

УДК 004.891

## СОЗДАНИЕ БАЗЫ ЗНАНИЙ ДЛЯ МЕДИЦИНСКОЙ ЭКСПЕРТНОЙ СИСТЕМЫ

<sup>1</sup>Бурцева А.Л., <sup>1,2</sup>Берестнева Е.В., <sup>2</sup>Степаненко Н.П.

<sup>1</sup>ФГБОУ ВПО НИ «Томский политехнический университет», Томск, e-mail: berestneva\_l@mail.ru;

<sup>2</sup>ФГБУ «Сибирский федеральный научно-клинический центр Федерального медико-биологического агентства России», Томск, e-mail: stepanenko\_nina62@bk.ru

Работа посвящена вопросам применения методов математической статистики для создания базы знаний медицинской экспертной системы (на примере научных исследований проблем ожирения детей). Актуальность данного исследования обусловлена тем, что несвоевременное и неадекватное лечение ожирения приводит к изменениям во всех внутренних органах и является причиной многих заболеваний органов и систем, инвалидизирующих людей в возрасте социальной активности. Обозначены основные проблемы создания баз знаний и экспертных систем в медицине. Представлены результаты статистической обработки и анализа клинических данных, полученных в Томском научно-исследовательском институте курортологии и физиотерапии. Для оценки эффективности лечения больных с ожирением в различные периоды года использовались непараметрические статистические критерии. Полученные результаты будут использованы при формировании базы знаний медицинской экспертной системы.

**Ключевые слова:** медицинская экспертная система, база знаний, ожирение

## BUILDING THE KNOWLEDGE BASE FOR MEDICAL EXPERT SYSTEM

<sup>1</sup>Burtseva A.L., <sup>1</sup>Berestneva E.V., <sup>2</sup>Stepanenko N.P.

<sup>1</sup>Tomsk Polytechnic University, Tomsk, e-mail: berestneva\_l@mail.ru;

<sup>2</sup>Tomsk SRI of health resorts and physiotherapy of the Federal Medical and Biological Agency, Tomsk, e-mail: stepanenko\_nina62@bk.ru

The work is dedicated to the application of mathematical statistics to create a knowledge base of medical expert system (for example, research problems of childhood obesity). The relevance of this study is due to the fact that late and inadequate treatment of obesity causes changes in all internal organs and is the cause of many diseases of organs and systems, disabling people aged social activity. Outlined the main problems of creating knowledge bases and expert systems in medicine. The results of the statistical processing and analysis of the clinical data obtained in the Tomsk Research Institute of balneology and physiotherapy. To evaluate the effectiveness of treatment of patients with obesity in different periods, the non-parametric statistical tests were used. The results will be used in the formation of the knowledge base of medical expert system.

**Keywords:** medical expert system, knowledge base, obesity

С развитием информационной деятельности человека и быстрым ростом объемов информации становится необходимым создание и разработка информационных технологий и эффективных интеллектуальных продуктов, способных справиться с обработкой больших объемов информации. Экспертная система – это интеллектуальная компьютерная система, построенная на фундаменте накопленных знаний и опыта специалистов-экспертов, о некоторой специализированной предметной области, способная принимать решения, которые не уступают по достоверности и качеству решениям эксперта-профессионала [6].

Одна из самых популярных областей применения экспертных систем – медицина. Это обусловлено тем, что диагностика большинства заболеваний должна проводиться оперативно, зачастую сроки варьируются до нескольких суток. Кроме того, медицинская экспертная система, которая содржит в себе знания и логику лучших

специалистов может обеспечить пациенту объяснение причины его недомогания, она способна предложить одну или несколько эффективных методик лечения.

По имеющимся данным Всемирной организации здравоохранения, к началу XXI века примерно 30% жителей планеты имели избыточную массу тела [9]. Почти у 30–60% взрослых, которые страдали ожирением в детстве, оно протекает тяжелее, а кроме того, продолжает прогрессировать и ведет к развитию различных серьезных осложнений, таких неинфекционных болезней, как диабет, сердечно-сосудистые заболевания, атеросклероз и ишемическая болезнь сердца. Поэтому особую актуальность в этом сложном вопросе представляет рост ожирения у детей. Несвоевременное и неадекватное лечение ожирения приводит к изменениям во всех внутренних органах и является причиной многих заболеваний органов и систем, инвалидизирующих людей в возрасте социальной активности [7, 9].

