

# ЭКСПЕРТНАЯ СИСТЕМА ОЦЕНКИ СОСТОЯНИЯ БОЛЬНЫХ

В. Г. Рудалев, М. А. Артемов, Р. В. Жуков

*Воронежский государственный университет*

Рассматриваются функциональные возможности экспертной системы оценки и прогнозирования состояния больных. Обосновывается структура системы, основными элементами которой являются подсистемы расчета диагностических показателей и аналитической обработки информации. Описываются алгоритмы установления статистической связи между количественными и качественными факторами заболевания.

## 1. ВВЕДЕНИЕ

При лечении различного рода тяжелых заболеваний возникает проблема оперативной обработки и правильной интерпретации данных терапевтического обследования, число которых может быть велико и измеряться десятками. Такого рода ситуации характерны в интенсивной терапии критических состояний, токсикологии, хирургической гастроэнтерологии, при черепно-мозговой и сочетательной травме и т.д. Значительную помощь здесь могут оказать системы поддержки принятия диагностических решений, позволяющие на основе данных обследования выдать оценку состояния больного и прогноз дальнейшего развития заболевания. Подобные системы должны носить, безусловно, специализированный характер, ориентированный на конкретные потребности лечебного заведения в терапии указанного круга заболеваний.

В данной статье описывается экспертная система оценки физического состояния больных и расчета схемы искусственного питания, внедренная в Воронежской ОКБ\*. Разработанный программный продукт функционально состоит из следующих подсистем:

- Подсистема оперативной обработки информации.
- Подсистема расчета основных диагностических показателей.
- Подсистема получения врачебных экспертных заключений и оценок.
- Подсистема редактирования базы знаний.
- Подсистема аналитической обработки информации.
- Подсистема безопасности.

© Рудалев В. Г., Артемов М. А., Жуков Р. В., 2007

\* Авторы выражают благодарность Э. А. Машарину и С. В. Ермоленко, любезно предоставившим детальное описание предметной области.

Большая часть функциональных возможностей реализована средствами клиент-серверной СУБД InterBase 7.0, пользовательский интерфейс выполнен в виде клиентского Windows-приложения, написанного в среде Delphi.

## 2. СТРУКТУРА ЭКСПЕРТНОЙ СИСТЕМЫ

Рассмотрим структуру системы более подробно.

Задачей подсистемы оперативной обработки информации является ведение базы данных пациентов и результатов их обследований. Здесь реализованы основные особенности предметной области и логика работы с данными. База данных включает таблицы, предназначенные для регистрации данных о пациенте и результатов обследований; справочные таблицы, содержащие необходимую базу знаний по рассматриваемой предметной области, на основе которой происходит выдача медицинских заключений.

Клиентская часть системы освобождена от выработки заключений. Ее задачи — сбор данных о пациенте (первичных данных, данных биохимических анализов и других дополнительных обследований), пересылка их серверу и получение от него экспертных заключений.

Особенностью разработанной системы является гибкость настройки — возможность изменения и пополнения базы знаний без изменения кода программы. Расчетные формулы, по которым производится расчет диагностических показателей (например, показателей энергетического обмена, прогностических индексов гипотрофии и др.) включают данные медицинских анализов и измерений и учитывают наличие у пациента факторов стресса. Значимость факторов стресса (травм, сопутствующих заболеваний, перенесенных операций, ИВЛ, ослож-

нений), температуры тела, степени ожогов и факторов активности измеряется в баллах. Баллы проставляются специалистами-экспертами и используются в формулах расчета диагностических показателей.

В целом расчетные формулы имеют полуэмпирический характер и могут видоизменяться в соответствии с достижениями медицинской науки. Потому все формулы хранятся изолированно от кода программы — в таблице Formules БД. При получении оценки состояния больного происходит поиск и выборка необходимых формул в таблице Formules.

Выбранные формулы затем используются для формирования запроса к таблицам (средствами динамического SQL и UDF) для расчета диагностических показателей, которые затем передаются на вход подсистемы выдачи экспертных заключений и оценок. Хранимая процедура GetConclusion (решатель) просматривает справочные таблицы, реализующие продукционную модель знаний, и выдает искомые заключения по вероятному течению заболевания, степени выраженности нарушений трофологического статуса, показателей белкового обмена и др. (всего 9 заключений). Для большей информативности происходит лексический анализ заключения на наличие слов, требующих повышенного внимания (тяжелый, средний, третья или четвертая степень) и текст выводится с соответствующей подсветкой. Заключения формируются в виде отчета и направляются на печать.

