

Технические науки

УДК 618.19

**Яковенко Алена Викторовна**

*кандидат технических наук,*

*доцент кафедры биомедицинской кибернетики*

*Национальный технический университет Украины*

*«Киевский политехнический институт имени Игоря Сикорского»*

**Yakovenko Alena**

*Candidate of Technical Sciences,*

*Associate Professor of the Department of Biomedical Cybernetics*

*National Technical University of Ukraine*

*"Igor Sikorsky Kyiv Polytechnic Institute"*

**Сорока Анастасия Ярославовна**

*студентка магистратуры*

*Национального технического университета Украины*

*«Киевский политехнический институт имени Игоря Сикорского»*

**Soroka Anastasiia**

*Master's Student of National Technical University of Ukraine*

*"Igor Sikorsky Kyiv Polytechnic Institute"*

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПРОГНОЗИРОВАНИЯ  
ДЛЯ ОЦЕНКИ РИСКА РАЗВИТИЯ ЛЕГОЧНЫХ ОСЛОЖНЕНИЙ  
COMPARATIVE ANALYSIS OF FORECASTING METHODS FOR  
RISK ASSESSMENT OF DEVELOPMENT OF PULMONARY  
COMPLICATIONS**

*Аннотация. В данной статье рассматривается необходимость прогнозирования возникновения легочных осложнений у детей с*

*врождёнными пороками сердца после операции на сердце с искусственной вентиляцией легких. Определены основные параметры, которые являются наиболее значимыми при построении математической модели. Построены модели определения возникновения легочных осложнений.*

**Ключевые слова:** *врожденные пороки сердца, искусственная вентиляция легких, лёгочные осложнения, бинарная логистическая регрессия, корреляция, дискриминантный анализ, метод группового учета аргументов.*

**Summary.** *This article discusses the need to predict the occurrence of pulmonary complications in children with congenital heart defects after heart surgery with artificial ventilation of the lungs. The main parameters that are most significant in the construction of a mathematical model are determined. Constructed models for determining the occurrence of pulmonary complications.*

**Key words:** *congenital heart defects, artificial ventilation of the lungs, pulmonary complications, binary logistic regression, correlation, discriminant analysis, the method of group accounting of arguments.*

**Вступление.** Врожденные пороки сердца – это распространенная патология, которая обосновывает высокую смертность детей в первый год их жизни. Среди врожденных пороков развития, врожденные пороки сердца находятся на первом месте [6, с. 505].

Выживаемость с этой патологией зависит, от того, как скоро был диагностирован порок и начато лечение. Единственным способом коррекции этой патологии является оперативное вмешательство.

Широкое распространение получили операции на открытом сердце в условиях искусственного кровообращения, но эти вмешательства являются не безопасными. В последнее время усилилось внимание к изучению

различных осложнений, в том числе и легочных. Знание механизмов развития этих нарушений и возможности своевременного их выявления при помощи современных методов исследования позволяют наиболее рано диагностировать и корригировать развивающиеся патологические состояния.

Тяжесть состояния обусловлена угнетения местного и общего иммунитета, и как следствие, частыми инфекционными заболеваниями легких, объединенных с разными хромосомными нарушениями. Все эти показатели в значительной степени обуславливают таких пациентов к возникновению осложнений в послеоперационном периоде.

Одной из важных проблем, возникающих в раннем послеоперационном периоде, является нарушение функции дыхания и дыхательная недостаточность. Их частота составляет 8 - 10% после операций на сердце в условиях искусственного кровообращения [4, с.207-210].

Целью работы является создание математической модели для прогнозирования наступления легочных осложнений у детей с врождённым пороком сердца в послеоперационном периоде.

**Материалы и методы.** Для достижения поставленной цели исследования проведен анализ результатов хирургического лечения 333 пациентов с врожденными пороками сердца, которым было выполнено хирургическое вмешательство. Все вмешательства выполнены в Государственном учреждении «Национальный институт сердечно-сосудистой хирургии им. Н.Н. Амосова» АМН Украины. База данных отображает состояние пациентов до операции и после операции, и наличие у пациентов легочных осложнений после операции.

Выборка в ходе исследования разделена случайным образом на обучающую выборку и тестовую выборку в соотношении 80%:20%. Обучающая выборка использовалась при построении математических

моделей, а при помощи тестовой проводилась оценка эффективности математических моделей. Средний возраст больных детей находится в диапазоне 18 месяцев: девочки от рождения до 37 месяцев, мальчики – от рождения до 36 месяцев.

Статистическая обработка материала проводилась с использованием пакетов программ IBM SPSS Statistics 23.0, GMDH Shell и Multifactor Dimensionality Reduction 3.0.2.

