

На правах рукописи.

КОТОВ Юрий Борисович

МЕТОДЫ ФОРМАЛИЗАЦИИ  
ПРОФЕССИОНАЛЬНОГО ЗНАНИЯ ВРАЧА  
В ЗАДАЧАХ МЕДИЦИНСКОЙ ДИАГНОСТИКИ.

Специальность 05.13.18 –  
Математическое моделирование, численные методы  
и комплексы программ

А в т о р е ф е р а т  
Диссертации на соискание ученой степени  
доктора физико-математических наук

Москва – 2002

Работа выполнена в ордена Ленина Институте прикладной математики им. М.В. Келдыша Российской академии наук.

Научный консультант: доктор физико-математических наук,  
профессор Малинецкий Г.Г.

Официальные оппоненты:

доктор физико-математических наук, академик РАН, Энеев Т.М.,  
доктор физико-математических наук, профессор Чернавский Д.С.,  
доктор физико-математических наук Смолянинов В.В.

Ведущая организация: Международный Институт Теории Прогноза  
Землетрясений и Математической Геофизики РАН.

Защита диссертации состоится « \_\_\_\_ » \_\_\_\_\_ 2002 г. в \_\_\_\_\_ час.  
на заседании диссертационного совета Д002.024.02 при Институте прикладной  
математики им. М.В. Келдыша РАН по адресу: Москва, Миусская пл., д. 4.

С диссертацией можно ознакомиться в библиотеке Института прикладной ма-  
тематики им. М.В. Келдыша РАН по адресу: 125047, Москва, Миусская пл., 4.

Автореферат разослан « \_\_\_\_ » \_\_\_\_\_ 2002 г.

Ученый секретарь диссертационного совета

К.ф.-м.н.

/Г.В. Устюгова/

**Актуальность.** Интерес к процессу принятия человеком решений в затруднительной ситуации существовал уже на заре кибернетики в 60-е годы. Очевидный парадокс состоит в том, что опытный специалист принимает правильное решение (выбирает правильное действие) в условиях видимого дефицита информации или при явной нехватке времени для принятия во внимание всех факторов ситуации. Следует разобраться, каким образом и какие именно особенности ситуации он успеваеt учесть в процессе выполнения профессиональной работы. В этом состоит один из аспектов проблемы исследования профессионального знания специалиста.

Медицинская диагностика представляет собой область интересных и своеобразных задач профессиональной выработки решений в сложных ситуациях или ситуациях с неполной информацией. Особенность работы врача состоит в том, что объект (больной) чрезвычайно сложен, а решение должно быть принято обязательно. Значительная часть информации о больном имеет невербальный характер. Формализация и структуризация хотя бы части используемой врачом информации может быть полезна для самого врача (часть вопросов упрощается и может быть решена формально, это освобождает сознание для решения более сложных профессиональных проблем). Кроме того, облегчается передача его опыта новому поколению специалистов.

Структуризация информации может заметно сократить ее объем. Еще в 1960-е годы И.М. Гельфанд и М.Л. Цетлин заметили, что в реально возникающих физических и технических задачах оптимизации многомерных функций существует локальная организация данных. Большая часть переменных (почти все) оказываются несущественными, т.е. минимум оценочной функции по ним достигается за сравнительно небольшое число шагов. Несколько (1-3) переменных оказываются существенными, т.е. минимизация по ним требует заметно большего числа существенно более крупных (нелокальных) шагов и дает заметно меньшие значения оценочной функции. Они связали это явление с особой структурой данных («хорошо организованная задача»).

Вскоре были обнаружены аналоги такой структуры данных при анализе построения произвольных движений млекопитающих. В современной теории нелинейных динамических систем существует метод «русел и джокеров», использующий близкие идеи о структуре переменных системы. Аналоги такого же представления можно обнаружить в классической гештальт-психологии, которая трактуеt восприятие некоторой ситуации каждым из ее участников в виде единого образа, организованного вокруг своего «центра интереса». Возможно, что в ходе эволюции механизмы работы нервной системы человека восприняли этот принцип организации информации, и следствием этого стал известный феномен «семь плюс или минус два», обнаруженный при исследованиях объема внимания.

Во всяком случае, важной частью структуризации и формализации медицинских данных будем считать поиск небольшого количества существенных переменных в каждой проблемной ситуации принятия специалистом решения в

его профессиональной области. В отличие от случая функции, заданной численно или формулой, поиск таких переменных в клинической задаче требует специфических методов.

Исторически определились два типа возможных носителей полученного знания: текст и программа для компьютера. Выбор одного из них диктовал в значительной мере характер и способ использования результатов. Если мы выбираем текст в качестве носителя нового знания, то нам естественно стремиться к максимально подробной вербализации знания, а методы исследования должны способствовать наилучшему словесному контакту с изучаемым специалистом. Одним из сильнейших достижений на этом пути является методика структуризации знания, использующая диагностические игры.

Методы получения готового знания от специалиста обычно ориентированы на передачу знания компьютерной программе в виде экспертного высказывания. При этом врач, выступающий в качестве эксперта, должен дать определенный ответ на все возможные варианты вопросов о больном. В практической работе врача такой случай крайне редок. Не интересующие его сочетания признаков и варианты заключений могут быть даже осмысленными, но поскольку они не относятся к его области интересов, это может привести к ответу случайному или даже невнятному.

Описанные два крайних подхода не охватывают всю проблему структуризации медицинского знания. С одной стороны, диагностические игры позволяют обнаружить состав и организационную структуру тех данных, которые имеют название, определены, введены в исследовательский оборот. Они позволяют также уточнять определения уже известных разновидностей данных. С другой стороны, логические конструкции, используемые в экспертных системах, пригодны для оперирования только полностью формализованными элементами.

Остаются, как минимум, два пробела. Данные, используемые врачом на интуитивном уровне, нужно обнаружить и описать. В рамках диагностических игр такая задача требует несоразмерных усилий и изощенной квалификации исследовательской группы. Кроме того, диагностическая информация об одном и том же процессе может существовать одновременно в нескольких формах, использующих разные наборы переменных, отражающих различные стороны процесса постановки диагноза. Разные правила принятия одного и того же решения могут использовать различные наборы сведений. Тем самым, формализация первичных наблюдений оказывается тесно связанной с построением диагностических правил, использующих каждый раз наиболее подходящий для данной ситуации набор сведений.

Можно кратко выразить цель формализации знания врача в данной работе: ***«вместе с врачом выработать новое формализованное знание, которым врач может пользоваться как еще одним источником информации в своей профессиональной деятельности»***. При этом следует использовать диагностические игры для анализа его подхода к проблеме, структуризации информации и изучения методов рассуждения. Также естественно решать вместе с ним

его частные проблемы, порожденные количественным характером некоторых типов данных и сложными логическими конструкциями, возникающими при анализе достаточно редких и разнообразных клинических случаев. В ходе всей совместной работы, разумеется, необходимо, чтобы врач имел возможность контролировать используемые данные, их преобразования и ход рассуждений, сохраняя свои самостоятельность и независимость, как исследователь. Это приводит к потребности иметь достаточно лаконичный язык манипулирования элементами рассуждений и заключений врача, позволяющий легко проследить связь логических конструкций с первоначальной информацией.

Работа возникла в русле исследований, проводившихся школой академика И.М. Гельфанда в 1970-90-е годы. Она использует наработки этой школы по диагностическим играм, анализу и формализации точных постановок медицинских диагностических исследований. В то же время, в ней использованы методы, возникшие в ходе решения специфических задач, общей стороной которых явилось большее внимание к деталям работы врача с его материалом (медицинскими сведениями, заключениями, промежуточными выводами и прогнозами). Ряд соображений был навеян работами других школ (М.А. Айзерман, М.М. Бонгард, Э.М. Браверман, В.Н. Вапник, Ю.И. Журавлев, О.И. Ларичев, И.Б. Мучник, В.С. Переверзев-Орлов, и др.).

**Цель исследования.** Разработка методов фиксации и вербализации опыта квалифицированного специалиста (опытного врача) для решения медицинских задач на естественном для врача языке.

#### **Задачи исследования.**

1. Найти методы уточнения постановки медицинской исследовательской задачи с учетом реально существующих ограничений и контингента больных.
2. Разработать методы выявления структуры профессионального знания специалиста, приемов оперирования данными, промежуточными результатами, выработки промежуточных целей исследования.
3. Разработать методы использования частных моделей ситуации и локальных прогнозов в качестве средства уточнения задачи и выявления проблем, возникающих перед специалистом в ходе исследования.
4. Найти методы выявления общей части (ядра) мнений группы экспертов при решении конкретной задачи. Рассмотреть возможные механизмы договоренности при формировании общей точки зрения.
5. Разработать методы компьютерной поддержки формализации профессионального знания.

**Научная новизна.** На ряде практических медицинских исследований разработаны, опробованы и доведены до практического применения диалоговые и человеко-машинные методы формализации знания специалиста.

- Предложена методика обнаружения проблем врача на довербальном уровне.
- Выработана методика построения моделей динамики процессов в связи с решениями врача в процессе наблюдения и лечения больного.
- Разработан язык логических симптомов для представления разнородных данных (количественного и категориального типа) и методы выявления их взаимосвязи.
- Разработан язык описания сценария развития заболевания, проведения лечебных мероприятий и накопления сведений и опыта врачом в ходе лечебной работы. Предложены методы прогнозирования развития ситуации на основе выявления критических событий.
- Предложен метод согласования экспертных выводов в новой для экспертов профессиональной задаче на основе параллельного моделирования экспертных решений и выявления наиболее затруднительных ситуаций для экспертов.

**Практическая значимость.** Все методы были разработаны в ходе решения практических медицинских задач и позволили достичь необходимого лечебного результата. На основе выявленных сведений были созданы рациональные методы ведения больных после хирургической коррекции коарктации аорты; детей, больных хроническим гломерулонефритом; предложена рациональная тактика подготовки к родам женщин с рубцом на матке. Были также предложены методы ранней диагностики катастрофического ухудшения состояния плода при беременности, отягощенной сахарным диабетом; критерии ранней оценки состояния новорожденных для нескольких вариантов патологии беременности. Разработан метод выделения наиболее тяжелой категории среди пострадавших от аварии на ЧАЭС с нарушениями иммунитета по результатам скринингового обследования.

**Апробация.** Материалы доложены и обсуждены на ряде отечественных и международных съездов, симпозиумов и конференций различных уровней.

Третья Всесоюзная научная конференция (12-14 июня 1984, Ереван); Всероссийский съезд кардиологов, Свердловск 22-24 мая 1985.

15 Всесоюзный съезд акушеров-гинекологов, Махачкала 28-30 ноября 1989; «Сахарный диабет и беременность», симпозиум, СПб, 28-29 ноября 1991 г.; «Ультразвуковая диагностика в акушерстве, гинекологии и педиатрии», I съезд Российской ассоциации врачей УЗ диагностики в перинатологии и гинекологии, Суздаль, 6-8 окт. 1992 г.; 1-й Северо-Кавказский съезд акушеров-гинекологов", 7-8 сент. 1994 г, Ростов; всероссийская научно-практическая конференция 15-17 ноября 1994 г., г. Чебоксары; 10-th Congress of European Association of Gynaecologists and Obstetricians, Monte-Carlo, March 29-Apr. 1, 1995; 27-th International Congress on Pathophysiology of Pregnancy; XV FIGO World Congress of Gynecology and Obstetrics, Copenhagen, Denmark, 3-8 Aug. 1997.

I съезд российской ассоциации специалистов перинатальной медицины. Суздаль, 1995; III Всероссийский съезд эндокринологов, 4-7 июня 1996г., Москва; Всероссийская конференция "Основы перинатальной медицины: плацента в этиологии, патогенезе и диагностике внутриутробно формирующихся заболеваний", Новосибирск 1996; II съезд Российской ассоциации врачей акушеров и гинекологов, Москва 1997; 13-th Congress of European Association of gynaecologists and Obstetricians (EAGO), Jerusalem, Israel, may 10-14, 1998; Первый Российский диабетологический конгресс, Москва 1998 г.; Международный симпозиум, Москва:1998; Всероссийский пленум Ассоциации акушеров и гинекологов 19-21 апреля 2000 г. Москва.

V Российский съезд специалистов по лабораторной диагностике, Москва, 24-26 мая 1995г.

Годичные конференции МИФИ, 1998-2000 гг.; V Всероссийская конференция "Нейрокомпьютеры и их применение", Москва, 17-19 февраля 1999.

V Российский национальный конгресс "Человек и лекарство", Москва 21-25 апреля 1998г; VI российский национальный конгресс "Человек и лекарство" 19-23 апреля 1999г, Москва; V Поволжская научно - практическая конференция Саратов, 1999.

3-я Московская международная конференция по исследованию операций (ORM2001), ВЦ РАН, М. 2001.

**Публикации.** По теме опубликовано 112 работ, в том числе 51 статья в центральной печати и рецензируемых изданиях.

**Структура диссертации.** Работа состоит из введения, 13 глав и заключения. Диссертация изложена на 399 страницах, включает 69 рисунков и 72 таблицы. В приложении приведены 2 вопросника. Библиография насчитывает 282 наименования.