Если вычисления по формуле не увенчались успехом, например, из-за не заполнения ряда полей при вводе данных обследования, то делается попытка подобрать другую формулу с меньшим объемом входной информации, а при неудаче выдается сообщение о недостаточности исходных данных.

Таблицы, позволяющие сделать диагностическое заключение по числовому показателю, построены по единому принципу. Поясним его на примере заключения о трофологическом статусе пациента по его индексу массы тела. Индекс массы тела ИМТ рассчитывается по известной формуле, выбранной из таблицы Formules. Таблица TrophStatus\_IMT имеет столбцы Range (нижняя граница индекса массы тела, допустимая для данного заключения) и Conclusion (заключение или следующий показатель). Значения столбца Range упорядочены по воз-

растанию. В качестве заключения принимается значение Conclusion, возвращаемое запросом

```
SELECT Conclusion FROM TrophStatus_IMT
WHERE Range = (SELECT Max (Range)
FROM TrophStatus_IMT WHERE :IMT >= Range)
```

Например, значение ИМТ в пределах 15—17 приводит к заключению «Гипотрофия второй степени» (тяжелое истощение).

Получение экспертных баллов из таблиц с факторами активности, стресса или аналогичных происходит несколько проще. Так как их первичные ключи являются внешними ключами в таблице обследований, то баллы выбираются из текущей записи упомянутых таблиц.

Для изменения базы знаний (формул, справочных данных и таблиц с диагностическими заключениями) используется подсистема редактирования с графическим интерфейсом. Очевидно, для полноценного сопровождения системы эксперт должен иметь возможность работать с ней в режиме «обратной связи», постоянно совершенствуя алгоритмы получения заключений по данным врачебных наблюдений. Первоочередными здесь являются вопросы о взаимосвязи диагностических показателей (как первичных, так и расчетных), факторов стресса и других сопутствующих факторов с возможным ходом течения заболевания; взаимосвязи диагностических показателей и факторов друг с другом. Ответы на данные вопросы позволяют выявить наиболее диагностически значимые показатели и факторы и исключить из рассмотрения второстепенные факторы, и факторы, полностью или частично дублирующие друг друга. В конечном итоге можно сократить размерность задачи и упростить процесс выдачи заключений. Взаимосвязь факторов носит статистический характер и может быть установлена на основе анализа совокупности данных.

Для установления статистической зависимости факторов используется подсистема аналитической обработки данных. Такого рода подсистемы оперируют данными, сгруппированными по заданным признакам [1,2]. Задача усложняется тем, что многие признаки (факторы) не имеют количественного выражения, а описываются словами «Имеется», «Отсутствует», «Ухудшилось», «Выше нормы» и т.п., и по сути являются качественными. Для анализа тесноты статистической связи качественных и количественных факторов данные представля-

ются в категоризованном виде, т.е. в виде частот наблюдений, попавших в некоторые категории или классы [3]. Данная статистическая связь возникает между двумя и более переменными, выраженными в категоризованной форме. Категоризованная переменная может либо представлять собой подходящую классификацию числовой переменной по диапазонам значений, либо быть качественной и не выражаться в терминах какой-либо исходной числовой переменной.

Вследствие того, что каждый из факторов разбивается на две и более категории, то для определения тесноты связи между двумя факторами используются таблицы  $r \times c$ , где  $r$  — количество категорий первого (входного) фактора, а  $c$  — количество категорий второго (выходного) фактора [3]:

|          |          |     |          |          |
|----------|----------|-----|----------|----------|
| $n_{11}$ | $n_{12}$ | ... | $n_{1c}$ | $n_{1*}$ |
| $n_{21}$ | $n_{22}$ | ... | $n_{2c}$ | $n_{2*}$ |
| ...      | ...      | ... | ...      | ...      |
| $n_{r1}$ | $n_{r2}$ | ... | $n_{rc}$ | $n_{r*}$ |
| $n_{*1}$ | $n_{*2}$ | ... | $n_{*c}$ | $n$      |

Здесь  $n_{ij}$  — количество пациентов с сочетанием двух факторов,  $n_{i*}$  — количество пациентов с первым фактором,  $n_{*j}$  — количество пациентов со вторым факторов,  $n$  — общее количество пациентов.

