

УДК: 004.021

© 2015 г. А.Е. Власенко,
Н.М. Жилина, д-р техн. наук,
Г.И. Чеченин, д-р мед. наук

(Новокузнецкий государственный институт усовершенствования врачей)

АЛГОРИТМ ПРОГНОЗИРОВАНИЯ СОСТОЯНИЯ РЕБЕНКА НА МОМЕНТ РОЖДЕНИЯ

В статье представлен алгоритм прогнозирования состояния младенца на момент родов, который с учетом динамики развития осложнений у женщины позволяет на различных сроках беременности рассчитывать вероятность рождения ребенка с отсутствием одного или более признаков живорождения.

Ключевые слова: здоровье новорожденного, базы данных, патологии беременности, прогнозирование, логистическая регрессия, управленческие решения.

Введение

Для снижения потерь здоровья населения необходимо выявить группы риска [1] и целенаправленно работать с ними, не допуская (отдаляя) развитие неблагоприятного состояния у каждого конкретного индивида. К настоящему времени в фондах информационно-аналитических отделов учреждений системы здравоохранения накоплены значительные объемы информации о здоровье населения [2, 3]. Необходимо более интенсивно проводить обработку баз данных, чтобы получить полезные знания для корректных, обоснованных управленческих решений, анализа и прогноза здоровья населения [4 – 6].

Потери здоровья не измеряются каким-то одним показателем, а выступают как совокупность статистических характеристик, которые можно разделить на три группы: потери жизненного потенциала, потери трудового потенциала и потери активности жизни [1, 7].

Целью исследования является оценивание и прогнозирование состояния здоровья новорожденного на основе анализа многолетних баз данных для поддержки принятия управленческих решений в муниципальном здравоохранении.

Задачи: 1) разработать алгоритм прогнозирования состояния новорожденного на момент родов, учитывающий динамику развития осложнений у женщины в различные сроки беременности; 2) реализовать программное приложение на основе построенного алгоритма для снижения трудоемкости и оптимизации времени подготовки управленческих решений; 3) провести оценку эффективности результатов исследования.

Материалы и методы. Исследование проведено на персонифицированных базах данных кустового медицинского информационно-аналитического центра «Рождаемость» и «Заболееваемость» населения г. Новокузнецка за ряд лет. Обучающая выборка содержит 45 437 случаев, пять контрольных выборок – по 2272 случая в каждой. Функции зависимости выходной величины от входных переменных на различных этапах беременности построены с помощью метода логистической регрессии. Применены итерационный алгоритм оценки коэффициентов Ньютона-Рафсона и метод определения статистической значимости влияния входной переменной на выходную величину с помощью статистики Вальда.

Научная новизна исследования заключается в создании алгоритма прогнозирования состояния новорожденного на момент родов, который отличается учетом динамики развития осложнений у женщины и позволяет на различных сроках беременности рассчитывать вероятность рождения ребенка с отсутствием одного или более признаков живорождения.

Практическая значимость. Разработанный алгоритм может применяться в здравоохранении и в службе родовспоможения непосредственно в работе с отдельным пациентом, для определения вероятности наступления неблагоприятного события и разработки управленческих решений для его предотвращения. Кроме того, алгоритм может применяться к группе пациентов (на уровне отделения, ЛПУ, города) для расчета потребности в медицинской помощи.

Одной из наиболее проблемных характеристик потерь здоровья в Новокузнецке является *мертворождаемость* [7]. Предлагаемый алгоритм разработан для снижения мертворождаемости. Многочисленными экспериментальными и клиническими исследованиями установлено, что внутриутробное развитие плода зависит главным образом от организма матери. Наиболее перспективным путем снижения мертворождаемости и перинатальной смертности является заблаговременное прогнозирование наступления неблагоприятного перинатального исхода.

Обобщенная схема управления течением беременности и формализованное описание параметров

Обобщенная схема управления течением беременности на основе прогнозирующего алгоритма представлена на рис. 1, где $W(t)$ – внешнее неконтролируемое воздействие в момент времени t ; $U(t)$ – управляющее воздействие; $\{X_1(t), \dots, X_n(t)\}$ – вектор характеристик беременной в момент времени t ; $Y(t + \Delta t)$ – состояние новорожденного в момент рождения; $\tilde{Y}(t + \Delta t)$ – прогнозируемое состояние новорожденного; $Y^*(t + \Delta t)$ – заданное состояние новорожденного.