**Личный вклад автора.** Включенные в диссертацию основные результаты получены лично автором. В разработке медицинских аспектов результатов, включенных в главы 6-12, принимали участие врачи, сотрудники различных организаций, которым автор выражает благодарность за плодотворное сотрудничество. В постановках геофизических задач из гл.13 и обсуждении результатов принимали участие сотрудники кафедры общей физики МИФИ, которым автор также чрезвычайно благодарен. Разработка вопросников, приведенных в приложениях, осуществлена неформальным коллективом с участием сотрудников Института физико- химической биологии им. А.Н. Белозерского при МГУ. Автор выражает благодарность д.м.н., профессору Селиваненко В.Т., д.м.н., профессору Федоровой М.В., д.м.н., профессору Ефимову В.С., д.ф.-м.н., профессору Горбачеву Л.П., д.м.н. Гришину В.Л., д.м.н. Петрухину В.А., к.м.н. Саляевой М.В., к.м.н. Гурьевой В.М., к.м.н. Шалаеву О.Н., к.м.н. Троицкой М.В., к.б.н. Яковлевой Н.И., к.б.н. Розкину М.Я., к.ф.-м.н. Алексеевскому А.В., к.ф.-м.н. Лукашенко С.Ю., к.ф.-м.н. Семеновой Т.А.

Автор благодарен Российскому фонду фундаментальных исследований за поддержку работ, гранты № 99-01-01091, 01-01-00628, 01-06-80435, 02-01-00844.

Автор благодарен также д.ф.-м.н., профессору Малинецкому Г.Г. и к.ф.-м.н. Шифрину М.А. за содержательное обсуждение работы.

## Содержание работы.

**Введение** к диссертации содержит подробный обзор истории вопроса. Прослеживается постепенное изменение постановок диагностических задач и изменение методов анализа доступных данных. Показано возникновение и углубление интереса исследователей к роли врача в решении проблем диагностики. Сформулированы основные еще не решенные вопросы формализации диагностического процесса. Поставлены задачи исследования. Описана структура диссертации.

**В главе 1** обсуждаются особенности медицинских задач и используемых в них данных. Обращено внимание на анализ числовых данных, их интерпретацию в терминах медицинской задачи.

Удалось выделить типовые подзадачи формализационного исследования, составляющие его «методическую базу». Введены определения используемых терминов: класс объектов, процесс, итог процесса, потенциальное знание, мнение участника работы, согласование мнений, рассеянные образцы и другие.

Рассмотрено использование диагностических игр, обсуждена высокая трудоемкость метода. Предложен метод рационализации технологии – предварительная подготовка рабочего документа, соответствующего запланированной цели игры, на основе накопленных предварительных баз данных.

Массивы численных данных, особенно на ранних этапах работы, малы, поэтому статистические оценки и сравнения классов можно делать только с помощью непараметрических критериев статистики. Краткое описание непараметрических критериев также помещено в главе 1.

По мере накопления материала появляется возможность анализировать эмпирические распределения переменных. Показано, что асимметричные распределения, встречающиеся в медицинских задачах, хорошо аппроксимируются распределением Вейбулла.

$$F(x) = 1 - e^{-\lambda x^\alpha}, \lambda > 0, \alpha > 0. \quad (1)$$

В главе введено понятие норматива численной переменной. Показано использование различных конструкций нормативов (простые пороги, нормативы, зависящие от параметра, серийные нормативы) в различных задачах структуризации медицинских понятий. Предложены методы выработки нормативов различных типов.

**Глава 2** развивает вопросы моделирования процессов. В контексте диссертации моделирование используется как средство выработки общего языка для врача и математика. Врач может легко сопоставить наглядное функционирование модели с наблюдаемыми изменениями состояния больного, а математик может использовать точные закономерности, лежащие в основе модели. В диссертации наряду с обычными имитационными использованы модели скользящих нормативов, позволяющие дать детальное описание отклонений течения процесса у индивидуума от типичного для эталонной популяции; использована также модель динамики точки зрения эксперта в процессе достижения общего мнения.

В начале главы рассмотрены две имитационных модели динамики наблюдаемых процессов.

Модель раскрытия родового канала (использованная далее, в гл. 8) в активной стадии родов позволяет объяснить наблюдаемый резко нелинейный характер процесса геометрическими свойствами взаимодействующих объектов. С ее помощью оказывается возможным придать точный смысл понятию «скорости раскрытия».

Модель описывает процесс прохождения твердого шара (аналог купола черепа плода) сквозь эластичное кольцо, оказывающее вязкое сопротивление (сила натяжения кольца пропорциональна скорости его удлинения).

Рассмотрим твердый шар радиуса  $R$  (рис. 1.), продвигаемый постоянной силой  $F$  через кольцо, оказывающее вязкое сопротивление деформации. Мгновенное состояние этой системы можно описывать величиной угла  $x$ , под которым виден радиус кольца  $r$  из центра шара и высотой  $z$  вершины шара над плоскостью кольца. Угол  $x$  подчиняется дифференциальному уравнению

$$\frac{dx}{dt} + \frac{F}{4\pi^2 \mu R} \cdot \frac{\sin x}{\cos^2 x} = 0. \quad (4)$$

Здесь  $\mu$  - коэффициент вязкости. Обозначим  $a = \frac{F}{4\pi^2 \mu R}$ , и, окончательно, получим:

чим:

$$\frac{dx}{dt} + a \cdot \frac{\sin x}{\cos^2 x} = 0. \quad (5)$$

Решение этого уравнения имеет вид:

$$a(t - t_0) = -\cos x - \ln \left| \frac{1 - \cos x}{1 + \cos x} \right|; \quad 0 < x < \pi/2. \quad (6)$$

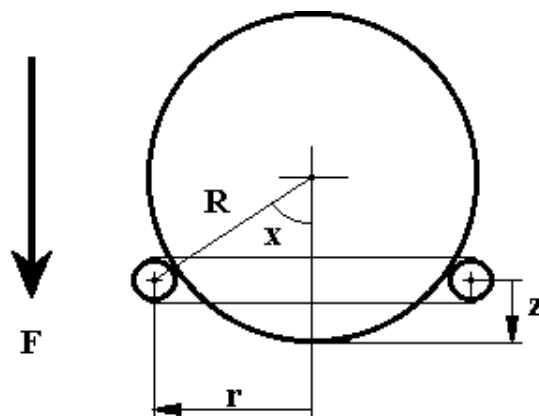


Рис. 1. Геометрия модели.

Вид его показан на рис.2. Качественно поведение диаметра кольца  $D=2r$  совпадает с наблюдаемой динамикой раскрытия родового канала в родах.

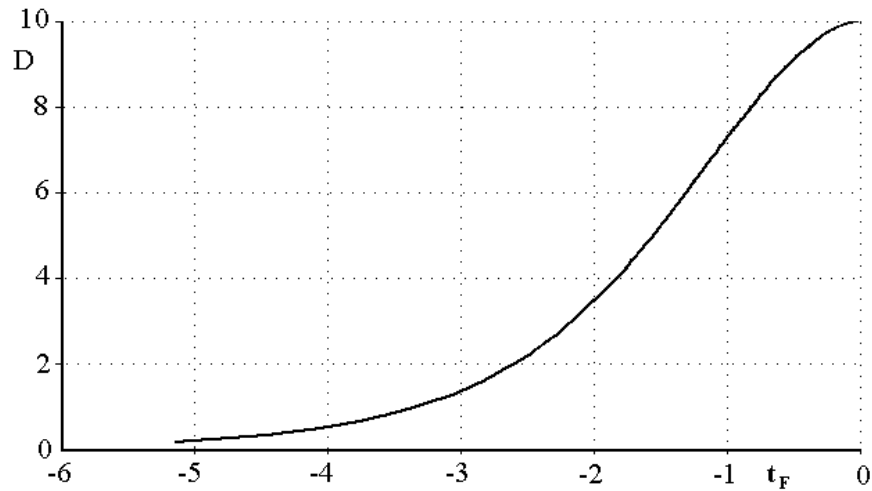


Рис. 2. Решение уравнения

Вторая модель описывает динамику регистрируемого сигнала при лабораторном исследовании процесса фибринолиза, который состоит в растворении (лизисе) нитей фибрина-полимера ферментами плазмы крови. Фибринолиз отвечает за ликвидацию ставших ненужными сгустков крови в кровяном русле. Повышенная или пониженная активность этих ферментов выражается в ускорении (замедлении) процесса лизиса и является одной из причин опасных нарушений гемостаза (состояния свертывающей системы крови) у больных.

Модель фибринолиза в тонком слое позволила связать наблюдаемую форму кривой с константой скорости реакции и параметрами измерительной установки. Представлены два варианта модели процесса, дающие близкие результаты. Первый вариант основан на представлении о расширении зоны лизиса фибрина, занимающей всю толщину слоя, второй – использует модель равномерного растворения фибрина на всей площади зоны лизиса под действием лизирующего агента. Оба варианта приводят к одному и тому же виду регистрируемого оптического сигнала

$$p(t) - p(t_0) \approx a \cdot \sqrt{k} \cdot \sqrt{t - t_0}, \quad (7)$$

где  $a$  – константа,  $k$  – константа скорости реакции,  $p(t)$  – регистрируемый сигнал,  $t$ ,  $t_0$  – текущий и начальный моменты времени. Из полученного соотношения видно, что традиционный метод трактовки наблюдений («площадь зоны через определенный интервал времени») некорректен. Уместнее использовать  $(\Delta p)^2/\Delta t$ .

В обоих случаях моделирования главным результатом оказалась выработка адекватного языка описания содержательного медицинского смысла зарегистрированных процессов. Удалось ввести корректную характеристику скорости процесса, опирающуюся на стандартные измерения.

Представлен класс моделей, используемых при построении скользящих нормативов. Модели описывают совокупность наблюдаемых реализаций процесса как основной динамический процесс с определенным распределением отклонений. Рассмотрен метод построения моделей этого класса по эмпирическим данным. В результате получаем распределение отклонений, функции, описывающие основной процесс, и статистические характеристики отклонений.

На рис.3 показан пример такого норматива для размера головы плода (БПР) относительно гестационного срока (срок беременности). Распределение в каждом из двухнедельных интервалов показано в виде прямоугольника (квартильный интервал) с «пояском» (медиана распределения). Линии демонстрируют сглаженное закономерное изменение размера со временем. Средняя кривая – медиана нормативных значений, крайние – квартили.

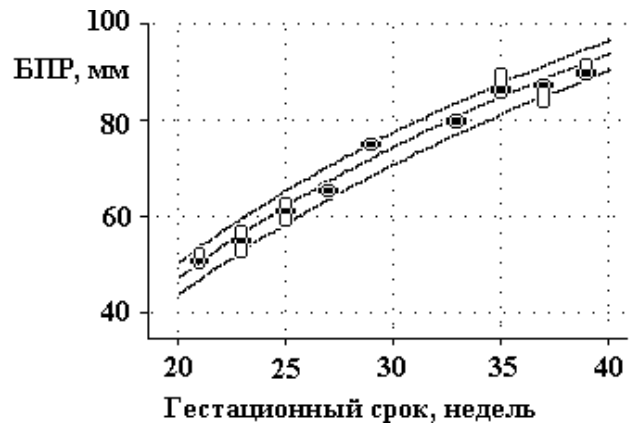


Рис. 3. Пример построения норматива БПР для популяции.

Рассмотрена модель консолидации точки зрения группы экспертов по мере анализа нового для них материала. На каждом этапе работы коллектива экспертов строится формальная модель «общей точки зрения». Эта модель выполняет две функции: точки отсчета для выделения отклоняющихся больных и оценки степени согласия экспертов. Описан метод, ускоряющий выработку общего мнения на основе модели достигнутой общей точки зрения.

**В главе 3** предложен метод анализа и моделирования заключений и выводов врача в процессе решения профессиональной задачи – метод логических симптомов.

Логическим симптомом называется переменная, принимающая 3 значения: ДА, НЕТ, НЕИЗВЕСТНО. Над логическими симптомами определены функции, аналогичные обычным логическим функциям. В отсутствие значения НЕИЗВЕСТНО эти функции переходят в обычные логические функции. Примеры операций над симптомами даны в таб.1.

Здесь НЕ – обычная одноместная операция отрицания; И, ИЛИ – аналоги конъюнкции и дизъюнкции со стандартными экстремальными свойствами; НТЖ – аналог операции «нетождественность», часто употребляемой для сравнения кодов. НАСЛ – операция «наследования», специфическая для логических симптомов. Она принимает «известные» значения только при совпадающих «известных» значениях обоих аргументов. Эта операция дает определенные удобства при формализации процессов выработки правил классификации.

Таблица 1. Результаты операций над симптомами.

у		НЕ			И			ИЛИ			НТЖ			НАСЛ		
	х	1	?	0	1	?	0	1	?	0	1	?	0	1	?	0
1		0	?	1	1	?	0	1	1	1	0	?	1	1	?	?
?					?	?	0	1	?	?	?	?	?	?	?	?
0					0	0	0	1	?	0	1	?	0	?	?	0

На этом языке каждому объекту  $j$  соответствует вектор значений всех симптомов  $\mathbf{x}_j = (x_{j1}, \dots, x_{jn})$ . Объекты, принадлежащие различным классам, могут иметь различные значения координат этих векторов и, следовательно, занимать различное положение в пространстве симптомов.

Для двух векторов  $(\mathbf{x}_a, \mathbf{x}_b)$  с симптомными значениями (симптомные векторы) введены два варианта меры близости: *относительное сходство*  $c(\mathbf{x}_a, \mathbf{x}_b) = G/n$  и *относительная неотличимость*  $d(\mathbf{x}_a, \mathbf{x}_b) = G/H$ . Они различаются тем, что в случае относительного сходства суммарное совпадение  $G$  (число координат с совпадающими значениями) отнесено к размерности вектора  $n$ , а в случае относительной неотличимости – к суммарному присутствию  $H$  (числу координат с «известными» значениями в обоих векторах). Показано, что  $c(\mathbf{x}_a, \mathbf{x}_b) \leq 1$ ,  $G \leq H$  и  $d \leq 1$ . В отсутствие значений НЕИЗВЕСТНО во всех координатах всех векторов обе эти величины переходят в  $1 - R_h/n$ , где  $R_h$  – расстояние по Хэммингу, а  $n$  – размерность векторов.

