

## ОСОБЕННОСТИ ПРОЕКТИРОВАНИЯ МЕДИЦИНСКОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ, ОСНОВАННОЙ НА ИНТЕЛЛЕКТУАЛЬНОМ АНАЛИЗЕ ДАННЫХ

Г.В. Кнышов<sup>1</sup>, А.В. Руденко<sup>1</sup>, Е.А. Настенко<sup>1,2</sup>, А.В. Яковенко<sup>2</sup>,  
С.О. Сиромеха<sup>1</sup>, С.С. Галич<sup>1</sup>

<sup>1</sup>Национальный институт сердечно-сосудистой хирургии им. Н.М. Амосова  
НАМН Украины

<sup>2</sup>Национальный технический университет Украины «Киевский  
политехнический институт»

Описаны структура и этапы проектирования информационной системы поддержки принятия врачебных решений при ведении больных с ишемической болезнью сердца. В основе информационной системы лежит интеллектуальный анализ данных, представленный математическими моделями прогноза развития острой сердечной недостаточности, что позволяет на дооперационном этапе выявить риск развития осложнения в раннем послеоперационном периоде. Разработанная информационная технология позволяет не только выявлять риск развития осложнения, но и принимать решения с целью коррекции лечебного процесса.

**Ключевые слова:** медицинская информационная система, принятие решений, интеллектуальный анализ данных, коррекция лечебного процесса.

Описано структуру і етапи проектування інформаційної системи підтримки прийняття лікарських рішень при веденні хворих з ішемічною хворобою серця. В основі інформаційної системи лежить інтелектуальний аналіз даних, наданий математичними моделями прогнозу розвитку гострої серцевої недостатності, що дозволяє на передопераційному етапі виявити ризик розвитку ускладнення в ранньому післяопераційному періоді. Розроблена інформаційна технологія дозволяє не лише виявляти ризик розвитку ускладнення, а й приймати рішення з метою корекції лікувального процесу.

**Ключові слова:** медична інформаційна система, прийняття рішень, інтелектуальний аналіз даних, корекція лікувального процесу.

### ВВЕДЕНИЕ

Современные медицинские информационные системы (МИС) обеспечивают сбор, регистрацию, структурирование, документирование данных, обмен, хранение и поиск информации, а также поддержку принятия решений (ППР) по оказанию медицинской помощи пациентам. Однако использование накопленной информации для снижения частоты осложнений и совершенствования лечебного процесса все еще является далекой от окончательного решения задачей.

Кроме того, часто в массивах данных содержатся недостающие для принятия успешных решений скрытые знания, получить которые можно с помощью современных направлений информационных технологий (ИТ) и методов интеллектуального анализа данных (ИАД).

Не менее важным остается формализация слабоструктурированных и

неструктурированных задач. Это позволяет создавать модели прогноза развития осложнений, анализировать структуры факторов, приводящих к осложнениям, выявлять пациентов повышенного риска, а также осуществлять динамическую корректировку лечебного процесса с целью минимизации риска развития и тяжести осложнений и снижения числа повторных обращений.

ИАД-ориентированные МИС требуют использования современных подходов и методов анализа данных. Особую актуальность приобретают современные методы многомерного статистического анализа, с помощью которых можно установить характер и структуру взаимосвязей между признаками изучаемого объекта [1].

**Целью** данной работы было обоснование необходимости и представление этапов проектирования информационной системы, основанной на ИАД на примере прогноза риска развития острой сердечной недостаточности (ОСН), в раннем послеоперационном периоде.

## **МАТЕРИАЛ И МЕТОДЫ ИССЛЕДОВАНИЯ**

Для решения поставленной цели был проведен анализ результатов хирургического лечения 4 809 пациентов с ишемической болезнью сердца (ИБС), которым было выполнено аорто-коронарное шунтирование (АКШ) в 2008–2012 гг. За исследуемый период было выполнено 3 683 операции АКШ на работающем сердце и 1 126 операций с переходом на искусственное кровообращение (ИК) (72 — с экстренным и 1 054 — с плановым переходом на ИК). Средний возраст пациентов составлял  $59 \pm 9$  лет. Соотношение мужчин и женщин составляло 85,2 % и 14,8 % (4 096 и 713 соответственно). У 9,3 % (223) пациентов развилась ОСН в раннем послеоперационном периоде.