В случае независимости переменных в таблице частота на пересечении  $i$ -й строки и  $j$ -го столбца равна  $n_{i*}n_{*j} / n$ . Следовательно, отклонение от независимости в этой клетке таблицы измеряется величиной

$$D_{ij} = n_{ij} - \frac{n_{i*}n_{*j}}{n}.$$

Определим коэффициент связи в терминах называемой квадратичной сопряженности  $\sum_{i,j} D_{ij}^2 / (n_{i*}n_{*j})$  и запишем величину  $X^2$  в виде [3]:

$$X^2 = \sum_{i,j} \frac{D_{ij}^2}{n_{i*}n_{*j} / n} \equiv n \left\{ \sum_{i,j} \frac{n_{ij}^2}{n_{i*}n_{*j}} - 1 \right\}.$$

Если выполнена гипотеза о независимости, то величина  $X^2$  имеет асимптотически  $\chi^2$ -распределение с числом степеней свободы  $(r - 1)(c - 1)$ .

Теснота связи качественных факторов определяется с помощью коэффициента взаимной сопряженности Пирсона

$$P = \left( \frac{X^2}{n + X^2} \right)^{\frac{1}{2}},$$

который изменяется в промежутке от 0 до 1. Чем ближе величина  $P$  к 1, тем связь между факторами теснее.

Анализ связи исхода заболевания как выходного качественного фактора и качественных факторов результатов обследований по прогностическим индексам трофологического статуса, соматического пула белка, висцерального пула белка и факторов стресса подтвердил их определяющее влияние на ход течения заболевания. Разработанный программный модуль поддержки принятия решений «не привязан» к каким-либо конкретным факторам, является универсальным и строит куб решений для произвольных факторов, передаваемых ему в качестве параметров. Это позволяет с большей точностью прогнозировать дальнейшее течение болезни и исключать показатели, оказавшиеся диагностически несостоятельными.

При разработке подсистемы безопасности с целью обеспечения конфиденциальности данных, защиты от несанкционированного доступа и надежного хранения базы знаний использована защита на основе ролей СУБД. Роли позволяют установить различные права доступа в зависимости от функций, выполняемых пользователем. Пользователь с административными правами имеет неограниченные возможности: полный доступ ко всем таблицам, возможность создавать новые таблицы и регистрировать в базе данных новых пользователей. Пользователь с правами «Регистратор» может добавлять новых пациентов. Роль «Доктор», имеет доступ к данным своих пациентов, имеет право модифицировать их данные, производить расчеты, но не имеет права редактировать анкетные данные пациента. Доступ к справочным таблицам и таблицам с формулами возможен в режиме «только для чтения». Пользователь, принадлежащий роли «Эксперт», обладает более высокой квалификацией и ему можно доверить, редактирование базы знаний, замену баллов, формул и т.п. Роль «Аналитик» дает полный доступ ко всей экспертной информации, но запрещает редактировать оперативные данные обследований.

### **3. ЗАКЛЮЧЕНИЕ**

Подходы и принципы, использованные при разработке системы, являются достаточно универсальными и могут применяться для автоматизации различных задач, связанных с поддержкой принятия решений.

#### **ЛИТЕРАТУРА**

1. *Коннолли Т., Бегг К.* Базы данных. Проектирование, реализация и сопровождение. Теория и

практика. 3-е издание / Т. Коннолли. — М.: Вильямс, 2003. — 1440 с.

2. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.* Методы и модели анализа данных: OLAP и Data Mining / А. А. Барсегян — СПб.: БХВ-Петербург, 2004. — 336 с.

3. *Айвазян С.А.* Прикладная статистика. Исследование зависимостей / С. А. Айвазян. — М.: Финансы и статистика, 1985. — 550 с.

*Статья принята к опубликованию  
25 декабря 2006 г.*