Исследование различных показателей обмена веществ у здоровых детей и детей с ожирением, а также разработка базы знаний и экспертной системы могут служить основой для определения наиболее эффективных методов реабилитации пациентов детского возраста с ожирением и повысить эффективность терапии в неблагоприятные временные промежутки годового цикла.

### Материалы и методы исследования

На базе детского отделения Томского НИИ курортологии и физиотерапии было проведено клиническое исследование детей с диагнозом конституционально-экзогенное ожирение [1, 7]. Дети обследовались в различные промежутки года [1]: 1 группа – август – ноябрь; 2 группа – декабрь – апрель; 3 группа – май – июль.

Определяли: объем талии и бедер, массу тела, индекс массы тела, диастолическое артериальное давление, индекс инсулинорезистентности, уровень холестерина и глюкозы.

**Целью** данной работы являлась оценка эффективности лечения больных с ожирением в различные периоды года с целью формирования базы знаний для медицинской экспертной системы.

### Экспертные системы и базы знаний

Основное достоинство экспертных систем (ЭС) заключается в возможности накопления знаний. Под знанием в данном случае принято понимать формализованную информацию, на которую ссылаются или применяют в процессе логического вывода [6].

При создании экспертной системы используются данные и знания, которые накапливаются в базе знаний, а также специальные механизмы вывода решений и новых знаний на основе имеющихся. Все

эти знания формируются различными экспертами, инженерами по знаниям и представлены в виде определенных правил для решения тех или иных задач [6]. Обобщенная структура экспертной системы представлена на рис. 1.

Процесс функционирования системы заключается в том, что пользователь, обратившийся к системе за информацией, посылает запрос системе с помощью пользовательского интерфейса. Решатель с помощью базы знаний генерирует информацию и выдает ответ, рекомендацию. При этом он выдает ход рассуждений с помощью подсистемы объяснений.

### База знаний экспертной системы

База знаний (БЗ) – это одна из главных частей ЭС. Ее отличительная особенность заключается в том, что эта часть системы может меняться и пополняться (рис. 2) [6].

База знаний содержит в себе различные факты, сведения, статистические данные о некоторой предметной области, а также правила или набор инструкций, благодаря которому можно получать новые факты на основе полученных знаний. База знаний создается экспертами, исследователями и консультантами какой-либо предметной области. Ценность всей экспертной системы на 90% определяется качеством созданной базы знаний. Для создания базы знаний существуют различные способы [2–6].

Принципиальное отличие знаний от данных заключается в их активности и структурности. Внесение новых фактов в базу знаний или нахождение новых связей может привести к изменениям в принятии решений. Проблема представления знаний состоит в том, что существует несоответствие между такими понятиями, как сведения о зависимостях в проблемной области, методами решения проблемы и возможностями представления информации в ЭВМ.

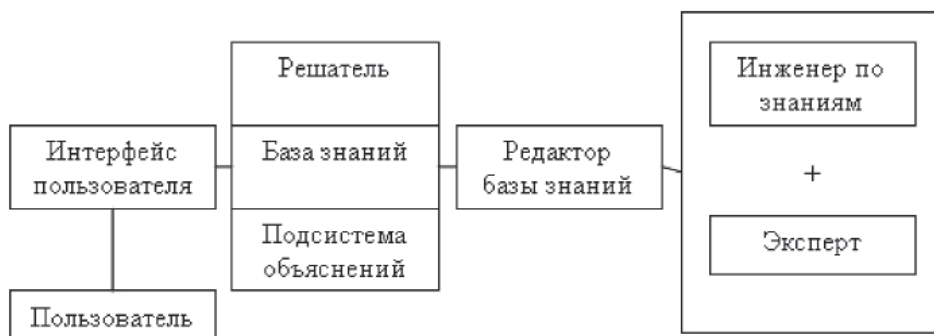


Рис. 1. Структура экспертной системы



Рис. 2. Структура базы знаний

Уровень представления знаний отличается более сложной структурой, особенностями в интерпретируемости, наличием классифицируемых связей от уровня представления данных. Кроме этого, для уровня представления знаний характерно наличие специальных процедур обобщения, а также пополнение имеющихся в системе знаний.

Следует помнить, что экспертные системы в основном разрабатываются не для универсальных задач, а для качественного решения задач в специализированной области. Главными критериями качества полученных новых знаний являются внутренняя согласованность, полнота информации, системность, историзм, однотипность и объективность. Информация должна быть четкой и непротиворечивой [6].