Все вмешательства выполнены в Государственном учреждении «Национальный институт сердечно-сосудистой хирургии им. Н.М. Амосова» НАМН Украины.

Анализ данных и выявление скрытых закономерностей проводились с использованием различных математико-статистических и энтропийных методов. Для выявления факторов риска были использованы методы бинарной логистической регрессии (БЛР) [2] и дискриминантного анализа (ДА) [3], а для нахождения взаимосвязей между признаками, представлялось перспективным применение метода снижения факторной размерности — Multifactor Dimensionality Reduction MDR [4].

Далее был проведен сопоставительный анализ факторов, выявленных каждым из методов [5] и для практического использования выбран наиболее эффективный метод.

В процессе проектирования и разработки МИС, определяя основные требования к системе, была построена функциональная модель в среде IBM Rational Rose [8]. Моделирование бизнес-процессов осуществлялось с использованием стандарта IDEF0 в среде BPwin [9]. В соответствии с установленными бизнес-понятиями и бизнес-событиями предметной области осуществлено проектирование БД в среде Microsoft SQL Server.

Для программной реализации системы применялась интегрированная среда разработки MS Visual Studio.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Для достижения успеха в принятии решения и ИАД необходимо иметь четкое представление о цели анализа, собрать релевантные данные, выбрать адекватные методы анализа и проверить предпосылки их применения; выбрать программно-технологические и математические средства, реализующие эти методы; выполнить анализ и принять решение об использовании результатов.

Объектом исследований была МИС для прогноза риска развития осложнений, на примере ОСН и ППР, которая обеспечивает следующие функции:

- возможность осуществления поиска пациента в базе данных (БД);
- ввод и сопровождение истории болезни поступившего пациента;
- прогноз развития осложнений;
- построение моделей динамики факторов, необходимых врачу для принятия решений;
- составление отчета по необходимым для врача критериям;
- сохранение полученных результатов в БД системы.

Определяя основные требования к МИС ППР, было проанализировано ее функциональное назначение и была построена диаграмма вариантов использования (рис. 1).

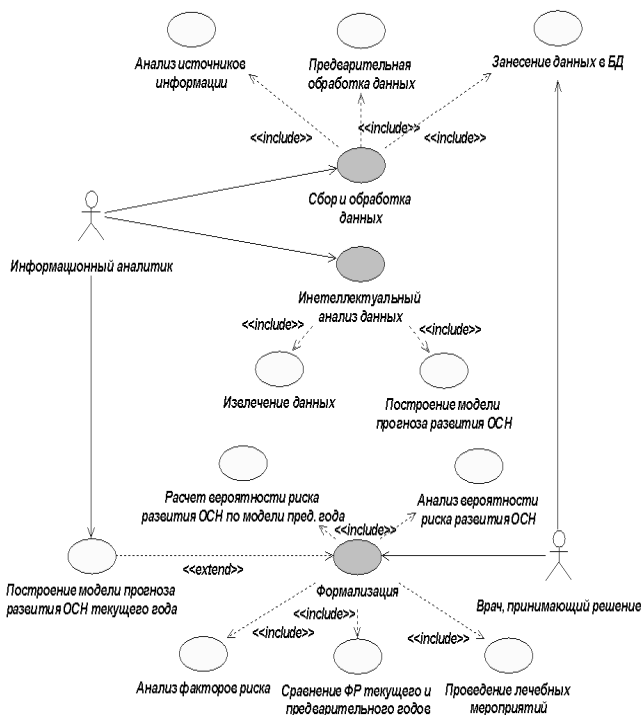


Рис. 1. Диаграмма вариантов использования МИС ППР: БД — база данных; ОСН — острая сердечная недостаточность; ФР — факторы риска

Исходя из функционального назначения, структура МИС ППР включает три блока, которые отвечают за сбор и накопление данных, ИАД и формализацию результатов (рис. 2).

