

УДК 531/534:[57+61]

В.Д. Семакина, В.С. Туктамышев

V.D. Semakina, V.S. Tuktamyshev

Пермский национальный исследовательский
политехнический университет

Perm National Research Polytechnic University

**ПРОГНОЗИРОВАНИЕ ПАТОЛОГИЙ
БЕРЕМЕННОСТИ НА ОСНОВЕ ДАННЫХ
ХИМИЧЕСКОГО СОСТАВА КРОВИ**

**PREDICTION OF PREGNANCY
PATHOLOGIES ON THE BASIS OF DATA
ON THE CHEMICAL COMPOSITION OF BLOOD**

Организм женщины дает будущему ребенку все необходимые вещества и микроэлементы через плаценту, которая является неким барьером между плодом и окружающей средой. Важна также среда, в которой плод находится. Необходимо, чтобы состояние вод и плаценты были в норме. Даже самые небольшие отклонения характеристик от нормы могут повлечь за собой тяжелые последствия, негативно отразиться на здоровье матери и будущего ребенка. Наиболее часто встречающимися патологиями при беременности является фетоплацентарная недостаточность (ФПН) и дородовое излитие околоплодных вод (ДИВ).

Ключевые слова: статистический анализ, химический состав крови, дискриминантные переменные, беременность.

A woman's body gives the fetus all the necessary substances and trace elements via the placenta, which is a kind of barrier between the fruit and the environment. It is also important to the environment in which the fruit is. It is necessary that the status of waters and placenta were normal. Even the smallest characteristic deviation from the norm can lead to serious consequences on the course and outcome of pregnancy and health of mother and unborn child. The most common pathologies in pregnancy is fetoplacental insufficiency (FPI) and prenatal rupture of membranes (DIV).

Keywords: statistical analysis, chemical composition of blood, discriminant variables, pregnancy.

Беременность является одним из самых прекрасных периодов в жизни женщины. Не секрет, что в это время в организме матери происходит значительная перестройка, начиная от общего состояния и самочувствия, заканчивая гормональными изменениями. Организм женщины дает будущему ребенку все необходимые вещества и микроэлементы через плаценту, которая является неким барьером между плодом и окружающей средой. Также важна среда, в которой плод находится. Необходимо, чтобы состояние вод и плаценты были в норме. Даже самые небольшие отклонения характеристик от нормы могут повлечь за собой тяжелые последствия, негативно отразиться на

здоровье матери и будущего ребенка. Наиболее часто встречающимися патологиями при беременности является фетоплацентарная недостаточность (ФПН) и дородовое излитие околоплодных вод (ДИВ). Причины возникновения рассматриваемых патологий до конца не изучены.

Актуальность выполняемой работы заключается в исследовании развития беременности с помощью методов математической статистики. Это поможет определить женщин, наиболее подверженных риску развития ФПН и ДИВ, и назначить необходимую терапию. Ранее исследования проводились исключительно в медицинской области, и другие сферы науки для изучения данной проблемы не были использованы.

Таким образом, целью настоящего исследования является поиск статистических зависимостей между патологиями беременности (ФПН, ДИВ) и показателями ее протекания.

Материалы и методы. Проводятся данные исследования с помощью дискриминантного анализа и математической статистики. Для достижения поставленной цели необходимо выполнение следующих задач:

- 1) выявление наиболее значимых характеристик содержания химических элементов в крови исследуемых женщин;
- 2) установление статистической зависимости развития тяжелых патологий течения беременности от отклонения от нормы найденных характеристик;
- 3) с помощью пакета прикладных программ STATISTICA построить дискриминантную функцию, позволяющую определить вероятность развития патологии по известным данным о содержании химических элементов в крови у беременной.

Программный пакет STATISTICA, используемый для выполнения работы, предназначен для статистического анализа, реализует функции анализа, управления, добычи и визуализации данных с привлечением статистических методов. С его помощью мы сможем установить, какие патологии развития беременности влияют на риск возникновения осложнений.