Логические симптомы пригодны для анализа наблюдений, заключений врача и правил постановки диагнозов в статическом случае (когда решение принимается одномоментно).

Задачу классификации объектов можно сформулировать как задачу получения разделяющей функции для объектов в пространстве симптомов, значения которой позволяли бы отнести новый объект к определенному классу. В практических задачах классификации эту функцию приходится строить «на примерах», т.е. используя наборы объектов, заведомо принадлежащих сравниваемым классам (обучающая выборка).

Будем считать, что справедлива *гипотеза делимости*: в каждом классе найдется хотя бы один объект, не совпадающий ни с одним объектом остальных классов.

Функцию принадлежности объекта определенному классу можно записать в виде логического выражения, связывающего значения симптомов. Например, для объектов двух классов А и В (таб. 2), функция  $f_1 = \overline{x_4} \wedge \overline{x_5} \wedge x_6$  равна 1 для объектов класса А и нулю – для объектов класса В. Цель стандартной классификационной задачи – получение такой функции для сравниваемых классов. Функция  $f_2 = \overline{x_1}$ , в качестве классификационной, дает противоречие на одном из объектов класса В. Это пример отказа правила. Другой способ задания принадлежности объекта классу дает процедура голосования.

Таблица 2. Примеры классификационных функций.

Класс	вектор					$f_1$	$f_2$
A	0	?	1	0	0	1	1
A	0	1	0	0	0	1	1
A	0	0	1	0	0	1	1
B	1	?	0	1	1	0	0
B	0	1	0	1	1	0	1
B	1	1	0	1	1	0	0

Процедура голосования задается списком участвующих симптомов, указанием для каждого симптома, какое его значение соответствует голосу за данный класс, и пороговым значением числа голосов, соответствующим принятию решения. В нашем формализме первые две группы сведений записывают в виде **маски**, т.е. вектора значений симптомов, соответствующих голосованию «ЗА» данный класс. Координаты, значения которых безразличны (например, принимают различные значения для объектов данного класса), будут закодированы в маске неизвестным значением (?). Пример голосования приведен в таблице 3.

Таблица 3. Выделение класса голосованием.

Класс	Вектор					S	
Маска	0	?	?	0	0	1	
A	0	?	1	0	0	1	4
A	0	1	0	0	0	1	4
A	0	0	1	0	0	1	4
B	1	?	0	1	1	0	0
B	0	1	0	1	1	0	1
B	1	1	0	1	1	0	0

Заданы два класса A и B, содержащие по 3 объекта. Задана маска, занимающая верхнюю строку таблицы. Каждая координата вектора объекта, совпадающая с маской, дает вклад 1 в сумму S для этого вектора. Координаты со значениями (?) в маске проверке не подлежат. Для объектов класса A все суммы равны 4, а для объектов класса B принимают значения 0, 1. Простое правило классификации позволяет отнести к классу A все объекты со значением суммы больше, скажем 3, а с меньшими значениями – к классу B.

По заданным образцам объектов, принадлежащих классу, можно сформировать маску класса и подобрать порог принадлежности. В этой главе предложены 3 алгоритма формирования маски для разделения двух классов: поразрядное наращивание маски, формирование маски с использованием критерия  $\chi^2_{2*2}$  и алгоритм наследования. Первый из них пригоден, скорее, для теоретического анализа, остальные – удобны на практике.

Приведены примеры моделирования диагнозов на основе симптомного описания в практических задачах.

**Глава 4** посвящена языку моделирования динамики решений врача или накопления данных по мере поступления информации во времени. Для этого разработан метод анализа сценария процесса с помощью фактов и событий, в котором для каждого наблюдения может учитываться предыстория процесса.

В каждый момент времени  $t$  ( $t = 1, 2, \dots, T$ ) мнение специалиста о состоянии объекта можно представить вектором  $\mathbf{x}(t) = (x_1(t), \dots, x_M(t))$ , где каждая координата  $x_i(t)$  (мнение специалиста об одном из свойств объекта в момент  $t$ ) имеет свой конечный алфавит возможных значений. *Сценарием процесса* называется полная совокупность всех векторов  $(\mathbf{x}(1), \dots, \mathbf{x}(T))$ , описывающая весь процесс. Сценарий допускает наглядное представление в виде таблицы из  $N$  строк и  $T$  столбцов, в клетках которой располагаются символы значений координат или пробелы в качестве значения «неизвестно».

Для формального описания элементов сценария введены понятия факта (наблюдение или заключение, привязанное к определенному интервалу или моменту времени) и события. *Факт*  $f(t^*, x_i(t), t_1, t_2, \mathbf{v}, a)$  задают сочетанием 5 элементов: момент реализации факта (момент наблюдения)  $t^*$ ; проверяемая координата  $x_i(t)$  вектора  $\mathbf{x}(t)$ ; интервал времени  $(t_1, t_2)$ ; образец набора значений  $\mathbf{v}$ ; правило  $a$  соответствия наблюдаемых значений значениям в образце. Если в интервале времени  $(t_1, t_2)$  значения проверяемой координаты  $x_i(t)$  могут быть соотнесены по правилу  $a$  значениям, приведенным в образце  $\mathbf{v}$ , факт принимает в момент  $t^*$  логическое значение «ДА» (1), во всех остальных случаях – значение «НЕТ» (0). Приведено описание простейших свойств фактов.

Назовем *итогом* процесса некоторый факт, поддающийся независимой верификации. В зависимости от решаемой задачи итог может быть окончательным решением специалиста после наблюдения всего изучаемого процесса, или фиксированным объективным состоянием процесса. Первый случай реализуется при анализе рассуждений специалиста с целью получения модели вывода его заключения. Второй – при выработке прогноза течения наблюдаемого процесса.

Событие – функция от совокупности фактов, определенным образом расположенных во времени. Будем называть *событием* совокупность фактов, присутствующих или отсутствующих в сценарии и имеющих общий момент реализации  $t^*$ , к которому и будет отнесено событие. Поскольку событие оперирует лишь наличием фактов, его можно представить логической функцией  $e(f_1, \dots, f_n)$  входящих в него фактов.

Событие представляет собой логическую конструкцию, которую можно интерпретировать, в зависимости от рода процесса, либо как наблюдаемую закономерную последовательность явлений (диагностическая интерпретация), либо как конструкцию рассуждения специалиста на основе постепенно поступающей информации о процессе.

События удобны для прогнозирования некоторого итога процесса, наступающего в момент времени  $t^{**}$ . Если момент события  $t^* < t^{**}$ , то в некоторых случаях возможно активное воздействие на процесс, изменяющее его дальней-

шее течение. В медицинских задачах эта ситуация соответствует возможности лечения.

Рассмотрим пример построения факта по наблюдаемым сценариям процесса в двух классах объектов (таб.4).

Таблица 4. Сценарии процесса для двух классов.

Кл	Обь	12	15	18	21	24	27	30	33	36
P	4004	...	...	...	...	..B	.B.	...	.B.	B..
P	4015	...	...	...	...	...	...	...	B..	CC.
P	4117	...	...	.A.	...	...	..B	..B	...	...
P	4199	...	...	B..	...	B..	..C	C..	..C	..C
Q	4255	...	...	A.A	...	.A.	..A	...	A..	.B.
Q	4324	.A.	...	..A	..A	...	.C.	C..	C..	.CC
Q	4325	...	...	...	A..	A..	A..	...	...	.C.
Q	4490	...	...	A..	A.A	...	AA.	AA.	.B.	B.B
Q	4554	...	..A	...	...	...	.A.	A..	...	.B.

В таблице приведены наблюдаемые значения переменной  $x$ , характеризующей процесс, в последовательные моменты времени. Каждая строка таблицы соответствует значениям переменной для одного объекта. Идентификатор (номер) объекта находится во втором столбце таблицы, в первом – размещается буквенное обозначение класса.

Столбцы 3-11 содержат наблюдаемые значения переменной, в каждом столбце – для трех последовательных моментов времени, начиная с указанного в заголовке столбца. Например, третий столбец с заголовком 12 содержит значения  $x$  для моментов 12-14. Буквы A, B, C являются значениями переменной  $x$ . Видно, что в классе Q имеется большое количество значений A, почти отсутствующих в классе P.

Выберем простейшее правило проверки (нестрогий список), образующее факт по каждому встреченному значению переменной. Тогда факт  $(x = A)$  уже достаточно хорошо различает классы P и Q. В самом деле, он встречается только у одного объекта класса P из четырех и у всех 5 объектов класса Q. Построим событие на основе рассмотренного факта, с нестрогим списком. Запишем факт в виде  $f_1 = (x = A[0,0])$ , где знак равенства в скобках означает принадлежность значений к нестрогему списку из одного элемента (A), а в квадратных скобках указан интервал проверки значений переменной относительно момента наблюдения  $t^*$ , здесь он совпадает с самим моментом наблюдения. Заметим, что у всех объектов класса Q этому факту предшествует такой же, но на 1-3 единицы времени раньше. Поэтому событие, характерное для объектов класса Q, можно записать в виде:  $e_1 = (x = A[-3,-1]) \wedge (x = A[0,0])$ . Запись означает обязательное наличие в сценарии двух фактов описанного выше типа, разделенных

интервалом времени от 1 до 3 единиц. Это событие выполняется для всех объектов класса Q и не выполняется для объектов класса P.

Использованная выше запись фактов и событий позволяет легко придавать им словесную форму, принятую для медицинских диагнозов и экспертных заключений.

**Глава 5.** Посвящена описанию программных средств, использованных в работе.

В главе 5 описан набор программ, разработанных автором и составивших минимальный комплект обслуживания медицинских задач. Программы опираются на файлы базы данных в формате dBASE III. Прежде всего, к ним относятся программы, расширяющие технические возможности базы. Это калькулятор произвольных формул, графический иллюстратор и программа подготовки специализированных текстов для диалога в ходе диагностических игр на определенную тему.

Калькулятор производит вычисления по заданным формулам (длина – до 1000 символов), используя входные числа из базы данных.

Графический иллюстратор визуализирует одномерные зависимости  $y = f(x)$ , где обе переменные – поля базы данных. В общем кадре экрана с равномерными шкалами по осям можно вывести одновременно до 20 графиков.

Программа подготовки текстов преобразует заданную информацию из комплекта файлов базы данных в текст, соответствующий обычным медицинским требованиям. Она позволила заметно уменьшить трудоемкость диагностических игр, сделать более стандартной информацию, подготавливаемую к очередному сеансу игры, и повысить четкость интерпретации результатов сеанса за счет контролируемого исключения посторонних «смыслов».

Программа контроля и первичного анализа данных выводит распределения дискретных переменных в табличном виде и позволяет проверять простейшие гипотезы об их связи.

Программа анализа числовых переменных предназначена для исследования эмпирических распределений. С помощью встроенных непараметрических критериев эта программа помогает найти наилучшие разделяющие границы сравниваемых классов, которые можно использовать для формулирования прогностических и диагностических симптомов и фактов.

Программа симптомного анализа оперирует логическими симптомами. Она позволяет строить и испытывать статические диагностические правила. Правило может моделировать закономерности наблюдаемой ситуации или рассуждение врача о ней, в зависимости от того, какие симптомы использованы в составе правила. Благодаря использованию «вложенных заключений» программа может оперировать правилами, сложность которых очень велика.

Программа регрессий и скользящих нормативов предназначена для поиска связей переменных из разнообразных численных массивов, включая небольшие массивы данных с резко нелинейными связями. Она позволяет в диалоговом режиме построить линию регрессии, описываемую произвольной достаточно сложной функцией и аппроксимировать распределение отклонений от нее.

Программа анализа сценариев рассматривает реально происходящий процесс (болезнь, лечение, выработку решения) как последовательность фактов, привязанных к определенным моментам времени, и событий, представляющих собой сочетания фактов с несовпадающими, в общем случае, интервалами проверки значений. Эта программа позволяет моделировать последовательность наблюдений во времени состояния больного или постепенную выработку решения врачом по мере накопления сведений о больном и болезни.

**Главы 6 – 12** содержат описание медицинских задач, решение которых существенно использует описанные выше методы. Отметим общие черты этих задач. Во-первых, точная постановка задачи, как правило, отличается от первоначального пожелания врачей. Формализация начинается с самого первого этапа работы над медицинским вопросом. Во-вторых, обычно в каждой из задач используются описанные в 1-5 главах методы в различных сочетаниях в зависимости от особенностей задачи. В-третьих, полученные правила используют небольшое количество переменных. В-четвертых, полученные правила удается формулировать, если и не прямо в медицинских, то, во всяком случае, в понятных врачу терминах.

Задачи относятся к различным областям медицины, и объединяет их интерес к рассуждению врача, приводящему к определенному решению, а также общий набор методов анализа данных и решений врача.

**Глава 6.** Задача о прогнозе артериальной гипертензии после хирургической коррекции коарктации аорты была исторически первой в рамках предложенного подхода. В ходе работы над ней выработывался стиль диалога с врачом, способы верификации диагноза, приемы анализа неполных данных. В диалоге со врачом было выработано понятие опасного класса больных. Диагностику больных по всей совокупности данных (включая все время послеоперационного наблюдения) врач выполнял дважды с интервалом в несколько месяцев. В течение этого времени проводились также обсуждения деталей операции и послеоперационного ведения больных. В опасный класс были включены больные, дважды отнесенные врачом к категории больных с ранней гипертензией.

На рис. 4 приведена диаграмма рассеяния для двух классов.

По оси абсцисс отложены значения разности ( $sys1 - sys0$ ) систолических давлений после и до операции, по оси ординат – отношение (ОЦК/ масса) объема циркулирующей крови до операции к массе тела больного.