Модуль базы данных (БД) позволяет структурировать и хранить информацию. Блок ППР, построенный на основе ИАД, обеспечивает прогноз риска развития осложнений и предоставляет информацию, необходимую для коррекции лечебного процесса. Формы пользовательских приложений дают возможность вводить, корректировать и анализировать данные по пациенту, рассчитывать риск развития осложнений в раннем послеоперационном периоде и формировать статистические отчеты.

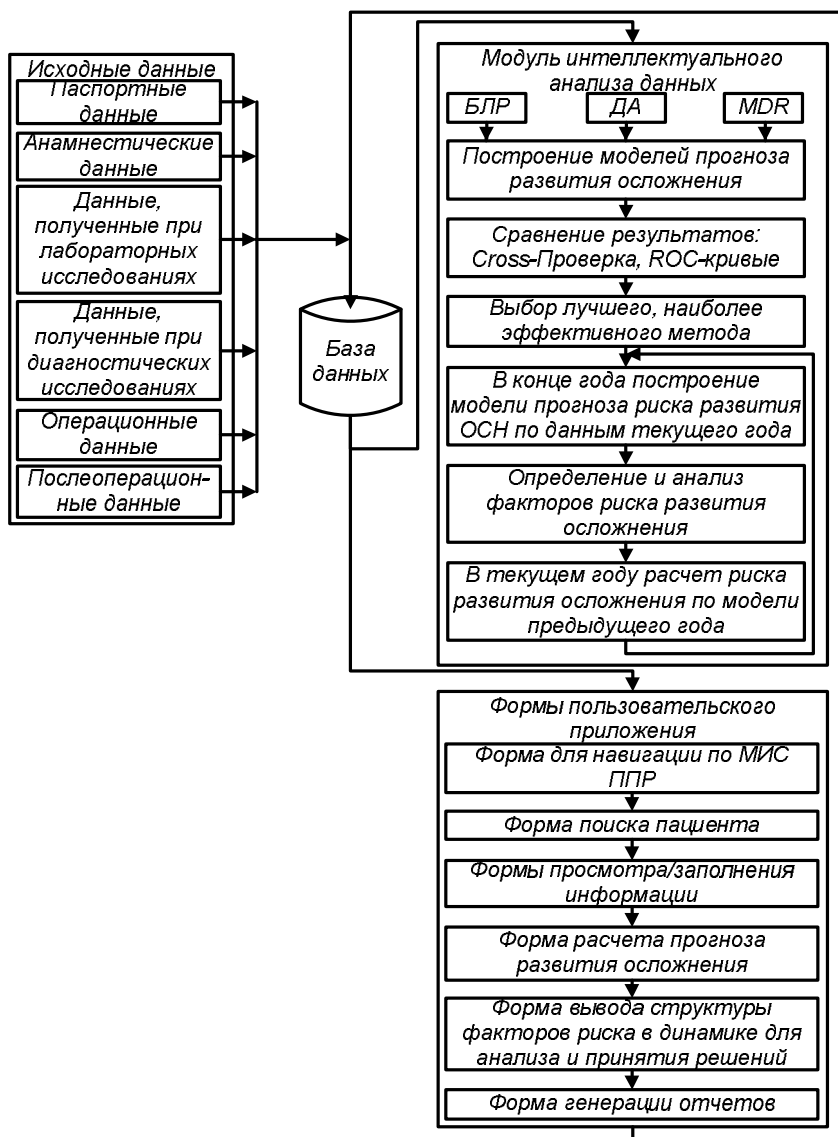


Рис. 2. Общая структура ИС: ОСН — острая сердечная недостаточность; БЛР — бинарная логистическая регрессия; ДА — дискриминантный анализ; MDR — Multifactor Dimensionality Reduction; МИС ППР — медицинская информационная система поддержки принятия решений

