІ ВПЛИВАЮТЬ НА ВИНИКНЕННЯ ПЕРІНАТАЛЬНОЇ ПАТОЛОГІЇ ЦЕНТРАЛЬНОЇ НЕРВОВОЇ СИСТЕМИ МЕТОДОМ НЕЙРОСЕТЬОВОГО МОДЕЛЮВАННЯ

Військово-лікарська комісія УМВС України в Донецькій області, Донецькій державний медичний університет ім. М. Горького

В роботі за допомогою методів нейросетєвого моделювання був проведений аналіз різноманітних факторів пре- та перенатального періоду у 102 дітей до 1 року з патологією ЦНС. Були виділені фактори, які найбільш часто зустрічаються, злучаються між собою та призводять до виникнення цієї патології. (Журнал психіатрії та медичної психології. — 2004. — № 2 (12). — С.69-74)

L.I. Grinenko, U.E. Lah, V.G. Gurianov

CHOICE OF THE MOST SIGNIFICANT OF THE ADVERSE FACTORS INFLUENCING ON OCCURRENCE PERINATAL PATHOLOGY OF CENTRAL NERVOUS SYSTEMS BY THE METHOD OF NEURONET MODELING

Health Protection Department of Interior Department of Ukraine in Donetsk region, Donetsk medical university by M. Gorkiy

In work with the help of methods нейросетєвое of modeling the analysis of the various factors pre- and perinatal period at 102 children till 1 year with a pathology of CNS is carried out. The meeting and combined among themselves factors are allocated most frequently which result in occurrence of this pathology. (The Journal of Psychiatry and Medical Psychology. — 2004. — № 2 (12). — P.69-74)

Литература

1. Бадалян Л.О. Детская неврология.— М.: Медицина, 2000.— 676 с.
2. Бадалян Л.О. и др. Руководство по неврологии раннего детского возраста/ Бадалян Л.О., Журба Л.Т., Всеволожская Н.М.— К.: Здоров'я, 1980.— 528 с.
3. Барашнев Ю.И. Гипоксически-ишемическая энцефалопатия новорожденных: вклад перинатальных факторов, патогенетическая характеристика и прогноз// Рос. вестн. перинатологии и педиатрии.— 1996.— Т. 41, № 5.— С. 29-34.
4. Евтушенко С.К. и др. Ранняя клиническая диагностика моторной, психической и речевой задержки у детей в возрасте до 1 года/ Евтушенко С.К., Шестова Е.П., Евтушенко О.С.— Донецк, 1995.— 66 с.
5. Журба Л.Т., Мастюкова Е.М. Нарушение психомоторного развития детей первого года жизни.— М.: Медицина, 1981.— 272 с.
6. Кирющенко А.П. Влияние вредных факторов на плод.— М.: Медицина, 1978.— 213 с.
7. Нейронные сети: Крат. рук. по SNN, STATISICA, StatSoft.— М., 1998.— 355 с.
8. Пальчик А.Б., Шабалов Н.П. Гипоксически-ишемическая энцефалопатия новорожденных.— СПб: Питер Ком, 2000.— 224 с.
9. Прикладная статистика. Классификация и снижение размерности: Справ. изд./ С.А.Айвазян, В.М.Бухштабер, И.С.Енюков, Л.Д.Мешалкин; Под ред. С.А.Айвазяна.— М.: Финансы и статистика, 1989.— 607 с., ил.
10. Применение математических методов в исследованиях по физиологии человека/ Под ред. В.Н.Казакова.— Донецк: Изд-во Донец. мед. ун-та, 2000.— 84 с.— (Очерки биолог. и мед. информатики).
11. Статистические методы для ЭВМ/ Пер. с англ.; Под ред. М.Б.Малотова.— М.: Наука, 1986.— 464 с.
12. Шабалов Н.П. Асфиксия новорожденных. — Л. Медицина. — 1990.— 192 с.

Поступила в редакцию 3.06.2004