В результате наблюдений экспертов формируются эмпирические знания. Они могут быть согласованными или несогласованными. Таким образом, в процессе извлечения знаний аналитикам необходимо, прежде всего, установить внутреннюю согласованность эмпирических знаний, которая характеризуется понятием модальности, противоречивости, неполноты. **Модальность** знания – это возможность существования знания в различных категориях. Необходимо четко различать такие модальности, как: эксперт знает, что...; *эксперт считает, что...*; *эксперт хочет, чтобы...* Под **противоречивостью** понимают естественное свойство эмпирических знаний, которое не всегда может быть устранено. Напротив, противоречивость может служить отправной точкой в рассуждениях экспертов. **Неполнота** знаний связана с невозможностью исчерпывающего описания какой-либо предметной области. **Системность** знаний базируется на определении места новых знаний в многоуровневой иерархической организации. Инженер по знаниям должен рассмотреть и изучить проблемную

ми [2–6]. Будем использовать следующие термины [6]: извлечение, получение, формирование, приобретение знаний и обучение БЗ. Определим сущность указанных терминов. Под извлечением знаний будем понимать процесс приобретения материализованных знаний из текстологических источников информации с помощью некоторой совокупности методов и процедур, позволяющих переходить от знаний в текстовой форме к их аналогам для ввода в базу знаний СИИ. Получение знаний – это процесс приобретения вербализуемых и невербализуемых знаний эксперта, основанный на использовании непосредственно им самим или инженером по знаниям приемов, процедур, методов и инструментальных средств.

Формирование знаний – это процесс автоматического приобретения (порождения) системой искусственного интеллекта или инструментальным средством нового и полезного знания из исходной и текущей информации, которое в явном виде не формируют эксперты, в целях освоения новых процедур решения прикладных задач на основе использования различных моделей машинного обучения.

В нашем случае для формирования базы знаний были использованы клинические данные [1, 7] и методы математической статистики [8].

### Результаты исследования и их обсуждение

Сравнение показателей у больных ожирением детей до и после лечения (оценка достоверности сдвига) проводилось с использованием непараметрических критериев (критерий знаков и критерий Уилкоксона) в пакете Statistica 8. Полученные результаты представлены в табл. 1 и 2

Таблица 1

Результаты сравнения физиологических показателей у детей с ожирением до и после лечения

	Критерий знаков		Критерий Уилкоксона	
	Уровень значимости (p)	Достоверность «сдвига»	Уровень значимости (p)	Достоверность «сдвига»
АД диастолическое	0,00039	Достоверно	0,0015	Достоверно
Глюкоза	0,052	Не достоверно	0,08	Не достоверно
Холестерин	0,02	Достоверно	0,025	Достоверно
НОМА	0,081	Не достоверно	0,09	Не достоверно
ИМТ	0,0001	Достоверно	0,0001	Достоверно
Масса тела	0,0001	Достоверно	0,0001	Достоверно
Объем бедер	0,0001	Достоверно	0,0001	Достоверно
Объем талии	0,0001	Достоверно	0,0001	Достоверно

область с точки зрения закономерностей не только целого, но и его частей. **Объективность** знаний определить практически невозможно, потому что процесс познания субъективен. Объективность некоторых закономерностей часто связывают с широтой области их применимости. Границы этой области можно установить экспериментальным или теоретическим путем, но далеко не всегда. В качестве косвенных свидетельств объективности иногда допускают совпадение представлений различных экспертов и подтверждение выдвинутых гипотез известными факта-

Как видно из табл. 1, лечение оказало наибольшее влияние на изменение показателей давления, холестерина, массы тела, индекса массы тела, объема талии и объема бедер для всей группы обследованных детей. Такие же результаты были получены для детей из первой группы. Для детей из второй сезонной группы лечение оказало наибольшее влияние на изменение показателей давления, массы тела и индекса массы