Задача состоит в том, чтобы по наблюдениям $\{X_1(t), \dots, X_n(t)\}$ определить управленческое воздействие $U(t)$, при котором выход системы $Y(t + \Delta t)$ – состояние новорожденного на момент родов – не превышает заданное значение $Y^*(t + \Delta t)$. Состояние новорожденного в данном случае описывается бинарной переменной (1 – отсутствует один или более признаков живорождения; 0 – наличие всех четырех признаков). Характеристиками беременной являются социально-демографические факторы (например, возраст, наличие вредных привычек) и медицинские осложнения беременности и родов.



Рис. 1. Обобщенная схема управления течением беременности на основе прогнозирующего алгоритма.

Все осложнения беременности и родов можно разделить на три группы. К первой группе относятся осложнения, связанные с общим состоянием здоровья женщины и реакцией организма на беременность, данные осложнения выявляются преимущественно в первую половину беременности. Ко второй группе относятся осложнения, связанные с ростом и развитием плода, такие осложнения выявляются во вторую половину беременности. И к третьей группе относят осложнения, связанные непосредственно с родами, они выявляются с момента начала родовой деятельности [7].

В зависимости от динамики развития осложнений беременности можно выделить три дискретных момента времени: t_1 – 10 недель, середина первой половины беременности (интервал прогноза $\Delta t_1 \approx 30$ недель); t_2 – 30 недель, середина второй половины беременности ($\Delta t_2 \approx 10$ недель); t_3 – начало родовой деятельности ($\Delta t_3 \approx 24$ часа).

Прогноз состояния новорожденного осуществляется в определенные дискретные моменты времени по формулам:

$$\tilde{Y}(t_1 + \Delta t_1) = f(X_1(t_1), U_1(t_1), X_{sd}(t_1)); \quad (1)$$

$$\tilde{Y}(t_2 + \Delta t_2) = f(X_2(t_2), U_2(t_2), X_1(t_1), U_1(t_1), X_{sd}(t_2)); \quad (2)$$

$$\tilde{Y}(t_3 + \Delta t_3) = f(X_3(t_3), U_3(t_3), X_2(t_2), U_2(t_2), X_1(t_1), U_1(t_1), X_{sd}(t_3)), \quad (3)$$

где $\tilde{Y}(t_i + \Delta t_i)$ – прогнозные значения состояния новорожденного на момент рождения; t_1, t_2 и t_3 – дискретные моменты времени соответственно 10-я, 30-я недели и начало родовой деятельности; $\Delta t_1, \Delta t_2$ и Δt_3 – интервал прогноза соответственно 30 недель, 10 недель и 24 часа; U_1, U_2 и U_3 – управленческие воздействия в момент времени t_1, t_2 и t_3 ; $X_{sd}(t_i)$ – вектор социально-демографических характеристик беременной; $X_1(t_i), X_2(t_i), X_3(t_i)$ – вектор осложнений, выявляемых соответственно в первую половину беременности, во вторую половину и с момента начала родовой деятельности.

В качестве управленческого воздействия предполагается госпитализация беременной с применением медикаментозного купирования выявленных ослож-

нений. Существует перечень заболеваний, выявление которых в период беременности женщины предполагает ее госпитализацию. Для учета влияния управленческого воздействия на выходную переменную, те заболевания, которые есть в перечне, кодируются тремя значениями: «осложнение отсутствует», «осложнение выявлено, пациентка не госпитализирована», «осложнение выявлено, пациентка госпитализирована». Осложнения, которые отсутствуют в перечне заболеваний, являющихся показанием для госпитализации беременной в стационар, кодируются как: 0 – «осложнение отсутствует»; 1 – «осложнение выявлено».