При проектировании систем, работающих с большими объемами данных, разработчики сталкиваются с двумя основными проблемами – обработка экспоненциально нарастающих объемов данных, поступающих в реальном времени и сокращение времени их обработки. В отделении хирургии ИБС Национального института сердечно-сосудистой хирургии им. Н.М. Амосова НАМН Украины врачи получают серию графических и видеофайлов, объем которых может занимать десятки гигабайт, а количество исследований постоянно возрастает. Поэтому выбор системы управления базой данных (СУБД) при разработке МИС являлся одной из важнейших задач на этапе проектирования.

При построении информационной системы потребовалось выбрать среду проектирования серверной и клиентской частей. Поскольку МИС должна обеспечивать обмен данными со многими пользователями, централизованная БД с сетевым доступом потребовала выбор клиент-серверной архитектуры, которая совмещает преимущества модели сетевых расчетов, высокие характеристики производства и облегчение администрирования.

В настоящее время активная конкуренция идет не между отдельными СУБД, а между их типами – реляционными и постреляционными [10].

К современным СУБД, поддерживающим модель сетевых вычислений «клиент-сервер», относятся Oracle, Informix, SyBase, Microsoft SQL Server и ряд других. Из всех параметров, по которым их можно анализировать и сравнивать было выделено несколько:

- надежность и защищенность (обеспечение безопасности и целостности данных, система ограничений и ограничителей, автоматическое архивирование, ведение системных журналов);

- быстродействие (наличие SQL-функций, хранимых процедур и оптимизация запросов);

- высокая совместимость с сетевой операционной системой;

- простота администрирования и наличие справочной литературы и документации [11].

Именно эти соображения позволили выбрать Microsoft SQL Server в качестве СУБД при разработке МИС ППР.

При разработке структуры БД МИС были сформулированы следующие принципы:

1. Система разрабатывается в условиях эксплуатации ряда программных средств.

2. Имеются значительные объемы накопленной информации, которая должна быть эффективно использована.

3. БД системы доступна из различных программных средств без использования специальных средств экспорта/импорта.

4. Накапливаемая в БД информация может формироваться в отчеты.

Результатом явилось создание БД, где центральной таблицей является таблица PATIENTS, содержащая паспортные данные больного. Сведения о пребывании больного в стационаре содержатся в таблице ADMISSION, связанной с предыдущей по уникальному ключевому полю. Именно к этой таблице могут быть подключены дополнительные таблицы, содержащие информацию об обследовании и лечении данного больного в отделении.

Массивы данных содержат последовательные наблюдения, которые были накоплены в ходе планового лечебного процесса и постепенно поступают в БД для последующего хранения, обработки и анализа, где хранятся в стандартном формате медицинских данных.

Интеграция статистических методов и СУБД позволяет создавать эффективные системы ППР различного назначения.

Так, вторым блоком, разрабатываемой МИС ППР является ИАД, в основе которого заложен разработанный алгоритм, что позволяет рассчитывать не только прогноз развития осложнения, но и определять годовые показатели значимых факторов риска (ФР).

Идея алгоритма заключается в том, что изначально анализировался ряд статистических методов, которые позволяют рассчитывать прогноз развития осложнения. Статистическая обработка материала проводилась с использованием БЛР с помощью модуля автоматического включения Binary Logistic и дискриминантного анализа с помощью модуля принудительного включения Discriminant Analysis пакета программ IBM SPSS Statistics 20.0 [12]. Нахождение взаимосвязей факторов риска между собой проводилось с помощью пакета Multifactor Dimensionality Reduction 2.0 Beta 8.4.

Прогноз риска развития ОСН осуществлялся на основе показателей, которые оказывают наибольшее влияние на ее развитие. Подробно построенные модели прогноза развития ОСН представлены в работе [13].