Таблица 2

Результаты сравнения групп по критерию Манна – Уитни

Переменные	Критерий Манна – Уитни			Различие	
	Группа 1	Группа 2	Группа 3	p	Достоверность «сдвига»
ДАД	-3,44	-3,06	-2,12	0,81	Не достоверно
Глюкоза	-0,11	-0,05	0,09	0,29	Не достоверно
Холестерин	-0,50	-0,32	-0,21	0,6	Не достоверно
НОМА	-0,40	0,39	0,21	0,57	Не достоверно
ИМТ	-1,19	-1,05	-0,88	0,03	Достоверно
Масса	-2,87	-2,56	-2,48	0,25	Не достоверно
ОБ	-2,55	-2,38	-1,8	0,25	Не достоверно
ОТ	-3,11	-2,21	-2,17	0,33	Не достоверно

тела. Таким образом, можно предполагать, что лечение не оказывает существенного влияния на данные показатели детей с ожирением, поскольку выявлено всего три достоверных различия. Для детей из третьей сезонной группы лечение оказало наибольшее влияние на изменение показателей давления, массы тела, индекса массы тела, объема талии и объема бедер. Таким образом, для детей данной группы лечение оказалось более эффективным, чем для детей второй группы (выявлено пять достоверных различий) [1].

Наилучшие показатели сдвигов наблюдаются у детей первой сезонной группы. Для сравнения эффективности лечения в разных группах был использован критерий Манна – Уитни [8].

Полученные результаты представлены в табл. 2.

В результате анализа данных из табл. 2 можно увидеть, что U-критерий выявил один достоверный показатель. Различие в полученных результатах можно объяснить малым объемом выборки.

Критерий Манна – Уитни выявляет достоверное различие оценок по такому показателю, как индекс массы тела. Следовательно, по результатам экспертных оценок можно сделать предположение, что дети из первой группы лечения имеют лучшую динамику изменения индекса массы тела, чем дети из группы 3.

### Заключение

Интеллектуальные и экспертные системы активно внедряются в различные сферы здравоохранения и медицины. Их появление считается большим научным достижением, потому что они позволяют быстро и эффективно решать задачи, связанные с использованием баз знаний, полученных как непосредственно от экспертов, так и на основе компьютерной обработки и анализа клинических данных. В нашем случае

получены новые знания по эффективности лечения больных с ожирением в различные периоды года с использованием различных схем лечения. Полученные результаты будут использованы при формировании базы знаний медицинской экспертной системы.

*Исследование выполнено при финансовой поддержке РФФИ в рамках проекта № 15-07-08922.*

### Список литературы

1. Агафонов А.А., Бурцева А.Л. Статистический анализ медико-биологических показателей в группах детей с избыточным весом и ожирением // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов II Международной конференции. – Томск, 2015. – С. 842–844.
2. Берестнева О.Г., Пеккер Я.С. Выявление скрытых закономерностей в сложных системах // Известия Томского политехнического университета. – 2009. – Т. 315. – № 5. – С. 138–143.
3. Берестнева О.Г., Шаропин К.А., Старикова А.В., Кабанова Л.И. Технология формирования баз знаний в медицинских информационных системах // Известия ЮФУ. Технические науки. – 2010. – № 8 (109). – С. 32–37.
4. Берестнева О.Г., Муратова Е.А. Построение логических моделей с использованием деревьев решений // Известия Томского политехнического университета. – 2004. – Т. 307. – № 2. – С. 154–160.
5. Берестнева О.Г., Осадчая И.А., Немеров Е.В. Методы исследования структуры медицинских данных // Вестник науки Сибири. – 2012. – № 1 (2). – С. 333–338.
6. Гаврилова Т.А., Хорошевский, В.Ф. Базы знаний интеллектуальных систем. – СПб.: Изд-во «Питер», 2001. – 276 с.
7. Кондратьева Е.И. Санаторно-курортная реабилитация детей с ожирением / Е.И. Кондратьева, Н.П. Степаненко, Г.А. Суханова, О.Б. Светлик и др. // Педиатрия. – 2010. – № 5. – С. 68–72.
8. Математические методы в психологии и педагогике / О.Г. Берестнева, Е.А. Муратова, И.Л. Шелехов, О.С. Жаркова, А.М. Уразаев. – Томск: ТГУ, 2012. – 276 с.
9. Ротов А.В., Берестнева О.Г., Медведев М.А., Гаврилов М.А. Динамика личностных и физиологических показателей при психокоррекции избыточного веса // Сибирский психологический журнал. – 1997. – № 5. – С. 134–136.
10. Семерякова Е.Г., Берестнева О.Г., Макарова Л.С. Математические методы в задачах медицинской диагностики // Современные проблемы науки и образования. – 2012. – № 6.