Была отобрана группа факторов, которые повлияли на развитие легочных осложнений в раннем послеоперационном периоде.

На первом этапе исследовались корреляционные связи факторов, влияющих на развитие легочных осложнений.

Для построения математической модели были использованы метод бинарной логистической регрессии, что позволяет рассчитывать весовые коэффициенты для качественных и количественных показателей, дискриминантный анализ, а также метод группового учета аргументов. Сравнив точность модели, была определена лучшая [1, с. 29-35].

Для анализа характера связи факторов был использован метод MDR - метод снижения факторной размерности. В нем используется алгоритм поиска, который оценивает все возможные комбинации признаков относительно исходного параметра. Влияние каждого фактора и их взаимодействие оценивается величиной энтропии, что выявлена для каждого узла и выражается в процентах, где 100% - ситуация, когда значение переменной однозначно определяет влияние, соответственно 0% - ситуация, когда переменная не влияет. На ребрах отображается значение энтропии между несколькими факторами. Все это позволяет воздействовать на исходную переменную.

**Анализ результатов.** Перед началом построения математической модели был проведен корреляционный анализ, что выявил наиболее значащие факторы. Результаты корреляционного анализа влияния

факторов риска на проявление легочных осложнений у пациентов представлены в табл. 1.

Таблица 1

**Корреляционный анализ влияния факторов риска появления легочных осложнений у пациентов**

Показатель	Легочное осложнение	
	r*	P
Вес, кг	0,289	0,000
Рост, см	0,270	0,000
Площадь поверхности тела	0,292	0,000
Заболевания дыхательной системы	1 – отсутствуют	0,310
	4 – пневмония	-0,283
Продолжительность ИВЛ	1 – 1-6 часов	0,335
	4 - >24 часов	-0,415
ДО мл (до перфузии)	0,258	0,000
C stat, комплайнс статический (до перфузии)	0,224	0,000

Источник: разработка автора

Выявлены факторы, для которых корреляция значима на уровне 0,01. Отрицательное значение коэффициента корреляции говорит о наличии обратной взаимосвязи между факторами риска и наступлением легочного осложнения.

Для построения математических моделей оценки риска возникновения легочных осложнений применены методы бинарной логистической регрессии, дискриминантный анализ и МГУА. Модели, полученные с помощью этих методов, дают возможность с одной стороны оценить структуру наиболее важных факторов риска наличия легочных осложнений среди больных, а с другой - определять вероятность, то есть прогнозировать возникновение легочных осложнений.

Для клинического применения выбрана математическая модель, что показала наибольшие показатели точности классификации.

Модель бинарной логистической регрессии (1) имеет следующий вид:

$$p = \frac{1}{1 + e^{-z}} \quad (1)$$

де  $z = a_m \cdot x_m + a_{m-1} \cdot x_{m-1} + \dots + a_0$ ;  $p$  – вероятность развития легочных осложнений в послеоперационном периоде;  $e$  – экспонента, основа натуральных логарифмов;  $z$  – показатель, что определяет степень влияния суммы прогностических факторов на развитие легочных осложнений;  $a_{1...m}$  – весовые коэффициенты уравнения регрессии;  $x_{1...m}$  – факторы, влияющие на развитие легочных осложнений [2].

Математическая модель прогнозирования вероятности развития легочных осложнений в раннем послеоперационном периоде после операции на обучающей выборке имеет следующий вид (2):

$$p_1 = \frac{1}{1 + e^{-z_1}} \quad (2)$$

где

$$z_1 = -0,054 \cdot x_1 + 0,010 \cdot x_2 + 6,473 \cdot x_3 + 1,704 \cdot x_4 - 0,904 \cdot x_5 + 1,129 \cdot x_6 - 0,409 \cdot x_7 + 0,021 \cdot x_8 - 0,208 \cdot x_9 - 1,769.$$

В модель вошли следующие факторы:

$x_1$  - возраст;

$x_2$  - рост;

$x_3$  - площадь поверхности тела;

$x_4$  - заболевания дыхательной системы, 1- отсутствуют;

$x_5$  - заболевания дыхательной системы, 4- пневмония;

$x_6$  - продолжительность искусственной вентиляции легких, 1- 1-6 часов;

$x_7$  - продолжительность искусственной вентиляции легких, 4 - >24 часов;

$x_8$  – диастолический объем мл (до перфузии);

$x_9$  - C stat, комплайнс статический (до перфузии).