Для оценки влияния факторов риска на послеоперационные осложнения у пациентов с ИБС использовались критерий Пирсона и коэффициенты корреляции Спирмена и Кендалла.

В модель были включены признаки, для которых уровень статистической значимости коэффициентов корреляции с ОСН составлял  $p < 0,05$ .

Вторым шагом алгоритма было сравнение построенных моделей, с целью выбора наиболее эффективного метода. Для оценки чувствительности и специфичности построенных моделей была использована кросс-проверка (cross validation test), т.е. метод оценки модели и ее поведения на независимых данных.

Для определения диагностической ценности результирующей прогностической модели использовалась ROC-кривая с последующим определением площади под ней (диагностически значимым является показатель, превышающий 0,70 [6, 7]).

С помощью наиболее эффективного метода строится математическая модель прогноза развития ОСН в раннем послеоперационном периоде, на основе собранных за год данных и отбираются значимые ФР для дальнейшего анализа и воздействия лечебными мероприятиями.

Каждый последующий год по модели предыдущего года рассчитывается риск развития ОСН, а в конце года строится новая математическая модель прогноза развития осложнения и сравниваются ФР. Это позволяет отслеживать данные, влияющие на развитие осложнения, вносить коррективы в лечебный процесс и анализировать, что удалось снивелировать (рис. 3).

Работа алгоритма для пользователя представлена третьим блоком, который содержит ряд форм пользовательского приложения. При запуске

открывается форма для навигации по МИС ППР, с которой можно попасть на форму поиска, форму просмотра/редактирования/внесения новых данных по пациенту и еще две формы: для расчета индивидуальной оценки риска развития осложнений у конкретного пациента и для анализа факторов риска развития ОСН по данным двух годов, на основе которых формируются решения для корректировки и совершенствования лечебного процесса.

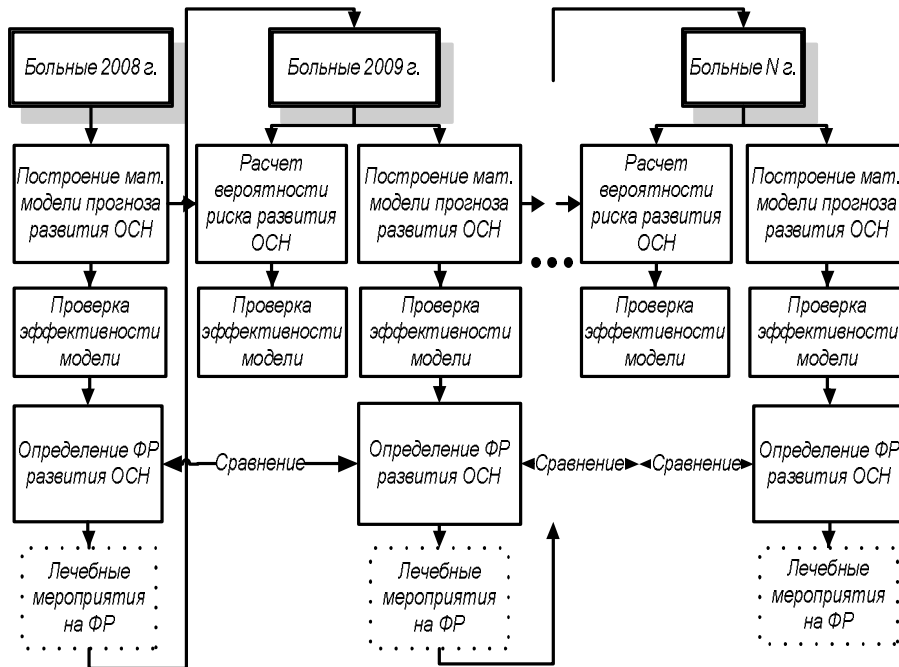


Рис. 3. Информационная модель алгоритма анализа расчета моделей и динамики данных: ОСН — острая сердечная недостаточность; ФР — факторы риска

Для врача представлена структура ФР, которые оказывали влияние на развитие ОСН в раннем послеоперационном периоде, на примере больных 2008 г. Видя ее, он может выделить ряд признаков на которые можно было бы тем или иным способом повлиять на дооперационном этапе, с целью снижения вероятности риска развития осложнения.