На начальном этапе сбора информации были определены самые часто встречающиеся виды осложнений. Основной этап работ был произведен с помощью дискриминантного анализа. Данный вид анализа представляет собой набор методов статистического анализа для решения задач, который используется для принятия решения о том, какие переменные дискриминируют (т.е. разделяют) возникающие наборы групп данных.

В 2003 г. в работах В.А. Дюка было представлено, как информационные технологии могут быть использованы в биологии и медицине [1, 2].

Ранее проводимые исследования в медицине были представлены в работе В.П. Колосова [3], в которой изучались пути прогнозирования в пульмонологии. Более поздней работой является исследование И.Б. Петрова в 2009 г. [4], где изучались математические модели в медицине и биологии.

В работе А.В. Вильдемана исследование проводилось методом дискриминантного анализа при измерении индекса моторики у больных с низким уровнем развития двигательных навыков [5, 6]. Были рассмотрены группы детей с детским церебральным параличом. В данной работе в число статистических переменных вошли возраст больного, текущие характеристики его состояния (обобщенные показатели кардиограммы и др.).

Новизна представленной работы заключается в применении метода дискриминации в акушерстве и гинекологии [7]. Ранее он не был задействован в данной области. Полученные в ходе исследования алгоритмы могут применяться в качестве дополнительных методов исследования с целью выявления патологий у беременных на ранних стадиях или же предотвращения их развития.

В настоящей работе в число статистических переменных вошел полный химический состав крови матери. Подготовкой к работе являлась выборка информативных характеристик, т.е. не рассматривались химические элементы, которые в исходных данных не отклонялись от нормы или отклонения были у незначительного процента женщин, что делало данный элемент неинтересным для статистики.

Обработка и измерения. В представленном исследовании был использован метод дискриминантного анализа. Дискриминантный анализ – это раздел статистики, содержанием которого является разработка методов решения задач дискриминации объектов наблюдения по определенным признакам. Впервые в биологии метод дискриминантного (статистического) анализа был применен в ботанике [7].

Рассматриваемый метод, предназначенный для изучения отличий между двумя или большим количеством групп объектов на основании данных о разнообразии нескольких признаков, отличающих эти объекты друг от друга.

Типичная для дискриминантного анализа задача – определение тех признаков, которые лучше всего дискриминируют или отличают объекты, разделяющие их на разные группы. После того как определены наилучшие способы дискриминации имеющихся групп, этот способ анализа позволяет проводить классификацию образцов, принадлежность которых к той или иной группе заранее неизвестна. Дискриминантный анализ разработал Рональд Фишер.

Пояснить, как может быть использован данный анализ, проще на примерах. Предположим, важны различия между женскими и мужскими скелетами. Рассмотрим совокупность скелетов мужчин и женщин и определим, какие признаки лучше всего разделяют эти группы. После этого можно использовать полученные результаты для определения половой принадлежности скелетов, пол которых нам изначально неизвестен.

В алгоритме дискриминантного анализа рассматривается многомерное пространство признаков, в котором расположены изучаемые объекты. В данном пространстве выбирается такая дискриминантная функция, которая в

наибольшей степени отражает различия между группами выбранных объектов. Данная процедура напоминает такую процедуру, которая используется при анализе главных компонент, за тем исключением, что при компонентном анализе выбираются главные компоненты, на которые проецируется большая часть информации о различии всех объектов, а дискриминантный анализ максимизирует отличия между заданными группами, известными заранее. После того как первая такая функция выбрана, на основании оставшейся информации выбирается вторая дискриминантная функция.

Исходные данные. Таблицы с данными, с которыми проводилась работа, содержат информацию о состоянии здоровья женщин, которые содержали в себе клинические данные, т.е. информацию о хронических заболеваниях женщины, а также о прошлых беременностях и абортах. Кроме того, присутствует информация о психическом состоянии беременных, включающая в себя возраст, уровень образования, семейное положение, занимаемую должность и полный анализ химического состояния крови.