Прямоугольник  $(-80,2 ; 50,92)$  выделяет область, не содержащую случаев гипертензии. Она соответствует логическому симптому

$$((\text{ОЦК/ масса}) \leq 92) \ \& \ ((sys1 - sys0) \leq 2).$$

Этот логический симптом позволяет прогнозировать уже в день операции опасность появления артериальной гипертензии впоследствии.

Для больных опасного класса найдено правило индивидуального прогноза, на основе построения одного логического симптома, использующего значения двух переменных.

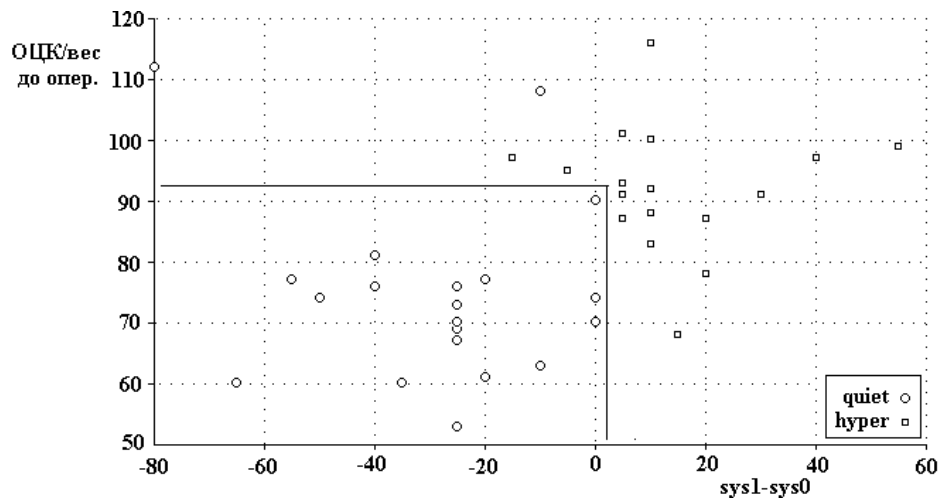


Рис. 4. Использование двух признаков для разделения классов.

**Глава 7** посвящена задаче о плацентарной недостаточности. Плацентарная недостаточность относится к категории болезненных состояний организма, которые еще не имеют единого определения, но, тем не менее, считаются достаточно опасными и требуют лечения.

В задаче о диагнозе и прогнозе возникновения плацентарной недостаточности основной акцент был сделан на разработку многоцелевого вопросника. В этой главе обсуждены основные вопросы сбора информации, предварительного анализа ее структуры и построения тематического вопросника. Вопросник содержит традиционные части «паспорт» и «анамнез», идентифицирующие больного и его состояние до начала лечения. Наряду с ними в вопросник включаются данные наблюдения за больным, текущего лечения с обязательными заключениями врачей, обосновывающими выбор лечения, и поэтапными оценками результатов.

Уточнение постановки задачи было достигнуто путем установления апостериорного критерия плацентарной недостаточности, не излеченной к моменту родов, - рождение ребенка со сниженной массой тела. При этом сам критерий потребовал дополнительного исследования для установления скользящего норматива массы тела новорожденного.

Ощутимый практический результат принесло решение задачи о прогнозе рождения маловесного ребенка по однократному ультразвуковому исследованию (УЗИ). Для решения задачи использована информация, собранная с помощью вопросника. На основе скользящих нормативов для ультразвуковых размеров тела плода получено несколько правил индивидуальной диагностики, применимых даже в ранних сроках беременности. Использование скользящих нормативов позволило перевести все переменные в безразмерные. Рассмотрение совместных распределений дало границы, отделяющие классы с различными результатами родов. Получены 3 логических симптома, реализующих три прогноза задержки внутриутробного развития плода. На их основе выработано правило прогнозирования возможной степени задержки развития плода в от-

сутствие лечения. Во всех правилах использованы 8 симптомов на основе 3 наблюдаемых переменных. Эти правила достаточно рано сообщают акушеру о начавшемся отставании роста плода и побуждают начать лечение.

**Глава 8.** Задача о выборе тактики родов (*самопроизвольные или оперативные*) у женщины с рубцом на матке представляет собой достаточно сложную медицинскую проблему. С точки зрения врача цель состояла в том, чтобы избежать опасности разрыва матки в родах по рубцу.

Прогноз разрыва требует недоступной информации о состоянии мышечной ткани стенки матки перед родами и в родах, поэтому информационную задачу поставили как прогноз *решения врача* об оперативном родоразрешении женщины или продолжении подготовки к самопроизвольным родам.

Для уточнения постановки задачи были использованы диагностические игры. При этом удалось выделить 3 этапа принятия решения врачом. Первый раз решение обязательно должно быть принято по результатам предварительного обследования больной сразу же после ее поступления в клинику в сроке 37-38 недель. Родоразрешение женщин с рубцом на матке обычно стараются проводить в плановом порядке, поэтому день родоразрешения выбирают заранее по состоянию матери и плода. Тем самым, обязательный второй момент – решение вопроса о методе родоразрешения накануне родов. Третий критический момент может возникнуть в активной фазе родов в случае затруднения или замедления изгнания плода. Эти три момента принятия решения оставались обязательными для всех рассмотренных случаев.

На каждом этапе решения врача моделировались сложными логическими симптомами. Специально для оценки роли некоторых сведений был разработан метод «экспертной экстраполяции». Он сводится к объединению в один класс больных, для которых врач принял одинаковое решение в одинаковых обстоятельствах, независимо от априорного типа больного.

Уже простая статистика показала, что на первом этапе состояние организма женщины в момент обследования теснее коррелирует с окончательным решением врача, чем имеющиеся сведения о прошлых операциях и осложнениях. Это определило метод построения решающих правил как логического обоснования решений врача объективными показателями состояния организма.

В результате, на первом этапе получили минимальный набор из трех логических симптомов (таб. 5): 1) низкая зрелость шейки матки (до 4 баллов Bishop), 2) разгибание головы плода, 3) малая толщина стенки нижнего сегмента матки (меньше 5 мм).

Наличия хотя бы двух симптомов из таб. 5 достаточно, чтобы рекомендовать операцию. Отметим, что все эти симптомы характеризуют степень готовности организма женщины к предстоящим родам.

Таблица 5. Симптомы, влияющие на решение врача на 1 этапе.

01	NUMBALL ≤ 4	Баллы зрелости шейки
02	RAZGIB == 2	Разгибание головы плода есть
03	MINTOL ≤ 4	Минимальная толщина матки

На втором этапе подготовки к родам оказались важны другие симптомы: бипариетальный размер головы плода (БПР) и отношение к нему трех размеров таза женщины из числа используемых в акушерской практике (таб. 6.)

Таблица 6. Симптомы, использованные на втором этапе.

сим	Условие	Комментарий
1	$PCB9 > 46.0$	Перцентиль БПР в родах
2	$TR\_B > 3.34$	<i>Distantia trohanterica</i> /BPR
3	$EX\_B < 2.25$	<i>Conjugata externa</i> /BPR
4	$DI\_B < 1.41$	<i>conjugata diagonalis</i> /BPR

Эти симптомы характеризуют относительные размеры (пересчитанные на момент родов) плода и родового канала, через который ему предстоит пройти. Если выполнены все эти условия, рекомендуется оперативное родоразрешение

Для выбора переменных, описывающих активную стадию родов, была разработана аппроксимационная модель открытия родового канала. Модель использует представление о продвижении твердого купола черепа плода через мягкие ткани, оказывающие вязкое сопротивление.

Зависимость величины раскрытия от времени в модели близка к наблюдаемой. Сравнение с записями американского исследователя Е.А. Фридмана показывает хорошее количественное совпадение (рис. 5). На рисунке кружками показаны измерения Фридмана, а линией – решение уравнения модели.

Это совпадение дает возможность определить скорость наблюдаемых родов в безразмерных единицах и использовать ее для оценки возможных нарушений процесса. Скоростью можно считать отно-

шение  $\Delta t_{\text{Фрид}}/\Delta t_{\text{набл}}$  интервала времени модели к интервалу астрономического

времени при изменении раскрытия от первого до второго значения. Все тяжелые случаи, требующие срочного вмешательства показывают относительную скорость родов не более 0.15, случаи осложнений менее опасных 0.15-0.28, и случаи благополучные – более 0.28.

На этом этапе решающее правило принимает во внимание динамику развития процесса и сводится к оценке результата по начальному отрезку процес-

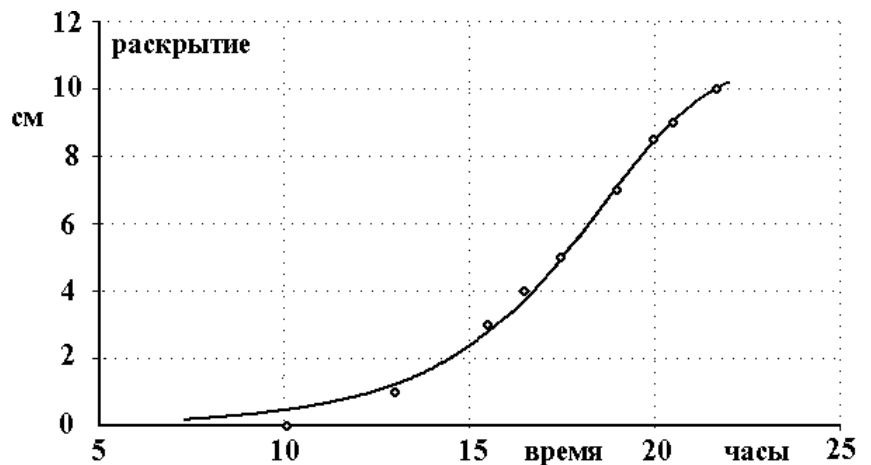


Рис. 5. Аппроксимация реального раскрытия в родах решением уравнения.

са. На основе этого правила разработана конструкция простого калькулятора из двух соосных дисков, позволяющего врачу решать вопрос о критичности состояния данной женщины непосредственно в родильном зале.

В результате исследования получено правило трехэтапного принятия решения врачом о методе родоразрешения конкретной женщины. Правило не противоречит обычной схеме действий врача и фиксирует его внимание на критических параметрах каждого этапа.

В задаче о прогнозе решения врача на различных этапах ее решения были использованы: диагностические игры (на предварительном этапе анализа задачи), сочетания симптомов на 1 и 2 этапах, имитационная модель для определения скорости процесса на 3 этапе, экспертная экстраполяция. Это позволило получить содержательные правила, использующие в совокупности менее 10 переменных из более, чем 450 первоначальных.

**Глава 9** содержит решение задачи о ведении и лечении беременной с сахарным диабетом. Цель лечения состоит в том, чтобы предотвратить возможную гибель ребенка. Трудность состоит в том, что лечить женщину нужно в течение всей беременности, а результат лечения становится известен лишь после родов.

Было предложено несколько правил прогноза существенного ухудшения состояния плода по сведениям, доступным на ранних сроках беременности.

Первый вариант был основан на наблюдаемом нарушении состояния свертывающей системы крови матери. Резкое уменьшение активности агрегации тромбоцитов в тесте с воздействием коллагена было обнаружено у 5 женщин из первых 30 наблюдавшихся.

Среди их детей выжил 1 и погибли 4 ребенка. В дальнейшем таким женщинам было показано лечение ингаляциями гепарина. Эффективность лечения видна уже из первых наблюдений (таб. 7). Было в дальнейшем обнаружено еще 5 женщин с нарушением агрегации тромбоцитов. Последних пятерых лечили ингаляциями гепарина. Хотя статистика мала (редко удавалось зафиксировать сравнительно кратковременные и редкие нарушения агрегации), но польза лечения заметна.

Здесь для оценки качества разделения использованы оценки достоверности по критериям Точного Метода Фишера (ТМФ)  $p_F$  и  $\chi^2$  (величина статистики критерия  $\chi^2$  и вероятность  $p$ ). Для выработки текущих прогнозов выживания плода был использован анализ сценариев хода лечения.

Таблица 7. Дети женщин с нарушением агрегации тромбоцитов.

	лечили	Не лечили	$p_F = 0.2$
Жив	4	1	$\chi^2 = 1.6$
Умер	1	4	$p = 0.2$

Сравнили распределения перцентиля размера головы плода (БПР) в классах выживших (L), умерших антенатально (до родов, A), интранатально (в родах, I) и постнатально (после родов, P). Граница (PCBPR=9.4) соответствует

наилучшему разделению классов (L:P) по критерию Смирнова, граница (PCBPR=47.5) – то же для классов (L:A). Значения PCBPR, встречающиеся с большей частотой в классах погибших (A и P), можно описать с помощью фактов (PCBPR<9.4) и (PCBPR>47.5). Эти факты соответствуют распространенному среди акушеров представлению о повышенном риске рождения «излишне маленьких» (всегда) и крупных (при диабете) детей. Аналогично формируются факты, описывающие ультразвуковые размеры плода, уровни гормонов и сахара крови, разграничивающие выживших и погибших детей. В таб. 8 приведены формулировки этих фактов.

Таблица 8. Факты, прогнозирующие гибель ребенка.

Факт	Условие	Пояснение
CPCEST=a	pcest > 71.6	Перцентиль эстриола
CPCPGS=a	pcpgs < 29.9	Перцентиль прогестерона
CPCCRT=ab	pcprt > 67.25	Перцентиль кортизола
CGU2=ab	glucose2 > 8	Уровень макс. сахара
PCBPR=ab	pcbpr < 9.4 pcbpr > 47.5	Перцентиль размера головы плода
CPCDG =b	pcdg > 60.65	Перцентиль диаметра груди плода
CPCDV =a	pcdv > 28.05	Перцентиль диаметра живота плода

Используя эти факты, сформулируем два прогностических события (таб. 9) для отрицательного результата лечения.