В течении последующего года рассчитывается риск развития ОСН и проводятся лечебные меры. Затем по накопленным данным текущего года рассчитывается модель прогноза развития осложнения, отбираются значимые ФР этого периода и сравниваются с ФР предыдущего. Таким образом, можно увидеть, что рядом лечебных воздействий удалось снивелировать некоторые факторы, часть ФР повторились в следующем году и часть новых факторов вошли в модель (рис. 4).

Возможно, повторившиеся ФР, не зависят от самого врача, ведь есть признаки, на которые нет возможности влиять на дооперационном этапе. Но также остается вероятность того, что лечебные воздействия не были эффективными. В этом случае целесообразно пересмотреть тактику ведения больных.

Исходя из этапов алгоритма, весь цикл шагов повторяется, что с каждым последующим годом дает возможность врачу определять и вводить в

постоянную практику ряд дополнительных лечебных мероприятий.

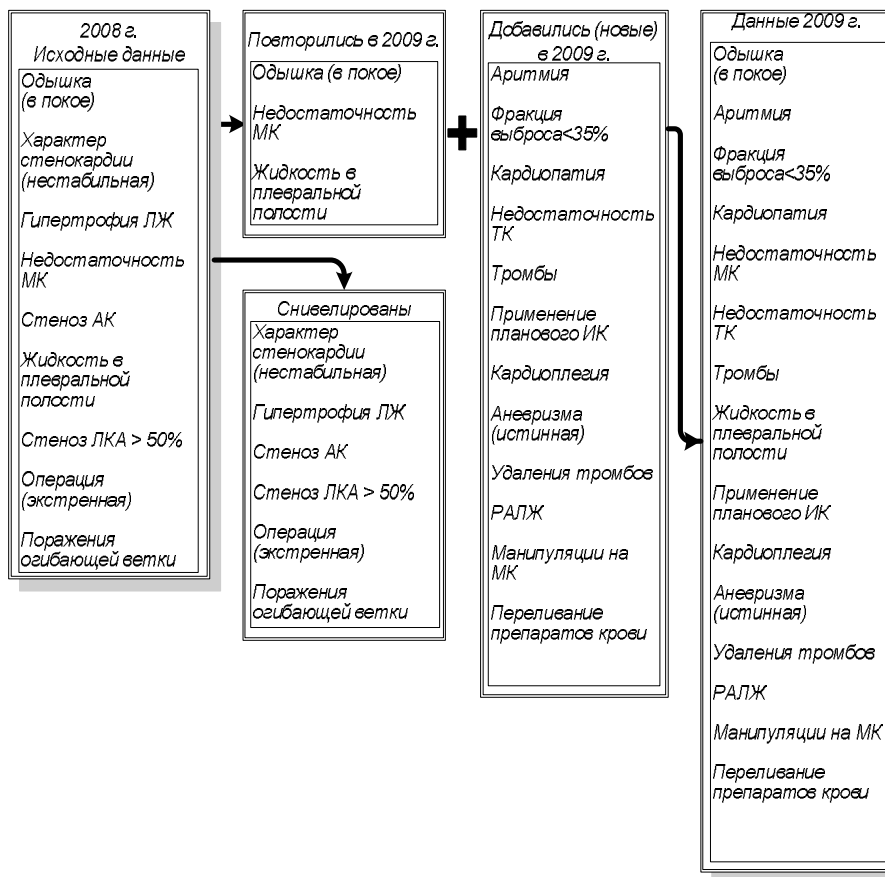


Рис. 4. Пример работы алгоритма для анализа динамики факторов риска:  
 ЛЖ — левый желудочек; МК — митральный клапан; АК — аортальный клапан;  
 ЛКА — левая коронарная артерия; ТК — трикуспидальный клапан;  
 ИК — искусственное кровообращение; РАЛЖ — резекция аневризмы  
 левого желудочка

## Выводы

Спроектированная информационная система полезна для сопровождения лечебного процесса, оценки риска развития осложнений в раннем послеоперационном периоде и для выявления структуры факторов-предикторов влияющих на развитие ОСН.