Ранее упомянутые таблицы, которые содержали полный анализ состояния здоровья у 57 женщин, были предоставлены Центром холистической медицины (г. Пермь). Одной из рассматриваемых проблем является то, что таблицы, используемые для данной работы как исходные данные, содержат огромное количество характеристик. Необходимо для удобной работы использовать только те показатели, которые отклоняются от нормы у достаточного количества беременных. Соответствующий отбор был проделан на начальном этапе исследования.

В настоящей работе были выбраны химический состав анализа крови и показатели наличия выбранных для дискриминации патологий беременности (ДИВ и ФПН). Общий объем исследуемой статистической выборки составляют наблюдения 57 беременных женщин, состояние которых описывалось 19 факторами (табл. 1).

Таблица 1

Дискриминантные переменные

№ п/п	Обозначение	Наименование	Диапазон изменения
1	x_1	Fe>	0–4
2	x_2	I>	0–4
3	x_3	K>	0–3
4	x_4	Cd>	0–3
5	x_5	Ca>	0–4
6	x_6	Co>	0–2
7	x_7	Co<	0–4
8	x_8	Si>	0–2
9	x_9	Mg>	0–4

Окончание табл. 1

№ п/п	Обозначение	Наименование	Диапазон изменения
10	x_{10}	Mn<	0–3
11	x_{11}	Cu>	0–3
12	x_{12}	As<	0–1
13	x_{13}	Na>	0–4
14	x_{14}	Se>	0–3
15	x_{15}	Sr<	0–3
16	x_{16}	P>	0–4
17	x_{17}	Cr	0–2
18	x_{18}	Zn>	0–3
19	x_{19}	Zn<	0–3
20	–	ФПН	0–1

Результаты. Построение математических моделей прогнозирования в рамках статистического подхода предполагает прохождение следующих этапов:

1) выбор прогнозируемого показателя, который наиболее полно отражает исследуемые аспекты изучаемого объекта;

2) определение группы факторов, оказывающих значимое влияние на прогнозируемый показатель;

3) проверка адекватности модели прогнозирования.

Функции G_1 и G_2 , приведенные ниже, показывают значение проведения дискриминации по выбранной группирующей переменной. В первом случае группирующей является фетоплацентарная недостаточность (ФПН). При этом первая функция ($G_1^{\text{ФПН}}$) показывает наличие ФПН, вторая ($G_2^{\text{ФПН}}$) – ее отсутствие. Определение дискриминантных функций или линейных комбинаций независимых переменных, которые наилучшим образом различают (дискриминируют) категории (группы) зависимой переменной, и есть дискриминантная функция [1].

$$G_1^{\text{ФПН}} = k_{11}x_1 + k_{12}x_2 + \dots + k_{1n}x_n,$$

$$G_2^{\text{ФПН}} = k_{21}x_2 + k_{22}x_2 + \dots + k_{2n}x_n.$$

Аналогичные функции были получены для группирующей ДИВ. Первая функция ($G_1^{\text{ДИВ}}$) показывает наличие патологии, вторая ($G_2^{\text{ДИВ}}$) говорит о ее отсутствии.

$$G_1^{\text{ДИВ}} = k_{11}x_1 + k_{12}x_2 + \dots + k_{1n}x_n,$$

$$G_2^{\text{ДИВ}} = k_{21}x_2 + k_{22}x_2 + \dots + k_{2n}x_n.$$

В результате применения многомерного дискриминантного анализа в задаче прогнозирования развития патологий фетоплацентарной недостаточности и дородового излития вод были получены таблицы коэффициентов (табл. 2, 3). Получение данных коэффициентов в представленной работе является наиболее важным этапом, так как в дальнейшем они используются в качестве компонентов дискриминантных функций.