Таблица 9. Формулы прогностических событий для гормонов, размеров плода и уровня сахара в крови матери.

@event= 1(A) ((CPCEST=a)&&(CPCPGS=a)&&(CPCCRT=ab))
@event= 2(B) ((CGU2=ab)&&(PCBPR=ab)&&(CPCDG =b)&&(CPCDV =a))

Появление хотя бы одного из этих событий в ходе беременности у женщины с сахарным диабетом служит сигналом опасности, угрожающей плоду. Таблица сопряженности для этого правила (таб. 10) показывает, что отсутствие этих событий является хорошей гарантией успешного развития ребенка, а их появление в половине случаев предвещает его гибель.

Таблица 10. Оценка качества прогноза. Все случаи - без лечения гепарином и без инфузионной терапии.

Признак (A∨B)	нет	есть	$p_F=4*10^{-4}$ $\chi^2=12.4$ $p=6*10^{-4}$
L (жив)	51	5	
AP (умер)	9	9	

Если теперь, кроме событий А,В ввести события, связанные со специфическим лечением после обнаружения угрозы, то можно оценить эффективность лечения (таб. 11). В этом случае использован контингент, наблюдавшийся после введения в практику методов лечения гепарином и инфузиями.

Таблица 11. Больные с отрицательным прогнозом. Лечение: гепарин, инфузии.

	Не было	Было	$p_F=0.02$
L (жив)	19	23	$\chi^2=4.64$
AP (умер)	24	9	$p=0.03$

Эти примеры показывают, что события, учитывающие формализованные опасения врача, могут служить для успешного прогнозирования результата лечения.

Для прогнозирования ухудшения состояния плода можно использовать не только данные анализов, но и более простые методы, не требующие специального оборудования и методик. Врачи, ведущие беременных с сахарным диабетом, знают, что избыточная масса тела матери осложняет внутриутробное развитие плода. При этом понятие «избыточная масса тела» остается нечетким. Более детальное изучение показало, что группы женщин с погибшими и выжившими детьми статистически неразличимы по абсолютной величине массы тела, по соотношению масса/рост, прибавке массы за время беременности, и т.д.

Динамический норматив массы тела по срокам беременности позволил получить прогноз лучше прежних, но еще недостаточно точный. Удовлетворительным оказался прогноз, использовавший два входа: норматив зависимости массы тела от роста («телосложение») и динамику изменения по срокам беременности. Выработанный при этом популяционный норматив давал для женщин с сахарным диабетом критерий «опасной полноты». Значение перцентильной оценки текущей массы тела женщины в данном сроке, большее 37.3, сигнализировало о возрастании риска гибели ребенка (таб. 12).

Таблица 12. Выживаемость детей в зависимости от перцентиля текущей массы тела матери при беременности.

	$p_{cm0h}<37.3$	$p_{cm0h}\geq 37.3$	$p_F=0.006$
L (жив)	43	56	$\chi^2=6.4$
AP (умер)	5	26	$p=0.01$

Строки таблицы сопряженности занимают классы живых (L) и умерших детей (AP). Столбцы соответствуют низкому и высокому значению перцентиля текущей массы тела у матерей с сахарным диабетом во время беременности. Видно, что в классе с малой текущей массой достоверно мало погибших детей.

Из таблицы видно, что доля женщин с меньшей массой среди сохранивших детей живыми (L) выше, чем у тех, чьи дети погибли (AP). Тем самым показана адекватность модели динамики массы тела для женщин с сахарным диабетом.

Для массы тела женщины в различные сроки беременности был построен скользящий норматив по всей популяции беременных, учитывающий телосложение. Отклонения индивидуальной массы тела были использованы для прогнозирования жизнеспособности новорожденного. Избыточная масса тела приводила к ухудшению состояния плода и последующей гибели ребенка.

Тем самым, удалось установить, какое понятие стоит на самом деле за утверждением о большем риске для жизни детей у полных матерей. Имеется в виду не конституционная полнота и не абсолютная величина прибавки массы за беременность, а слишком большая динамическая прибавка по сравнению с индивидуальным нормативом. Высказывание о риске, очевидно, имеет рекомендательный характер, а не указание на абсолютную опасность. Возможно, что если женщине удастся удерживать массу тела в рекомендованных пределах, она просто относится к группе меньшего риска по каким-то жизненно важным параметрам.

В ходе работы над вариантами прогноза ухудшения состояния плода удалось обнаружить важную особенность развития плода при сахарном диабете у матери. Использование скользящих перцентильных нормативов для размера головы плода позволило установить факт возникновения плацентарной недостаточности при беременности, отягощенной сахарным диабетом.

Рассмотрим в качестве характеристики внутриутробного роста плода значение перцентильной оценки индивидуального размера головы плода (БПР). Если рост плода происходит в соответствии с физиологическими закономерностями, то перцентильная оценка размера должна оставаться постоянной в течение всей беременности, поскольку стандартная динамика роста уже содержится в перцентильном нормативе. Если при этом размер соответствует нормативу, константа должна быть близка к 50, во всяком случае, для половины плодов она должна лежать между 25 и 75.

При сахарном диабете приходится наблюдать совсем другую картину (рис. 6). На рисунке приведена диаграмма зависимости перцентилей размера головы плода

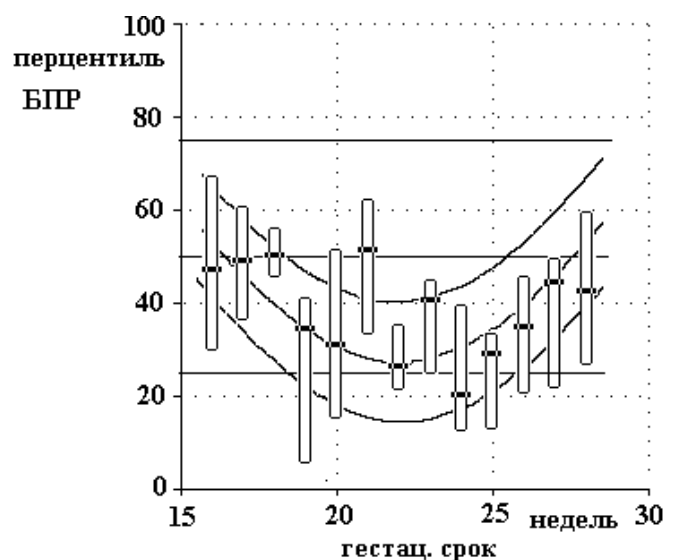


Рис.6. Изменение перцентилей БПР во 2 триместре беременности

(ордината) от гестационного срока (абсцисса). Горизонтальные линии на уровнях 25, 50 и 75 показывают, где должны были бы располагаться нижний квартиль, медиана и верхний квар-

ть распределений размера головы плода, если бы равномерность его развития не нарушалась. Прямоугольники с «пояском» демонстрируют квартильные интервалы (от нижнего квартиля до верхнего) изменения наблюдаемого перцентиля размера головы плода в пределах недельных отрезков гестационного срока. Поясок соответствует медиане распределения значений перцентиля БПР на недельном отрезке.

При равномерном развитии плода все эти прямоугольники должны были бы располагаться между крайними горизонталями так, чтобы пояски лежали на средней горизонтали. Однако, кроме начального отрезка (16-18 недель) и единичного прямоугольника в 21 неделю, это не так. На рисунке отчетливо виден сдвиг прямоугольников в диапазоне 19-26 недель вниз.

Для наглядности демонстрации этого сдвига проведены кривые, аппроксимирующие положение этих прямоугольников (нижний квартиль, медиана и верхний квартиль).

При решении задачи о жизнеспособности ребенка при диабете у матери были использованы: формализация решений врача-лаборанта, анализ симптомов и сценариев лечения, прогнозирование результата по данным до лечения, построение скользящего норматива массы тела женщины.

**Глава 10** посвящена проблеме выработки критериев оценки состояния новорожденных от матерей с сахарным диабетом. Эта оценка используется при подведении итогов лечения женщины и позволяет вырабатывать более точные критерии назначения лечения во время беременности.

Основой для построения оценки послужили две группы новорожденных, единодушно оцененные лечащими врачами как вполне благополучные или крайне плохие. При сравнении этих групп были построены логические симптомы, принимавшие различные значения на этих группах. Были приняты во внимание качественные признаки, характеризовавшие «крайне плохих».

Часть симптомов отбраковали из-за относительно редкой встречаемости. В работе использовали оставшиеся 29 симптомов. По ним выделили 9 классов детей, приблизительно следующих шкале оценки состояния их нервной системы. Градации шкалы были выбраны «с избытком» – они вмещали практически все варианты врачебных оценок. Некоторые из смежных классов оказались похожи друг на друга по перечню симптомов и одновременно по словесным заключениям врачей о состоянии детей. Такие смежные классы были объединены. Осталось 5 классов (4 укрупненных, А-D, в порядке ухудшения и, дополнительно, Е - умершие постнатально, после рождения).

Для этих классов заново построили симптомы, различающие их. Из 131 симптома отобрали 91, дающие различие не хуже  $p < 0.05$  по критерию хи-квадрат. На этих симптомах построили маски классов. В маску соответствующего класса включали симптом, встречающийся не менее, чем у 80% больных этого класса, и не более, чем у 20% каждого из остальных. Для каждого объекта было вычислено относительное несовпадение  $1 - c(x_i, x_j)$  вектора объекта с маской каждого класса. Несовпадение с маской благополучного класса А оказалось удачным классификационным признаком (таб. 13).

Таблица 13. Формальная классификация и ранее полученные классы.

Клас-сы	Диапазоны относит. несовпадения					df=16 Y=102 p=10 <sup>-13</sup>		
	0-0.045	0.045-0.333	0.333-0.526	0.526-0.737	0.737-1.000			
А	6	0	0	0	0	Совпад.	40	77 %
В	0	5	0	0	0	Выше на 1	3	6 %
С	0	2	6	1	0	ниже на 1	9	17 %
Д	0	0	5	12	2			
Е	0	0	0	2	11			

Столбцы таблицы соответствуют диапазонам значений величины относительного несовпадения векторов объектов и маски класса А, строки – первоначальным классам.

Из 13 наиболее часто встречающихся симптомов были сформированы факты (с учетом моментов их проявления). Четыре события, основанные на этих фактах, позволяют не только сформулировать еще один удобный вариант прогноза, но и дают возможность увидеть динамику выявления опасного состояния детей по дням наблюдения. Почти все дети, принадлежащие опасным классам (С-Е), обнаруживаются по прогностическим событиям к концу первых суток наблюдения (38 случаев из 41).

Полученное правило полностью решает задачу сравнительно экономными средствами: использованы всего лишь 4 события, основанные на наблюдении за 7 параметрами состояния ребенка, отобранными из более чем 640 первичных показателей.

Тем самым показано, что использование формального аппарата логических симптомов позволяет ввести объективную классификацию больных по наблюдаемым особенностям их состояния.

**В главе 11** описано решение задачи обучения коллектива экспертов, анализу новой для них ситуации в пределах профессиональной области. Эксперты проводят несколько сессий оценки состояния свертывающей системы крови у больных. После каждой сессии, по ее итогам, вырабатывается по формальному правилу распознавания «общая точка зрения». Образцы, оценка которых расходится с «общей точкой зрения», обсуждают, приводя аргументы. Динамика изменения отклонений от «общей точки зрения» служит удобным индикатором прогресса в согласовании мнений.

Исследованию подвергли группу детей с гломерулонефритом, у которых был установлен характерный для этого заболевания дисбаланс между показателями гемостаза. Всего обследовано 52 ребенка в возрасте от 3 до 14 лет, у которых различали смешанную, гематурическую и нефротическую формы заболевания. Дети находились на традиционном лечении (кортикостероиды, цитостатики и дезагреганты) с применением у 13 из них ингаляций антикоагулянта (гепарина).

Три высококвалифицированных эксперта по системе гемостаза были привлечены для оценки состояния детей в ходе лечения. Их советы были необходимы для корректного лечения детей ингаляциями гепарина. Эксперты обладали солидным опытом диагностики системы гемостаза, но не занимались раньше этой специфической группой больных практически. Требовалось мобилизовать профессиональные знания и опыт экспертов, и создать им условия для возможно более точной оценки состояния системы гемостаза у каждого больного в каждый момент времени. Критерии для отнесения больных к классам гипо-, нормо- или гиперкоагуляции предстояло выработать, используя опыт и знания экспертов по коагулологии.

Для получения и проверки формальных выводов были использованы 170 карточек, каждая из которых содержала результаты проведенных анализов и измерений для одного больного в один день. Из этих карточек часть (55 штук) была использована для выработки решающего правила, а остальные составили «экзаменационный» массив для независимой проверки эффективности правила.

На первом этапе эксперты работали независимо, а решения экспертов признавались значимыми только при условии единогласия. Наличие особого мнения у одного из экспертов было достаточным для отнесения карточки в категорию «неизвестно».

В результате первой экспертизы карточки разделились следующим образом: норма-5 (N), гипо-13 (D), гипер-8 (U), «неизвестно»-29. Карточки трех первых категорий были использованы для построения формального разделяющего правила (*правило 1*).

Поскольку реальное использование отдельных переменных экспертами нам неизвестно, сравним распределения для каждой переменной у каждой пары классов. Для каждой пары классов выберем порог наилучшего разделения по критерию Смирнова и образуем симптомы для решения классификационной задачи в каждой паре классов в отдельности. Пример симптомов, различающих классы гиперкоагуляции и нормокоагуляции, сформированные по результатам экспертизы приведен в таб.14.