Прогнозируемая переменная (состояние ребенка) является бинарной, таким образом, решается задача построения модели бинарного выбора. В таких моделях прогнозируется не сама зависимая переменная, а вероятность положительного или отрицательного исхода. Значение выходной переменной определяется в зависимости от полученного значения вероятности:

$$\tilde{Y} = \begin{cases} 1, & \text{если } P(Y) \geq 0.5, \\ 0, & \text{если } P(Y) < 0.5, \end{cases}$$

где \tilde{Y} – прогнозируемое состояние новорожденного; $P(\tilde{Y})$ – вероятность того, что у ребенка на момент родов будет отсутствовать один или более признаков живорождения, лежит в диапазоне от 0 до 1.

Зависимость прогнозируемой вероятности от входных факторов определяется логистической функцией

$$P(\tilde{Y}) = (1 + \exp(-b_0 - b_1x_1 - b_2x_2 - \dots - b_nx_n))^{-1}, \quad (5)$$

где $P(\tilde{Y})$ – вероятность наступления неблагоприятного исхода; b_0 – свободный член уравнения регрессии; b_1, \dots, b_n – коэффициенты уравнения; x_1, \dots, x_n – входные переменные; n – количество входных переменных

Уравнение (5) является уравнением логистической регрессии, с учетом которого формулы (1) – (3) принимают следующий вид:

$$\tilde{Y}(t_2 + \Delta t_2) = f(X_3(t_3), U_3(t_3), X_2(t_2), U_2(t_2), X_1(t_1), U_1(t_1), X_{sd}(t_1))), \quad (3)$$

$$P(\tilde{Y}(t_1 + \Delta t_1)) = (1 + \exp(-b_0(t_1) - B_{sd}(t_1)X_{sd}(t_1) - B_1(t_1)X_1(t_1)))^{-1}, \quad (6)$$

$$P(\tilde{Y}(t_2 + \Delta t_2)) = \quad (7)$$

$$= (1 + \exp(-b_0(t_2) - B_{sd}(t_2)X_{sd}(t_2) - B_2(t_2)X_2(t_2) - B_1(t_1)X_1(t_1)))^{-1},$$

$$P(\tilde{Y}(t_3 + \Delta t_3)) = \quad (8)$$

$$= (1 + \exp(-b_0(t_3) - B_{sd}(t_3)X_{sd}(t_3) - B_3(t_3)X_3(t_3) - B_2(t_2)X_2(t_2) - B_1(t_1)X_1(t_1)))^{-1},$$

где $P(\tilde{Y}(t_i + \Delta t_i))$ – вероятность рождения ребенка с отсутствием одного или более признаков живорождения; t_1, t_2 и t_3 – дискретные моменты времени, соответственно 10-я, 30-я недели и начало родовой деятельности; $\Delta t_1, \Delta t_2$ и Δt_3 – интервал прогноза соответственно 30 недель, 10 недель и 24 часа; $b_0(t_i)$ – свободный член уравнения логистической регрессии; $X_{sd}(t_i)$ – вектор социально-демографических характеристик беременной; $B_{sd}(t_i)$ – вектор коэффициентов при X_{sd} ; $X_1(t_i)$ – вектор осложнений, выявляемых в первую половину беременности; $B_1(t_i)$ – вектор коэффициентов при X_1 ; $X_2(t_i)$ – вектор осложнений, выявляемых во вторую половину беременности; $B_2(t_i)$ – вектор коэффициентов при X_2 ; $X_3(t_i)$ – вектор осложнений,

выявляемых с момента начала родовой деятельности; $B_3(t_i)$ – вектор коэффициентов при X_3 .

Для оценки статистической значимости влияния каждой входной переменной на выходную величину рассчитывается показатель Вальда. Если данный показатель больше критического значения, то влияние переменной на выходную величину статистически значимо. Критическое значение показателя Вальда определяется из распределения χ^2 и зависит от числа степеней свободы и заданного уровня статистической значимости, в данном случае это 3,84 (одна степень свободы и уровень значимости 0,05).