Таблица 2

i	x_i	k_{1i}	k_{2i}
1	Fe>	-0,74	-0,97
2	I>	-0,01	0,21
3	K>	0,67	1,05
4	Cd>	2,53	2,89
5	Ca>	0,54	0,07
6	Co>	0,67	2,47
7	Co<	-0,35	0,79
8	Si>	2,85	1,89
9	Mg>	1,61	0,09
10	Mn<	0,49	0,57
11	Cu>	-0,11	0,44
12	As<	-0,18	2,26
13	Na>	0,07	0,37
14	Se>	0,39	-0,02
15	Sr<	2,72	1,71
16	P>	1,33	1,43
17	Cr	0,68	-0,35
18	Zn>	0,49	0,69
19	Zn<	0,66358	0,21401

Таблица 3

i	x_i	k_{1i}	k_{2i}
1	Fe>	-0,69566	-0,88736
2	I>	0,26482	-0,10559
3	K>	0,71461	0,90932
4	Cd>	1,67013	3,48207
5	Ca>	0,29738	0,08592
6	Co>	1,07987	2,29407
7	Co<	0,47953	0,10170
8	Si>	2,29440	1,48289
9	Mg>	0,31636	0,82359
10	Mn<	1,08977	-0,41961
11	Cu>	0,52230	-0,24306
12	As<	0,76925	1,79434
13	Na>	0,28569	0,14712

Окончание табл. 3

i	x_i	k_{1i}	k_{2i}
14	Se>	-0,09783	0,42086
15	Sr<	1,50401	2,26246
16	P>	0,72413	1,92143
17	Cr	-0,02741	0,10767
18	Zn>	-0,03078	1,37205
19	Zn<	0,44892	0,16264

На завершающем этапе работы с помощью обучающей выборки была выполнена проверка полученных результатов на адекватность. Зная заранее наличие у беременной женщины патологии, показания анализов вносились в дискриминантные функции, а результаты сравнивались с действительностью. В итоге было установлено, что у всех изучаемых женщин статистика верна по группирующей ФПН в 91 % случаев, по группирующей ДИВ в 92 % случаев.

Таким образом, метод многомерного дискриминантного анализа представляет собой эффективный инструмент статистического отбора, что было подтверждено применением представленного метода при прогнозировании развития тяжелых осложнений у беременных женщин.

В процессе исследования получены коэффициенты, которые являются константами в дискриминантных уравнениях, оказывающие влияние на степень достоверности статистического анализа.

Преимущество использования многомерного дискриминантного анализа химического состава крови беременных сводится к возможности прогнозирования возникновения патологий в процессе протекания беременности. Использование данного метода открывает большие возможности для медицины, в частности для области акушерства и гинекологии, при прогнозировании возможных осложнений.

Список литературы

1. Дюк В.А. Обработка данных на ПК в примерах. – СПб.: Питер, 1997. – 240 с.
2. Дюк В.А., Эммануэль В.Л. Информационные технологии в медико-биологических исследованиях. – СПб.: Питер, 2003. – 525 с.
3. Колосов В.П., Перельман Ю.М., Ульянычев Н.В. Пути построения прогнозных моделей в пульмонологии // Информатика и системы управления. – 2005. – № 2 (10). – С. 64–71.
4. Петров И.Б. Математическое моделирование в медицине и биологии на основе моделей механики сплошных сред // Труды МФТИ. – 2009. – Т. 1, № 1. – С. 5–16.

5. Вильдеман А.В., Ташкинов А.А. Математические модели индивидуального прогнозирования в медицине // Прикладная математика и механика: тез. докл. науч.-техн. конф. студентов и молодых ученых. – Пермь: Изд-во Перм. гос. техн. ун-та, 2007. – С. 44–46.

6. Вильдеман А.В., Ташкинов А.А., Бронников В.А. Многомерный метод индивидуального прогнозирования индекса моторики // Информационные технологии и вычислительные системы. – 2010. – № 3. – С. 79–85.

7. Гирко В.Л. Многомерный статистический анализ. – Киев: Вища школа, 1988. – 320 с.

Получено 15.04.2017

Семакина Валерия Дмитриевна – студентка, кафедра «Теоретическая механика и биомеханика», факультет прикладной математики и механики, Пермский национальный исследовательский политехнический университет.

Туктамышев Вадим Саитзянович – кандидат физико-математических наук, доцент кафедры «Теоретическая механика и биомеханика», факультет прикладной математики и механики, Пермский национальный исследовательский политехнический университет.