Таблица 14. Симптомы гиперкоагуляции по сравнению с нормой (NU).

@symptom 1	#1	mA <64
@symptom 2	#2	Рек < 120.5
@symptom 3	#3	Тол < 137
@symptom 4	#4	АТЗ < 86
@symptom 5	#5	Белмоч < 2.9
@symptom 6	#6	Эритр <102
@symptom 7	#7	Креа <1.45

Первые два столбца таблицы содержат заголовки (названия) симптомов, третий – условия. Поясним смысл переменных, упомянутых в условиях. Амплитуда тромбоэластограммы (mA), времена рекальцификации (Рек) и толерантности к гепарину (Тол), концентрация белка в моче (Белмоч), количество

эритроцитов (Эритр), уровень креатинина в крови (Креа). Положительные значения симптомов («ДА») соответствуют классу гиперкоагуляции по результатам экспертизы.

По построению, симптомы этого набора должны часто встречаться в векторах карточек, отнесенных экспертами к классу гиперкоагуляции, и редко – в векторах, отнесенных к классу нормокоагуляции. Идеально было бы, чтобы они отсутствовали во всех векторах второго из упомянутых классов.

Аналогично строятся наборы симптомов на парных сравнениях «гипо – гипер» и «гипо – норма».

В каждом парном сравнении участвующие классы разделяются полностью, а класс, не участвовавший в нем, дробится на подклассы, тяготеющие к участвующим классам. Иногда результат классификации оказывается неопределенным.

Правило 1 можно сформулировать кратко, как отнесение карточки к некоторому классу по совпадению частных решений, сравнивающих этот класс с каждым из остальных. Например, к классу U относим карточки получившие категорию U по обоим правилам DU и NU. Исключениями являются 8 карточек, отобранных в категории «ошибка» и «отказ».

Выделенные 8 карточек были представлены экспертам для обсуждения без указания причин предъявления. Экспертам было предложено обсудить эти карточки, мотивируя свои решения. В результате обсуждения изменена принадлежность 4 карточек - их решено перенести в категорию «неизвестно». Остальные 4 карточки сохранили свою принадлежность первоначальному классу (класс N – 1, класс D – 0, класс U – 3). Тем самым, класс N сохранил все карточки, и в нем осталось 5, класс D потерял 1 карточку и в нем осталось 12, а класс U потерял 3 карточки и в нем осталось 5. Эти числа зафиксированы в начале второго этапа работы.

Решающее правило на этапах 2 и 3 строилось по принципу «дифференциального диагноза», содержало 6 частных решений (вместо 3 на первом этапе) и основывалось на парном сравнении совокупности признаков по принципу разделения: «норма-гипо» (ND), «норма - гипер» (NU), «гипо- гипер» (DU), «норма - остальные» (N\*), «гипо-остальные» (D\*), «гипер - остальные» (U\*).

Были заново сформулированы симптомы для разделения каждой пары классов, только теперь уже было создано 6 наборов симптомов. Как и при выработке правила 1, были составлены таблицы значений симптомов и проведен подсчет баллов и относительной неразличимости.

Правило 3 использует показатели тромбоэластограммы, время рекальцификации и время толерантности к гепарину, концентрацию эндогенного гепарина, активность антитромбина III, максимальную амплитуду и наклон кривых агрегации тромбоцитов и эритроцитов.

Таким образом, после трехкратного повторения процедуры оценки экспертов практически стабилизировались и сблизились с формальной оценкой. Разумеется, факт совпадения формальной оценки с экспертной на классе обучения

не гарантирует пригодности формального правила для самостоятельной оценки больных. Оно основано на слишком малом массиве данных (55 карточек).

Правило 3 было использовано для классификации 115 карточек, не участвовавших в предшествующей части работы (чистый экзамен). Одновременно эти же карточки были предложены для классификации экспертам. Результат экспертной классификации считался действительным по большинству голосов экспертов (2 из 3). Результаты сравнения приведены в таб. 15.

Столбцы таблицы соответствуют градам формальной классификации, а строки – экспертной. Видно, что лишь 64 карточки из 115 отнесены экспертами и формальным правилом к одинаковым категориям. Обратим внимание на второй столбец. Категория «гипо» по формальной классификации целиком вошла в экспертную категорию «гипо». Лечение с применением антикоагулянтов вызывает у специалистов опасение получить гипокоагуляцию вследствие возможной ошибки, и они предъявляют повышенные требования к оценке ситуации.

Таблица 15. Сравнение формальной и экспертной классификаций на материале «чистого экзамена»

Оценка Экспертов	Формальная классификация				Сумма
	Норма	“гипо”	“гипер”	“неизвестно”	
Норма	6	0	0	0	6
Гипо	22	42	0	3	67
гипер	5	0	12	5	22
неизвестно	12	0	4	4	20
Сумма	45	42	16	12	115

Пройдя описанную стадию «знакомства» с новым для них контингентом больных, эксперты в дальнейшем ходе работы смогли обнаружить положительный эффект лечения гепарином и сформулировать отличия состояния гемостаза при длительной ремиссии от состояния у здоровых детей. Решающее правило 3 сконцентрировало параметры системы гемостаза и критические их значения, сыгравшие существенную роль в диагностике состояния больных. Заметим, что была применена формальная, универсальная процедура согласования мнений с параллельным построением решающего правила, ускорившая освоение экспертами специфики больных хроническим гломерулонефритом.

Рассмотрим теперь использование переменных для описания больных на различных этапах построения классификации больных (таб. 16).

Строки таблицы соответствуют 19 переменным, использованным для описания состояния больных (параметры тромбозаграммы R, k, (R+k), mA, времена рекальцификации крови (Рек) и толерантности к гепарину (Тлр), концентрации в крови эндогенного гепарина (Эгеп) и антитромбина 3 (Ат3), амплитуда (АТрМ) и наклон (АтрА) кривой агрегации тромбоцитов, амплитуда (АэрМ) и наклон (АэрА) кривой агрегации эритроцитов, время по тесту Квика (Квик), скорость фибринолиза (Фблз), количество эритроцитов (Эрит) и тром-

боцитов (Трбц) в крови больного, концентрации белка (Белок), креатинина (Креа) и кокарбоксилазы (Кок).

Таб. 16. Параметры коагулограммы, использованные при построении разделяющих правил и вошедшие в минимальное описание классов.

(N-норма, D-гипокоагуляция, U-гиперкоагуляция, \*-остальные).

	Прав.1				Правило 2						Правило 3						Мин. Описан		
	N D	U D	N U	N *	N D	U D	N U	U *	N *	D *	N D	U D	N U	U *	N *	D *	U	N	D
<b>R</b>	@	@		@	@	@	@	@		@		@	@	@					
<b>К</b>												@	@	@	@				
<b>R+k</b>												@	@	@					D
<b>МА</b>			@	@														N	
<b>Рек</b>	@	@	@	@	@	@				@	@	@		@	@	@			
<b>Тлр</b>	@	@	@	@	@	@	@		@	@	@	@	@	@	@	@	U	N	D
<b>Эгеп</b>	@	@		@		@				@		@	@	@		@	U		D
<b>АтЗ</b>	@	@	@	@	@				@	@	@				@	@			
<b>АтрМ</b>	@	@		@	@	@				@	@				@	@			D
<b>АтрА</b>											@				@				
<b>АэрМ</b>	@	@		@			@		@							@			
<b>АэрА</b>															@				
<b>Квик</b>				@		@		@											
<b>Фблз</b>	@	@							@										
<b>Эрит</b>	@		@	@			@		@										
<b>Трбц</b>		@																	
<b>Белок</b>		@	@																
<b>Креа</b>			@																
<b>Кок</b>		@																	

Группы столбцов соответствуют решающим правилам, использованным для аппроксимации выводов экспертов (правила 1, 2, 3). Столбцы внутри групп соответствуют частным решениям, полученным попарным сравнением классов (ND-сравнение нормы с гипокоагуляцией, NU- сравнение нормы с гиперкоагуляцией, N\*- сравнение нормы со всеми остальными, и т.д.).

Пустая клетка на пересечении строки и столбца означает, что данная переменная в рассматриваемом частном решении не участвует. Символ @ в клетке указывает на использование данной переменной в частном решении. Последние три столбца показывают на использование переменных в итоговом описании классов «Норма» (N), «Гипокоагуляция» (D) и «Гиперкоагуляция» (U).

Обращает на себя внимание систематическое уменьшение числа используемых переменных по мере выработки экспертами общей точки зрения. После первого обсуждения становятся ненужными количество тромбоцитов (Трбц),

белок (Белок), креатинин (Креа) и кокарбоксилаза (Кок). После второго – количество эритроцитов (Эрит), фибринолиз (Фблз), время Квика (Квик). В минимальное описание окончательных классов (не менее 85% объектов своего класса и не более 6% из совокупности остальных «известных» классов,  $p < 10^{-4}$ ) вошли лишь 2 параметра тромбоэластограммы (R+k, mA), 2 параметра состояния гепариновой системы (Тлр, Эгеп) и амплитуда (АТрМ) агрегации тромбоцитов (5 переменных из первоначальных 19).

Отметим, что эксперты не только пришли к общему мнению, но и сумели использовать для указания принадлежности коагулограммы определенному классу минимальное количество переменных.

**Глава 12** затрагивает трудный вопрос для медицинских исследований. Если есть отклонение состояния, выраженное достаточно слабо, как определить характер его связи со здоровьем пациентов? Сплошные обмеры, скрининговые исследования дают зашумленный материал. На примере обследования пострадавших при аварии на Чернобыльской АЭС предложены методы, ориентирующиеся на выявление максимально пострадавшего контингента.

Простейшим методом (*метод 1*) представления и визуального сравнения данных может служить отображение распределения значений наблюдаемой переменной в виде диаграммы значений в последовательные годы (в сравнении с нормативами для здоровых людей). Среди жителей пострадавших районов иногда можно наблюдать систематическое отклонение вверх или вниз отдельных количественных показателей. Изображенные в виде диаграммы измененных значений, они удобны для изучения специфики территорий с разными уровнями радиационного загрязнения. Рассмотрим, например, такую диаграмму для Р – белка (рис. 7).

На рисунке представлены 4 диаграммы долей низких, нормальных и высоких значений Р–белка в популяциях жителей 4 регионов с различными уровнями загрязнения. На каждой диаграмме по оси абсцисс отложены годы наблюдения (с 1986 по 1993). Пустыми оставлены годы, когда экспедиции в эти регионы не посылали.

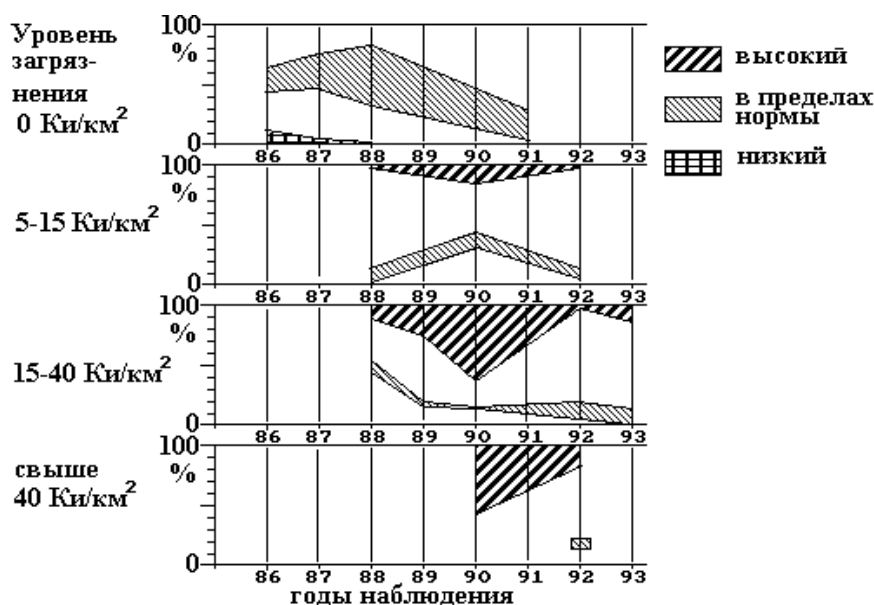


Рис. 7. Динамика Р – белков по годам наблюдения

По оси ординат отложены доли низких, нормальных и высоких значений концентрации Р–белка в крови пациента. Интервалы значений концентрации не

соприкасаются, т.е. есть промежуточные значения между нормальными и низкими и между нормальными и высокими. Для каждого года вертикальный отрезок делится на 5 частей пропорционально доле наблюдений каждого интервала (очень малые значения, умеренно малые, в пределах нормы, умеренно большие, очень большие). На диаграмме окрашены только «нечетные» диапазоны (норма и «крайние»). Следовательно, более широкая область соответствует преобладанию представителей соответствующего интервала. Например, из диаграмм видно появление заметного количества высоких значений после 1988 года, причем в районах большего загрязнения их доля выше. Такая закономерность поддерживает гипотезу о повышении концентрации Р-белка как возможной реакции на загрязнение.

Еще один метод анализа (*метод 2*) состоит в отсечении слабо измененного большинства, предположительно пострадавших в наименьшей степени. Для этого подвергается анализу динамика "хвостов" распределений показателей по годам наблюдения. Предполагая наличие достаточно мощного скрытого фактора, например, интенсивности индивидуального радиоактивного повреждения, можно искать одинаковые тенденции изменения среди соответствующих наборов показателей. В дальнейшем эти наборы можно подвергнуть более тщательному исследованию.