Постановка задачи разработки алгоритма прогнозирования состояния ребенка на момент рождения

Дано: 1) зарегистрированные случаи родов с указанием социально-демографических характеристик матери (X_{sd}), выявленных у нее во время беременности осложнений (X_1 – осложнения первой половины беременности; X_2 – осложнения второй половины; X_3 – осложнения, выявляемые с момента начала родовой деятельности) и состояния ребенка в момент родов ($Y = 1$ – отсутствует один или более признак живорождения; $Y = 0$ – наличие всех четырех признаков);

2) обучающая выборка, содержащая 45 437 случаев;

3) пять контрольных выборок по 2272 случая в каждой;

4) функции зависимости выходной величины от входных переменных на различных этапах беременности;

5) итерационный алгоритм оценки коэффициентов Ньютона-Рафсона;

6) метод определения статистической значимости влияния входной переменной на выходную величину с помощью статистики Вальда;

7) ограничения – число итераций при оценке коэффициентов не менее 100 и значение статистики Вальда, рассчитанной для каждой входной переменной, не менее 3,59;

8) критерий – минимизация среднемодульного отклонения

$$\frac{\sum_{i=1}^N (y - \tilde{y})}{N} \rightarrow \min, \quad (9)$$

где y – фактическое значение выходной переменной; \tilde{y} – прогнозное значение; N – объем выборки.

Требуется построить алгоритм прогнозирования состояния ребенка на момент рождения в зависимости от социально-демографических и медицинских характеристик женщины, с учетом динамики развития осложнений во время беременности:

на обучающей выборке с помощью алгоритма Ньютона-Рафсона оценить коэффициенты уравнения логистической регрессии на каждом этапе беременности с учетом представленных ограничений;

с помощью теста Вальда отобрать переменные, влияние которых на выходную величину статистически значимо;

проверить работоспособность алгоритма (рассчитать среднемодульное отклонение) на обучающей и контрольных выборках.

Результаты решения поставленной задачи

В табл. 1 представлены численные значения коэффициентов уравнений (6) – (8), полученные в результате решения поставленной задачи. Представлены оценки только для тех переменных, влияние которых на выходную величину статистически значимо.

Таблица 1

Характеристики беременной	Этап беременности		
	t_1	t_2	t_3
<i>Социально-демографические характеристики (X_{sd})</i>	<i>Коэффициенты B_{sd}</i>		
Паритет родов (1 – первый ребенок; 2– второй и более)	-0,29	-0,27	-0,35
Возраст (от 14 до 55)	0,15	0,00	0,00
Состоит в браке (1 – да, 0 – нет)	-0,25	-0,24	0,00
Наличие высшего образования (1 – да, 0 – нет)	-0,40	-0,31	-0,20
Курение (1 – да, 0 – нет)	0,70	0,49	0,37
Злоупотребление алкоголем (1 – да, 0 – нет)	1,20	1,17	0,71
Наркомания (1 – да, 0 – нет)	1,00	0,84	0,27
<i>Факторы, выявляемые в 1-й половине беременности (X_1)</i> (1 –осложнение выявлено; 0 – не выявлено)	<i>Коэффициенты B_1</i>		
Послеоперационный рубец матки (без госпитализации)	0,54	0,49	0,59
Послеоперационный рубец матки (с госпитализацией)	0,01	0,02	0,02
Существовавшая ранняя эссенц. гипертензия (без госпитализации)	0,89	0,91	1,20
Существовавшая ранняя эссенц. гипертензия (с госпитализацией)	0,48	0,49	0,58
Миопия (без госпитализации)	0,41	0,26	0,34
Миопия (с госпитализацией)	0,27	0,17	0,00
Предполагаемые наследственные болезни у плода	0,73	0,58	0,44
Кровотечение в ранние сроки беременности	1,19	0,80	0,57
Чрезмерная рвота беременных	-0,67	-0,64	0,00
Инфекционные болезни, выявленные в 1-й половине	0,55	0,55	0,70
<i>Факторы, выявляемые во 2-й половине беременности (X_2)</i> (1 –осложнение выявлено; 0 – осложнение не выявлено)	<i>Коэффициенты B_2</i>		
Многоводие (без госпитализации)	-	0,51	0,41
Многоводие (с госпитализацией)	-	0,01	0,02
Тяжелая преэклампсия (без госпитализации)	-	1,11	0,76
Тяжелая преэклампсия (с госпитализацией)	-	1,20	0,71
Неправильное предлежание плода (без госпитализации)	-	0,54	0,34
Неправильное предлежание плода (с госпитализацией)	-	0,03	0,04
Преждевременная отслойка плаценты (без госпитализации)	-	2,30	1,66
Преждевременная отслойка плаценты (с госпитализацией)	-	1,31	1,01
Признаки внутриутробной гипоксии плода	-	0,91	0,90
Избыточный рост плода	-	0,31	0,54
Гипотензивный синдром	-	-0,77	-0,91
Ложные схватки	-	-0,49	-0,63
Преэклампсия средней тяжести	-	0,55	0,40
Инфекционные болезни, выявленные во 2-й половине	-	0,28	0,37
Пол ребенка мужской (1 – да, 0 – нет)	-	0,18	0,27
<i>Факторы, выявляемые с момента начала родовой деятельности (X_3)</i>	<i>Коэффициенты B_3</i>		
Срок родов (от 24 до 50 недель)	-	-	-0,37
Обвитие пуповины вокруг шеи	-	-	0,64