Разработанная ИТ построения прогностических моделей на основе ИАД позволяет неограниченно расширять спектр прогнозируемых осложнений и дает возможность выявлять пациентов повышенного риска. Кроме того, она является инструментом для врача в принятии решений с целью коррекции лечебных мероприятий и минимизации риска развития осложнений.

Система показала свою эффективность при тестовом использовании на ретроспективных данных, а также в реальных условиях лечебного процесса. Разработанные модели позволяют с высокой надежностью (более 90 %) прогнозировать вероятность развития осложнения в раннем



послеоперационном периоде. В статье приведены результаты для тестовой выборки 2008 г.

1. Ланг Т.А. Как описывать статистику в медицине / руководство для авторов, редакторов и рецензентов / Т.А. Ланг, М. Сесик. — М. : Практическая Медицина, 2011. — 480 с.
2. Григорьев С.Г. Логистическая регрессия. Многомерные методы статистического анализа категориальных данных медицинских исследований : уч. пособие / С.Г. Григорьев, В.И. Юнкеров, Н.Б. Клименко. — СПб, 2001. — С. 10–21.
3. Дискриминантный анализ / Дж.-О. Ким [и др.] // Факторный, дискриминантный и кластерный анализ. — М. : Финансы и статистика, 1989. — С. 78–138.
4. Jakulin, A. Quantifying and Visualizing Attribute Interactions / A. Jakulin, I. Bratko // An Approach Based on Entropy. PKDD. — 2004. — Vol. 3. — P. 229–240.
5. Дюк, В. Data Mining: учебный курс / В. Дюк, А. Самойленко. — СПб. : «Питер». — 2001.
6. Risk stratification and therapeutic decision making in acute coronary syndromes / E.M. Ohman, C.B. Granger, R.A. Harrington, K.L. Lee // JAMA. — 2000. — Vol. 8. — 284 p.
7. Бирман Э.Г. Сравнительный анализ методов прогнозирования / Э.Г. Бирман. — НТИ. Сер. 2. — 1986. — № 1. — С. 11–16.
8. Дерябкин В.П. Проектирование информационных систем по методологии UML : методические указания к учебно-лабораторному практикуму / сост. В.П. Дерябкин, В.В. Козлов. — Самара : Самарск. гос. арх.-строит. ун-т, 2008. — 42 с.
9. Маклаков С.В. BPwin ERwin CASE-средства разработки информационных систем / С.В. Маклаков. — М. : Диалог-МИФИ, 2001. — 304 с.
10. Кузнецов С.Д. Основы баз данных / С.Д. Кузнецов — 2-е изд. — М. : Интернет-университет информационных технологий; БИНОМ. Лаборатория знаний, 2007. — 484 с.
11. Гарсиа-Молина Г. Системы баз данных. Полный курс / Г. Гарсиа-Молина, Дж. Ульман, Дж. Уидом. — М. : Вильямс, 2003. — 1088 с.
12. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей : пер. с нем. — СПб. : ООО «ДиаСофтЮП», 2005. — 608 с.
13. Выявление структуры факторов риска развития острой сердечной недостаточности в раннем послеоперационном периоде / А.В. Яковенко, А.В. Руденко, Е.А. Настенко, Н.Л. Руденко, В.А. Павлов // Восточно-европейский журнал передовых технологий. — 2013. — № 3/10 (63). — с. 4–8.

Получено 19.06.2014