Распределения иммунологических параметров в пострадавших регионах слабо отличаются от распределений в регионе сравнения (регион 0). «Хвосты» выделяли по границе распределения в регионе 0. Увеличение порога, уводящее границу диапазона значений переменной в сторону «хвоста» распределения, уменьшает количество отобранных наблюдений, но увеличивает относительное «участие» более тяжелого региона.

Метод имеет качественный характер, его цель – обратить внимание исследователя на корреляции, не выявляющиеся при более простых методах обработки. В силу достаточно произвольного выбора порогов получение количественных оценок затруднено.

Метод анализа распределений (*метод 3*) уточняет результаты анализа измененных значений и вводит количественную меру отклонения показателя в группе.

Он использует результаты выделения верхнего и нижнего краев распределений для нескольких показателей. Заметим, что пороги для различных показателей можно выбирать независимо, но они должны быть едиными для всех групп пациентов.

Если какой-то показатель превышает верхнюю границу диапазона нормативных значений, говорят о напряжении показателя, если оказывается ниже нижней границы - об угнетении. Распределения количества напряженных (*parp*) и угнетенных (*ugnet*) показателей демонстрируют отличие регионов друг от друга (рис. 8.).

На рисунке представлены распределения числа напряженных (верхний график) и угнетенных (нижний график) параметров у жителей регионов с разными плотностями загрязнения (0 – нет, 1 – от 5 до 15 Ки/км<sup>2</sup>, 2 – от 15 до 40

Ки/км<sup>2</sup>, 3 – более 40 Ки/км<sup>2</sup>). По оси абсцисс отложены числа напряженных (соотв. угнетенных) показателей из 11, по оси ординат – процент обследованных жителей региона, имеющих соответствующее количество измененных параметров иммунитета.

Для нулевого региона характерны большое количество (48%) жительниц, не имеющих ни одного напряженного показателя, и очень малая доля женщин с тремя и более напряженными показателями. По мере перехода к регионам более

сильного поражения (1-2-3), доля женщин с 3 и более напряженными (и угнетенными) показателями систематически растет. Это указывает на корреляцию количества напряженных показателей с уровнем загрязнения в регионе.

Более детальные сведения позволяет получить четвертый из предлагаемых методов (*метод 4*) - анализ корреляционной связи количественных показателей с числом повышенных или пониженных значений.

Рассмотрим систему  $n$  центров, продуцирующую  $n$  веществ, причем каждый центр продуцирует одно вещество и состоит из  $m$  элементов, в принципе подверженных воздействию радиации. Пусть доля  $p$  поврежденных элементов достаточно мала ( $p \ll 1$ ). Количество поврежденных элементов в одном центре, следовательно, составит  $pm$ . Будем также считать, что продукция каждого вещества изменяется (например, увеличивается) пропорционально числу поврежденных элементов в центре, производящем это вещество. Назовем состояние

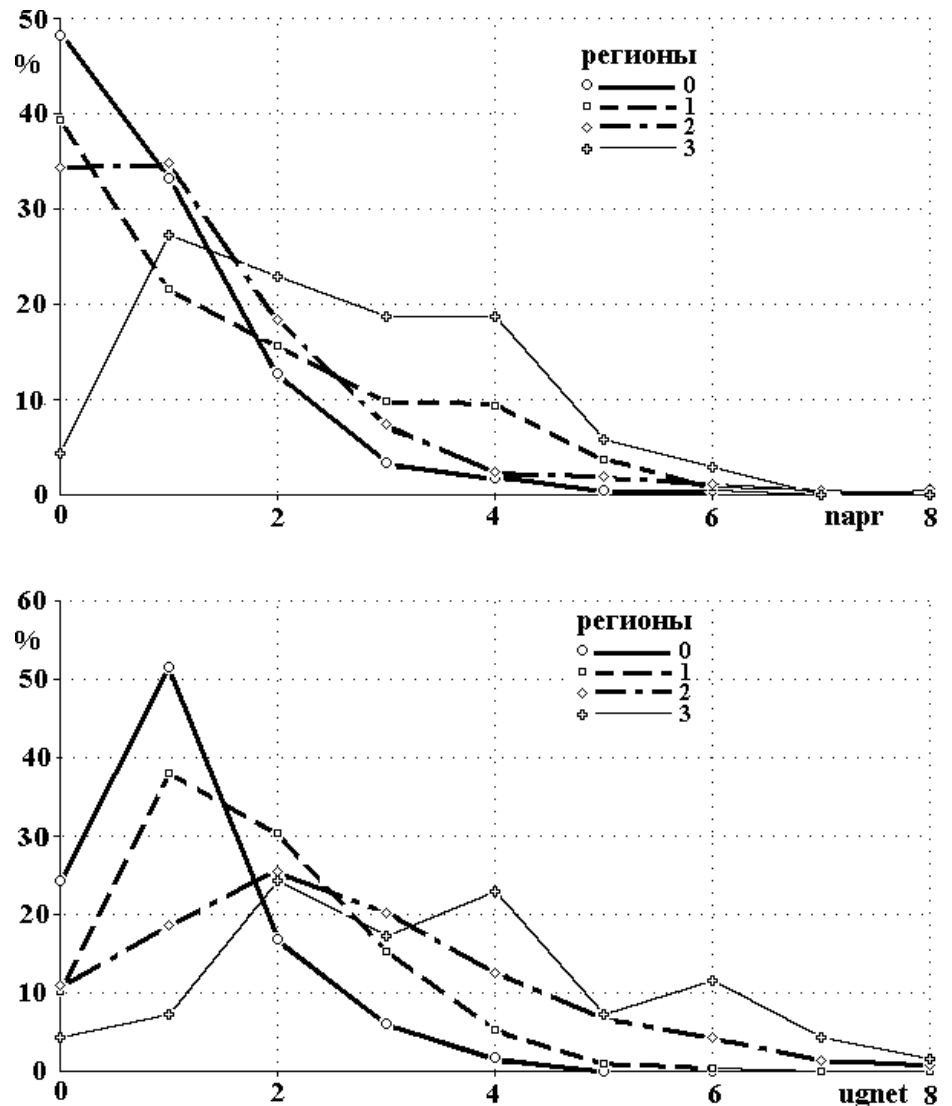


Рис. 8. Распределения числа напряженных и угнетенных показателей.

центра напряженным, если в нем повреждено более  $z$  элементов ( $z < m$ ). Оценим число напряженных центров при заданной доле повреждений.

Очевидно, что при одной и той же доле поврежденных  $p$  число центров в напряженном состоянии будет зависеть от конфигурации повреждений, т.е. от распределения поврежденных элементов по центрам. Наибольшее количество напряженных центров окажется, если в максимально возможное количество центров попадет по  $(z + 1)$  повреждению, т.е. оно равно

$$l_{\max}(p) = \frac{pm}{z + 1}. \quad (8)$$

Наименьшее количество – если максимальное количество центров получит  $z$  повреждений (ниже порога), а остальные повреждения будут сконцентрированы в минимальном количестве центров

$$l_{\min}(p) = \begin{cases} 0, & p < z/m \\ \frac{pm - z}{m - z}, & p \geq z/m \end{cases}. \quad (9)$$

Поскольку всегда  $l_{\min}(p) \leq l(p) \leq l_{\max}(p)$ , обе оценки – неубывающие функции  $p$ , а зависимость  $l(p)$  лежит между ними, область ее возможных значений будет смещаться в сторону увеличения значений по мере роста  $p$ .

Поэтому увеличение доли поврежденных элементов должно в большинстве случаев сопровождаться ростом числа повышенных значений, уменьшение – пониженных. В обоих случаях должен возникать сдвиг распределения наблюдаемых значений переменной. Группы, неоднородные по отклику, могут демонстрировать одновременно оба эффекта в соответствующих подгруппах. Одновременно, увеличение доли поврежденных элементов должно вести к увеличению числа разнородных центров, находящихся в напряженном состоянии, т.е. к увеличению количества переменных с измененными значениями. Пример диаграммы, демонстрирующей корреляционную связь распределения количественной переменной с числом напряженных переменных на примере уровня иммуноглобулина А в крови женщин, приведен на рис. 9.

На рисунке представлены два блока диаграмм, относящихся к региону 2 (15-40 Ки/км<sup>2</sup>) и региону 1 (5-15 Ки/км<sup>2</sup>). Внутри блока диаграммы сгруппированы по годам, по 4 строки на каждый год.

Строка представляет собой диаграмму статистики по соответствующему классу: звездочка отмечает медиану распределения, а линия указывает квартильный интервал. Слева от строки отмечено число напряженных переменных, справа – численность выборки.

По значениям медиан одного года проведена регрессионная прямая для наглядной демонстрации наблюдаемых тенденций. Видно, что в регионе 2 в 1988, 1989 и 1992 годах проявилась связь уровня иммуноглобулина IgA с числом напряженных переменных, а в остальные годы в этом регионе и всегда в регионе 1 такую связь не наблюдали. Для 1993 г нецелесообразно рассматривать такую связь из-за крайне малого числа наблюдений.

По-видимому, изменения затрагивают систему выработки иммуноглобулина «А» только при достаточно интенсивном воздействии загрязнений и, возможно, при определенном их характере, который особенно сильно проявился в 1988-89 годах. Именно для этих лет следует использовать уровень IgA как характеристику степени поражения отдельного человека.

Совокупность таких переменных позволяет отобрать для каждого периода времени и региона жителей, пострадавших от загрязнения в наибольшей степени, и, используя результаты измерения различных их параметров, получить описание наиболее характерных процессов в организме человека.

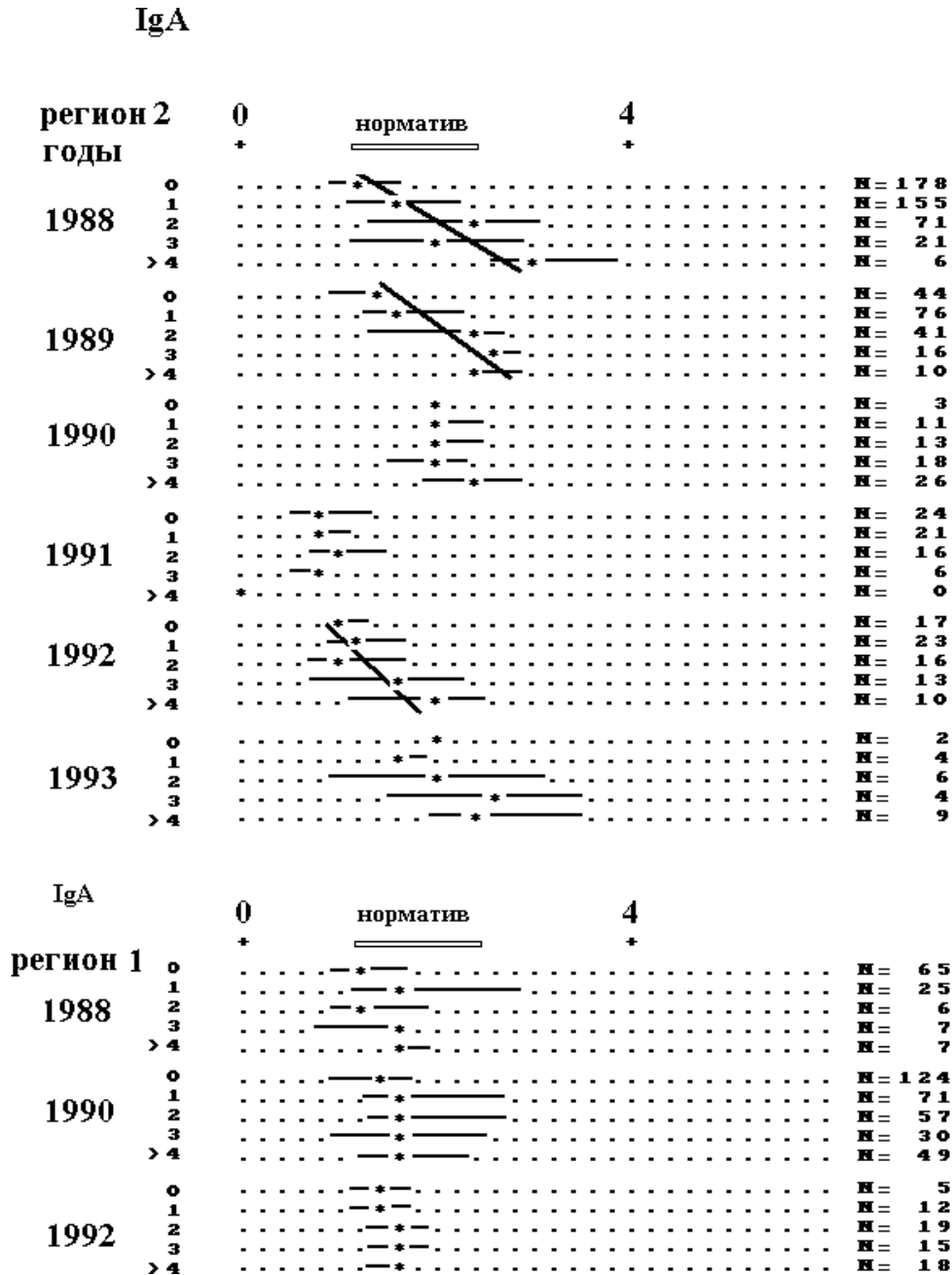


Рис. 9. Связь уровня иммуноглобулина А с количеством напряженных показателей у жителей двух регионов.

Коэффициент корреляции проведенной прямой может служить ориентировочной характеристикой степени проявления связи переменной для определенного года с числом напряженных показателей.

Использована простейшая модель корреляции наблюдаемых признаков с тяжестью поражения организма. Исследование совместных распределений разных характеристик тяжести поражения дает набор наиболее чувствительных показателей и одновременно наиболее отклоняющийся класс.