Характеристики беременной	Этап беременности		
	t ₁	t ₂	t ₃
Выход микония в амниотическую жидкость	-	-	0,84
Признаки стресса плода	-	-	0,53
Разрыв промежности при родоразрешении	-	-	-1,33
Первичная слабость родовой деятельности	-	-	0,48
Вторичная слабость родовой деятельности	-	-	0,69
Запутывание пуповины	-	-	0,70
Неудачная попытка стимуляции родов	-	-	1,63
Неудачная попытка вызвать роды	-	-	1,00
Затрудненные роды вследствие неправильного предлежания	-	-	0,85
Изменение частоты сердечных сокращений плода	-	-	1,51
Биохимические признаки стресса плода	-	-	1,56
Операция кесарева сечения	-	-	-0,22
Константы (B ₀)	-0.01	0.24	13.2

Показатели качества модели в зависимости от величины интервала прогноза представлены на рис. 2. Минимальная ошибка прогноза составляет 13% на обучающей и $13,3 \pm 0,23\%$ – на контрольных выборках, при интервале прогноза 24 часа. С ростом интервала ошибка увеличивается, качество прогноза снижается.

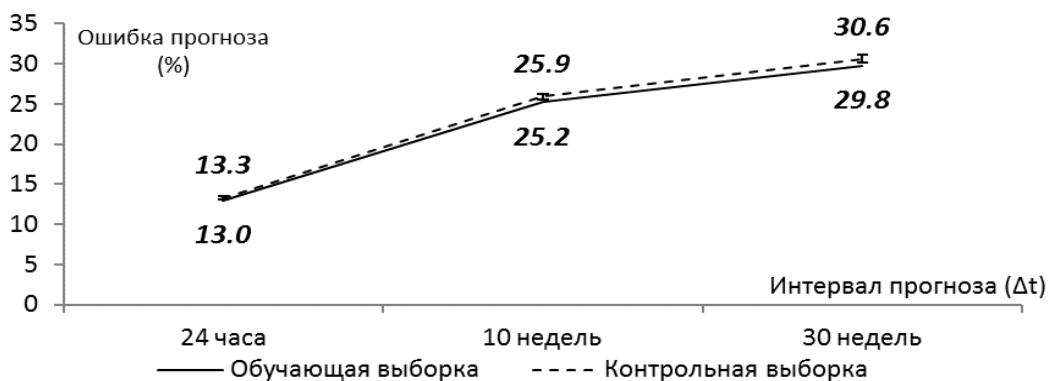


Рис. 2. Показатели качества модели в зависимости от интервала прогнозирования.

Поскольку выработка управленческого воздействия на основе однократного прогноза приводит к значительным отклонениям от заданного значения выходной величины, непосредственное воздействие на объект осуществляется не на всем горизонте прогнозирования $[t + \Delta t]$, а только на его малой начальной части $\Delta > 0$, существенно меньшей Δt .