На обучающей выборке было рассчитано чувствительность математической модели – 70.8% и специфичность – 86.4%. Общий процент правильно классифицированных значений составил 85.1%. Для тестовой

выборки чувствительность – 66.6%, специфичность – 83.3%, точность – 81%.

Вторая модель, представленная дискриминантной функцией (3) имеет вид:

$$z = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_0; \quad (3)$$

где  $x_{1...n}$  – показатель, что влияет на появления легочного осложнения;  $b_{1...n}$  – весовые коэффициенты уравнения.

Математическая модель прогнозирования вероятности наступления легочного осложнения, имеет следующий вид:

$$z_2 = -0,207 \cdot x_1 - 0,146 \cdot x_2 + 0,682 \cdot x_3 - 0,310 \cdot x_4 + 0,203 \cdot x_5 + 0,240 \cdot x_6 - \\ - 0,595 \cdot x_7 + 0,017 \cdot x_8 - 0,247 \cdot x_9.$$

$z_2$  - показатель, который определяет степень влияния факторов риска на развитие легочных осложнений;

$x_{1..n}$  - фактор, который влияет на развития легочных осложнений.

В модель вошли следующие факторы:

$x_1$  - возраст;

$x_2$  - рост;

$x_3$  - площадь поверхности тела;

$x_4$  - заболевания дыхательной системы, 1- отсутствуют;

$x_5$  - заболевания дыхательной системы, 4- пневмония;

$x_6$  - продолжительность искусственной вентиляции легких, 1- 1-6 часов;

$x_7$  - продолжительность искусственной вентиляции легких, 4 - >24 часов;

$x_8$  – диастолический объем мл (до перфузии);

$x_9$  - C stat, комплайнс статический (до перфузии).

На обучающей выборке было рассчитано чувствительность математической модели – 57.8% и специфичность – 91.6%. Общий процент правильно классифицированных значений составил 80.4%. Для тестовой

выборки чувствительность – 57.1%, специфичность – 92.8%, точность – 81%.

Третья модель, представлена методами моделирования в программе GMDH Shell, при помощи метода МГУА. Наилучшие результаты моделирования были получены путем применения метода МГУА с пошаговым добавлением аргументов [3].

Общий вид построенной модели:

$$Y1 = 0.118 + X1 \times X2 \times 0.006 + X3 \times X4 \times 0.135 + X5 \times X6 \times 0.108 + X7 \times X8 \times 0.033 + X8 \times X1 \times 0.049 + X8 \times X9 \times 0.028 + X5 \times X10 \times 0.005 + X5 \times X11 \times (-0.277) + X12 \times X11 \times (-0.457)$$

Где:

X1 - заболевания дыхательной системы, 1 – отсутствуют;

X2 - FiO<sub>2</sub>, % (начало операции);

X3 – пол;

X4 - продолжительность ИВЛ, 1 – 1-6 часов;

X5 - заболевания дыхательной системы, 2- ОРВИ;

X6 - C stat, комплайнс статический (до перфузии);

X7 – вес;

X8 - заболевания дыхательной системы, 3 – бронхит;

X9 – продолжительность ИВЛ, 2 7-12 часов;

X10 - прижимание Ао или фибрилляция;

X11 - хромосомная и генетическая патология, 1-нет;

X12 - площадь поверхности тела.

На обучающей выборку было рассчитано чувствительность математической модели – 71.8% и специфичность – 94.2%. Общий процент правильно классифицированных значений составил 79.6%. Для тестовой

выборки чувствительность – 72.7%, специфичность – 92.9%, точность – 83.9%.

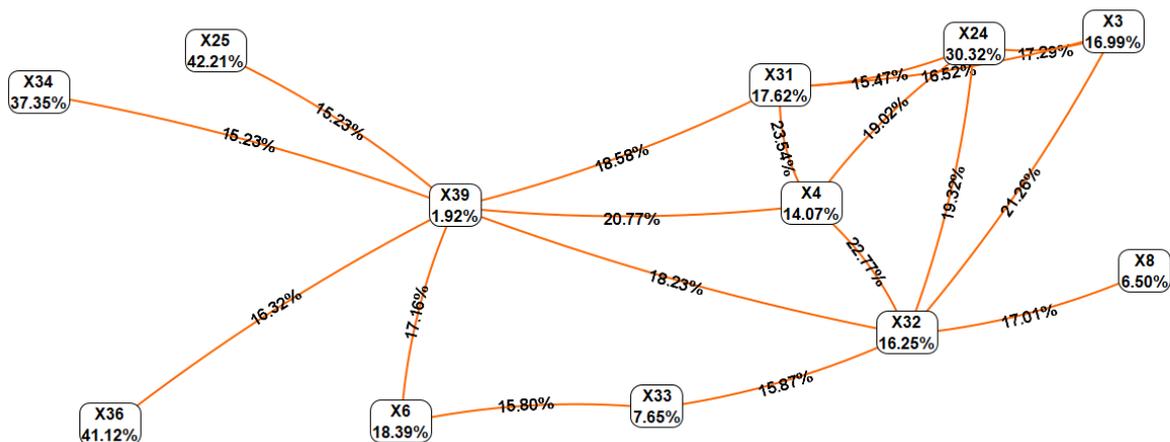
Сравнив результаты точности математической модели для экзаменационной выборки, можно сделать вывод, что модель построена методом МГУА с пошаговым добавлением компонентов имела наибольшую точность, и она равна 83.9%.