После реализации управления в момент $t + \Delta$ осуществляется новый прогноз поведения объекта с горизонтом предсказания $\Delta t - \Delta$ и решается задача выбора оптимального управленческого воздействия уже на отрезке $[t + \Delta; t + 2\Delta]$, причем начальным условием для прогнозирующего алгоритма выступают характеристики объекта управления в момент времени $t + \Delta$. Разработанное управленческое решение применяется к объекту на отрезке $[t + \Delta; t + 2\Delta]$, и далее процесс повторяется. Для первой половины беременности величина Δ составляет 1 месяц, для второй половины – 2 недели, начиная с момента начала родовой деятельности параметр Δ принимается равным 1 часу.

Программная реализация разработанного алгоритма

В качестве языка программирования выбран язык C#. Программная реализация алгоритма прогноза состояния новорожденного на момент родов позволяет вычислить вероятность рождения ребенка с отсутствием одного или более критериев живорождения в зависимости от социально-демографических характеристик матери, выявленных медицинских осложнений на различных этапах беременности. *Пример* задания входных параметров и выходной результат даны на рис. 3.

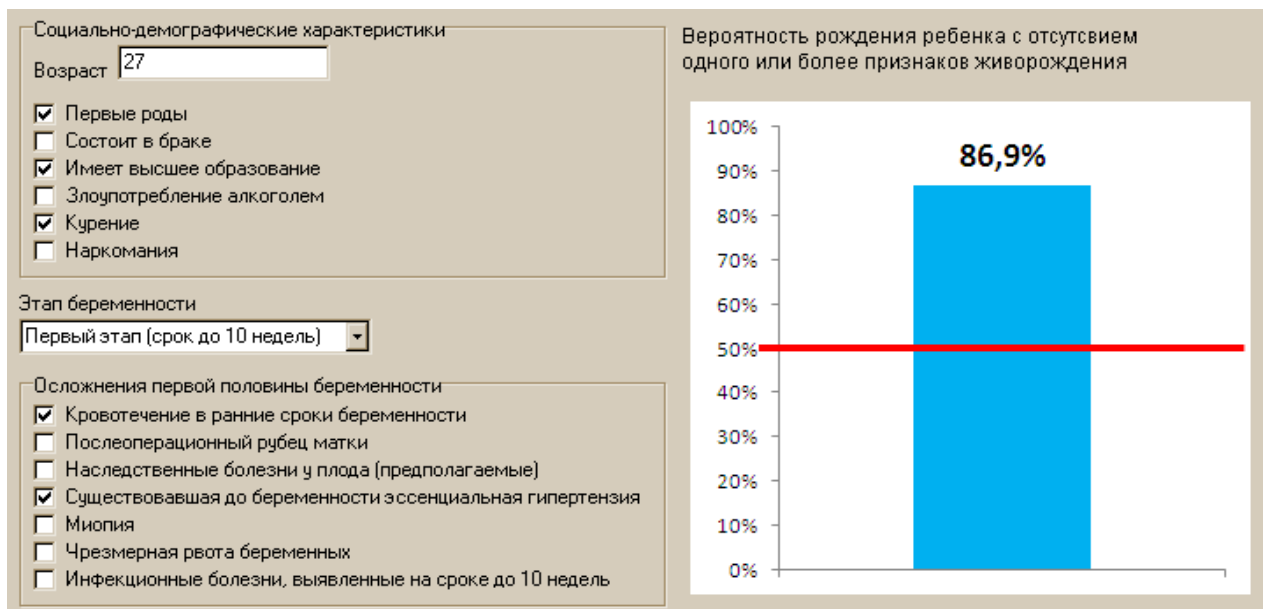


Рис. 3. Фрагмент экранной формы задания входных параметров и вывода результата.

Разработанное программное приложение предназначено для врачей, ведущих прием в женских консультациях и стационарах, для управления состоянием конкретного пациента. Пользователями приложений также могут быть и специалисты управления здравоохранением, главные врачи, заведующие ЛПУ (для выделения групп риска среди населения).

Оценка эффективности результатов исследования

Далее проводится расчет эффективности управленческих решений, принимаемых на основе прогнозирующего алгоритма. Особенность здравоохранения заключается в том, что нередко медицинские мероприятия лечебного и профилактического характера могут быть экономически невыгодны, однако медицинский и социальный эффект требует их проведения. Описание затрат и количественная оценка предотвратимых потерь, рассчитанные на данных о здоровье населения г. Новокузнецка за 2012 г., представлены в табл. 2.

Таблица 2

Затраты	Предотвратимые потери	
	социальные	экономические
Затраты на реанимацию и выхаживание ребенка	Снижение мертворождаемости (38 случаев)	Предотвращение потерь ВВП в результате преждевременной смертности (648624 тыс. руб.)

Заключение

Разработан алгоритм прогнозирования состояния новорожденного на момент родов, позволяющий определять вероятность рождения ребенка с отсутствием одного или более признаков живорождения, с учетом динамики развития осложнений у женщины во время беременности. На первом этапе беременности (интервал прогнозирования 30 недель) ошибка прогноза на контрольных выборках составляет $30,6 \pm 0,48\%$, на втором этапе беременности (интервал прогнозирования 10 недель) – $25,9 \pm 0,41\%$, а при расчете прогноза в момент начала родовой деятельности ошибка составляет $13,3 \pm 0,23\%$.

Разработана программная реализация предлагаемого алгоритма для снижения трудоемкости и оптимизации времени подготовки управленческих решений, позволяющая в диалоговом режиме рассчитывать вероятность развития у беременной неблагоприятных последствий с учетом динамики их развития (для разработки и принятия медицинских и управленческих решений).

Анализом эффективности алгоритма установлено, что применение его на практике позволит обеспечить как *социальную эффективность* (предотвратить 38 случаев мертворождаемости в Новокузнецке в год), так и *экономическую эффективность* (выгода составила 648 624 тыс. рублей в год).

Достоверность полученных результатов подтверждается корректностью методик исследования, достаточным объемом экспериментального материала, положительными результатами участия авторов в работе отечественных и международных конференций, а также документами о внедрении результатов работы практику.

ЛИТЕРАТУРА

1. *Артюхов И.П., Сульдин С.А., Протасова Н.П.* Методические подходы к оценке факторов риска здоровья населения // Сибирское медицинское обозрение. – 2012. – № 6. – С.80-85.
2. *Балыгин М. А.* Ситуация с младенческой смертностью в России // Врач. – 2001. – № 2. – С.40-42.
3. *Жилина Н.М.* Алгоритм интеграции показателей в применении к автоматизированной информационной системе социально-гигиенического мониторинга // Информатика и системы управления. – 2009. – № 4(22). – С.150-152.
4. *Власенко А. Е.* Комплексная оценка потерь здоровья населения Новокузнецка // Современные наукоемкие технологии. – 2013. – № 8 (часть 1). – С. 62-64. URL: www.rae.ru/snt/?section=content&op=show_article&article_id=10001619 (дата обращения: 12.09.2013).
5. *Власенко А.Е., Жилина Н.М., Чеченин Г.И.* Развитие аналитической функции в системе охраны здоровья жителей г. Новокузнецка // Сборник трудов XII Всероссийской конференции «Проблемы информатизации региона» (ПИР-2011). – Красноярск: СФУ, 2011. – С.51-57.
6. *Гареева И.А.* Теоретические и методические основы исследования системы здравоохранения // Власть и управление на Востоке России. – 2009. – № 4. – С. 133-139.
7. *Власенко А.Е., Жилина Н.М., Полукаров А.Н., Чеченин Г.И.* Тенденции рождаемости и состояние здоровья матери и ребенка в 2002 – 2011 гг. в г. Новокузнецке // Мать и дитя в Кузбассе. Спецвыпуск. – Кемерово: ИД «Медицина и просвещение», 2012. – С.4-7.

Статья представлена к публикации членом редколлегии Ю.М. Перельманом.

E-mail:

Власенко Анна Егоровна – Vlasenkoanna@inbox.ru;

Жилина Наталья Михайловна – zhilina.ngiuv@yandex.ru;

Чеченин Геннадий Ионович – postmastergiduv@rambler.ru.