Отсутствие в модели некоторых значимых факторов, говорит о том, что отсутствующие факторы поддерживают значимые показатели на должном уровне.

Анализ признаков, непосредственно связанных с факторами риска, приводит к необходимости самостоятельного анализа большого числа вариантов для построения большого числа моделей разного уровня. Этого можно избежать с помощью метода MDR, что помогает выявить иерархию взаимосвязей факторов риска и сами систематические связи.

Граф между факторных взаимодействий в группах больных отражает силу влияния отдельных факторов риска на развитие легочных осложнений с помощью показателя энтропии. Наибольшее влияние имеют факторы с максимальным процентом энтропии, представлены на рис. 1.

Анализ уровня энтропии (рис. 1), показал, что наиболее значимыми факторами риска развития легочных осложнений являются те же параметры что и при построении модели методом МГУА.



**Рис. 1. Граф между факторных взаимодействий, которые влияют на возникновения легочных осложнений**

Обозначения: X3 – пол; X4 – вес; X6 – площадь поверхности тела; X8 – C stat, комплайнс статический (до перфузии); X24 – прижатие A<sub>o</sub> или фибрилляция; X25 – заболевания дыхательной системы, 1 – отсутствуют; X31 – заболевания дыхательной системы, 2- ОРВИ; X32 - заболевания дыхательной системы, 3 – бронхит; X33 – продолжительность ИВЛ, 1 – 1-6 часов; X34 – продолжительность ИВЛ, 2 7-12 часов; X36 - хромосомная и генетическая патология, 1-нет; X39 - FiO<sub>2</sub>, % (начало операции).

Источник: разработка автора

Метод является удобным, во-первых, для определения факторов, которые сильно влияют на исходную признак - наступление легочных осложнений; во-вторых, для определения факторов, влияющих на факторы, связанные с наступлением легочных осложнений, и подвергаются терапевтической коррекции.

Таким образом, метод значительно облегчает разработку системы терапевтических мероприятий, направленных на предотвращение возникновения легочных осложнений среди больных с вмешательством на сердце в анамнезе.

**Вывод.** Оценка риска наступления легочных осложнений дает возможность правильно назначать лечение и предотвращать развитие в послеоперационном периоде.

Построена математическая модель определения вероятности возникновения легочных осложнений у детей с врождёнными пороками

сердца, полученная методом МГУА. Эффективность данной модели – 83.9%. Построен граф между факторных взаимосвязей методом MDR, что дает полную информацию о взаимосвязи факторов, влияющих на возникновения легочных осложнений в послеоперационном периоде.

### **Литература**

1. Койчубенков Б. К., Сорокина М. А., Мхитарян К. Э. Математические методы прогнозирования в медицине. – 2014. – №4.
2. Логистическая регрессия и ROC-анализ — математический аппарат [Электронный ресурс]. – basegroup – Режим доступа: <https://basegroup.ru/community/articles/logistic>.
3. Метод группового учёта аргументов [Электронный ресурс]. – Режим доступа:  
[http://www.machinelearning.ru/wiki/index.php?title=Метод\\_группового\\_учета\\_аргументов](http://www.machinelearning.ru/wiki/index.php?title=Метод_группового_учета_аргументов).
4. Мошковская Л. В., Труба Я. П., Бойко С. Н. и др. Анализ защиты легких у детей с врождёнными пороками сердца на протяжении операции с искусственным кровообращением / Весник сердечно-сосудистой хирургии. – 2014.
5. СПСС (SPSS): искусство обработки информации / Под редакцией А. Бьюль, П. Цёфель. – Москва, Санкт-Петербург, Киев: ТИД «DiaSoft», 2005.
6. Non—lobar atelectasis generates inflammation and structural alveolar injury in the surrounding healthy tissue during mechanical ventilation / J. Retamal, B. Bergamini, A. R. Carvalho [et al.] // Crit. Care. — 2014. — Vol. 18.