Сравнение этого контингента с остальными (условно «непострадавшими») позволяет выявить эффект воздействия.

**В главе 13** методы, разработанные для медицинских задач, использованы при решении геофизических задач анализа крайне низкочастотных электромагнитных сигналов от ядерных взрывов.

Рассмотрен случай анализа процессов при наличии теоретической модели. Аппроксимационная модель наблюдаемых записей процесса, использующая частные особенности, почерпнутые из теоретической модели, позволяет оценить параметры объекта. Рассмотрение проведено на двух физических задачах оценки регистрируемых сигналов при ядерных взрывах.

Рассмотрены две задачи: ионосферный и подземный ядерные взрывы. В обоих случаях существуют удовлетворительные приближенные аналитические модели генерации геомагнитных возмущений (ГМВ). Нужда в моделировании процессов возникла при сопоставлении с результатами экспериментов.

1. В теоретических решениях присутствует большое количество характеристик условий проведения экспериментов (ионосферные проводимости, свойства грунта, и т.п.). Многие из этих констант известны, в лучшем случае, грубо приближенно, некоторые лишь по порядку величины, между тем их влияние на результат вполне ощутимо. Это обстоятельство крайне затрудняет корректную интерпретацию результатов детального численного моделирования процессов методом сеток.
2. Во время самого взрыва и процессов, порожденных взрывом, многие свойства среды существенно меняются, и не всегда известно, насколько.
3. Границы слоев в теоретической модели выбраны резкими, ради возможности доведения расчетов до конечных результатов. Это предположение в реальном эксперименте не соблюдено. И даже по одной этой причине интересно провести оценки свойств среды по результатам экспериментов.

Для проверки основных результатов теоретической модели мы провели математическую обработку записей ГМВ от взрыва "Старфиш". Целью анализа экспериментальных данных был поиск соотношений, связывающих частоты и времена релаксации ГМВ с характеристиками гиротропной среды. В качестве экспериментальных данных были использованы 5 записей ГМВ, содержащих короткопериодные колебания.

В выбранных записях исследовали начальный отрезок сигнала общей длительностью около 20-30 с (см., например, рис. 10а), соответствующий короткопериодным колебаниям.

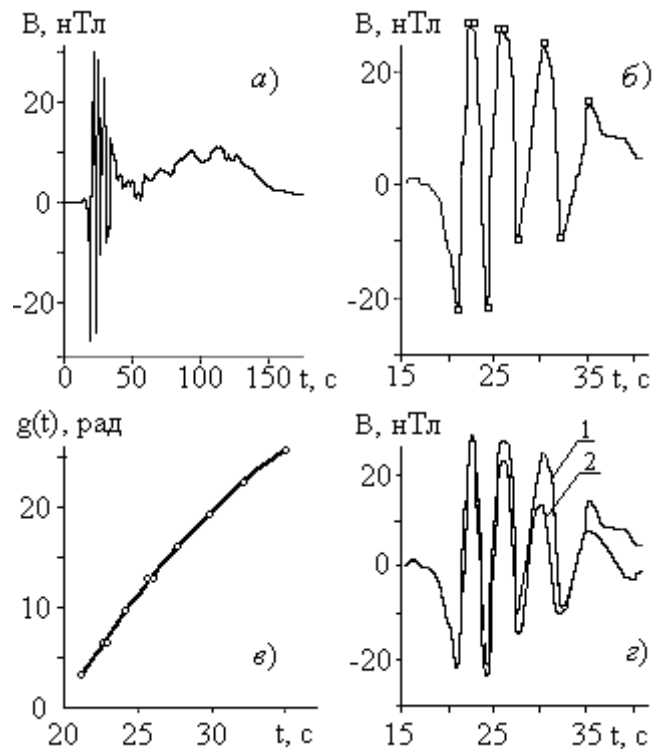
Предлагаемый метод анализа использует модель осциллирующего сигнала с плавным изменением амплитуды и длительности последовательных отклонений:

$$f(t) = A(t) * \cos(g(t)) + x(t) \quad (10)$$

где  $A(t)$  и  $g(t)$  - непрерывные функции времени  $t$ , а  $x(t)$  - шумовое слагаемое. Фазу  $g(t)$  будем считать монотонно неубывающей, а функцию  $A(t)$  (амплитуду сигнала) - монотонно невозрастающей. Эта модель охватывает широкий класс процессов.

Аппроксимация реальной записи сигнала ГМВ в рамках модели (10) сводится к подбору функций  $A(t)$  и  $g(t)$ . Рассмотрим сначала вторую из них. Влияние шумов и неизвестный вид  $A(t)$  затрудняют поточечное восстановление функции  $g(t)$  по записи сигнала. Относительно надежно удастся идентифицировать, особенно при довольно высоком уровне шумов, лишь окрестности экстремумов.

Рис. 10. Запись геомагнитного возмущения в Пало-Альто: *а* - полная запись сигнала; *б* - короткопериодная часть записи с точками, расставленными в области экстремумов; *в* - график зависимости фазы от времени; *г* - результат аппроксимации короткопериодной части; 1 - эксперимент, 2 - аппроксимация



При высокой зашумленности они могут оказаться отрезками или наборами разрозненных точек. Каждой такой точке (отрезку, набору) можно условно приписать

значение фазы:  $0, 2\pi, \dots, 2k\pi$  (максимумы) или  $\pi, 3\pi, \dots, (2k+1)\pi$  (минимумы). Если шум не слишком велик и отсутствует медленный дрейф нулевой линии, можно использовать "нули" графика (точки, отрезки, наборы точек), приписывая им  $\pi/2, 3\pi/2, \dots, (2k+1)\pi/2$ . Построенную таким способом эмпирическую временную зависимость фазы  $g(t)$  аппроксимируем подходящей гладкой кривой в соответствии с предполагаемой моделью сигнала.

Построение функции амплитуды  $A(t)$  выполняется обычным способом - по абсолютным значениям отклонений в экстремальных точках сигнала, уточненных с помощью  $g(t)$ .

Рассмотрим метод аппроксимации на примере сигнала (рис. 10а), записанного рубидиевым магнитометром на станции Пало-Альто при взрыве "Старфиш". Запись содержит относительно "высокочастотное" (визуально) начало за

которым следует серия низкоамплитудных колебаний увеличивающейся длительности. Во второй части сигнала отчетливо видно медленное "дрейфовое" движение длительностью около 100 с. Применим описанный выше метод к анализу начальной части сигнала.

Выделим максимумы и минимумы с помощью программы поиска экстремумов оцифрованной кривой (рис. 10б). Последовательным экстремумам, начиная с первого максимума, поставим в соответствие значения фазы  $0, \pi, 2\pi, \dots, 8\pi$ . Эмпирические значения фазы в соответствующие моменты времени показаны точками на рис. 10в. Видно, что быстрый первоначальный рост фазы замедляется к концу исследуемого сегмента сигнала. По точкам рис. 10в подбирали различные аппроксимирующие функции и сравнивали коэффициенты корреляции  $r$ . Наилучшая аппроксимация (с наибольшим коэффициентом корреляции) зависимости фазы от времени на всем протяжении сигнала достигается при использовании гиперболы.

Соответствие экспериментальных точек гиперболической аппроксимации  $g(t) = a/(t-t_0) + g_0$  для всех исследованных записей показывает рис. 11. На нем приведены графики зависимости  $\ln(g_0 - g(t))$  от  $\ln(t-t_0)$ . Точки всех записей ложатся на прямые единичного отрицательного наклона, демонстрируя хорошее подтверждение гиперболической аппроксимации. Прямые 3 и 4, полученные для станций Джакамарка и Пало-Альто соответственно, слились в единую прямую.

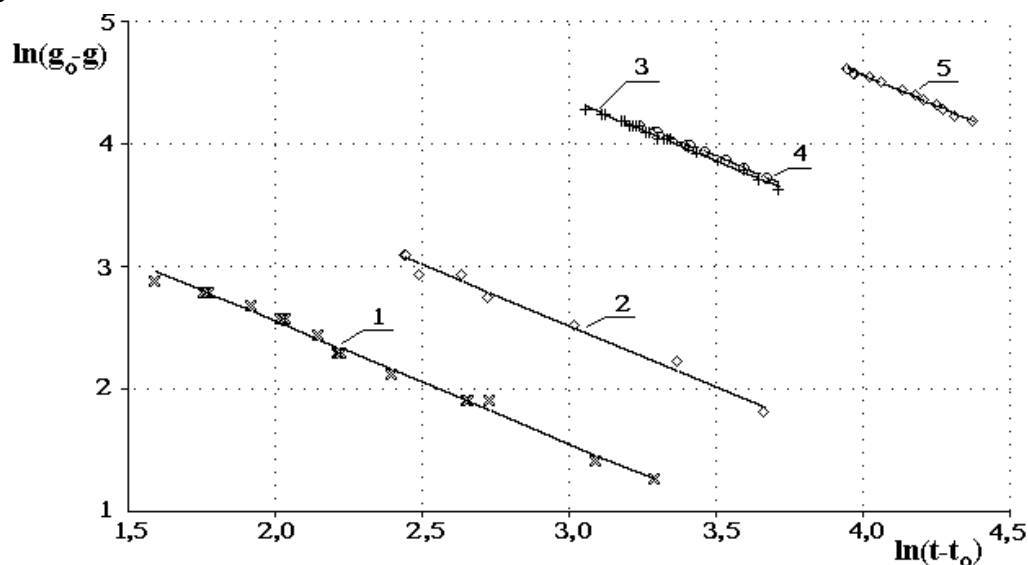


Рис. 11. Зависимость натурального логарифма фазы от натурального логарифма времени для пяти станций: 1- Пуако, 2- Тутуила, 3- Джакамарка, 4- Пало-Альто, 5- Уанкайо

Обнаруженная гиперболическая зависимость фазы от времени соответствует теоретическим выражениям для компонент токов  $j_\phi$ ,  $j_r$  в начальной стадии диффузии.

$$\begin{aligned}
 j_\varphi &= rM_1 \sqrt{\frac{\mu_0 \sigma_2}{\pi t}} \exp\left(-\frac{\mu_0 \sigma_1 z^2}{4t}\right) \cos\left(\frac{\mu_0 \sigma_2 z^2}{4t} - \frac{\varphi}{2}\right), \\
 j_r &= rM_1 \sqrt{\frac{\mu_0 \sigma_2}{\pi t}} \exp\left(-\frac{\mu_0 \sigma_1 z^2}{4t}\right) \sin\left(\frac{\mu_0 \sigma_2 z^2}{4t} - \frac{\varphi}{2}\right).
 \end{aligned} \tag{11}$$

Здесь  $\operatorname{tg}\varphi = \sigma_2/\sigma_1$ , где  $\sigma_1$  и  $\sigma_2$  - компоненты тензора проводимости,  $M_1$  - амплитуда токов;  $r, \varphi, z$  - цилиндрические координаты,  $\mu_0$  - магнитная постоянная.

Построена квазиколлебательная аппроксимационная модель сигнала. Показано удовлетворительное соответствие ее вариантам записанных сигналов. Применение модели к записям сигналов, выполненным в различных регионах Земли, позволило установить однотипный характер сигнала, не зависящий от расстояния и геомагнитных особенностей пункта наблюдения. Подтвержден гипотетический механизм генерации ГМВ при надионосферных взрывах.

Полученные значения констант позволяют оценить усредненные значения электромагнитных констант ионосферы (проводимости Холла  $\sigma_2$  и Педерсена  $\sigma_1$ ). Обычные методы дают лишь ориентировочные значения этих констант.

Вторая часть главы посвящена случаю геомагнитного сигнала от подземного ядерного взрыва без выхода продуктов взрыва на поверхность (камуфлетный взрыв). В этом случае сигнал имеет вид большого импульса приблизительно «треугольной» формы с крутым передним и пологим задним фронтами, иногда с еще одной волной существенно меньшей амплитуды. Для сигнала от подземного ядерного взрыва роль аппроксимационной модели оказалась даже больше, чем в предыдущем случае.

Теоретический расчет дает для азимутальной компоненты поля

$$E_\varphi = \sin\theta \cdot E(r, t), \quad \text{где}$$

$$\begin{aligned}
 E(r, t) &= \frac{M_0}{8\pi^{3/2} \sigma r^4} \times \\
 &\times \left\{ \lambda^{5/2} \exp(-\lambda/4) - 4\lambda_0 \left( \sqrt{\lambda} \exp(-\lambda/4) + \sqrt{\pi} \left[ 1 - \Phi\left(\sqrt{\lambda/2}\right) \right] \right) \right\} \\
 &\text{при } \lambda \geq \lambda_0 \quad (0 \leq t \leq \tau); \\
 E(r, t) &= \frac{M_0}{8\pi^{3/2} \sigma r^4} \left\{ \lambda^{5/2} \exp(-\lambda/4) - 4\lambda_0 \left( \sqrt{\lambda} \exp(-\lambda/4) - \right. \right. \\
 &\left. \left. - \frac{\lambda_0^{1/2}}{(t/\tau - 1)^{1/2}} \exp\left[-\frac{\lambda_0}{4(t/\tau - 1)}\right] + \sqrt{\pi} \left[ \Phi\left(\sqrt{\frac{\lambda_0}{2(t/\tau - 1)}}\right) - \Phi\left(\sqrt{\frac{\lambda}{2}}\right) \right] \right) \right\} \\
 &\text{при } 0 \leq \lambda \leq \lambda_0 \quad (t \geq \tau).
 \end{aligned} \tag{12}$$

Здесь  $t/\tau = \lambda_0/\lambda$ ,  $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$  и введены обозначения

$$\lambda = \mu_0 \sigma r^2 / t, \\ \lambda_0 = \mu_0 \sigma r^2 / \tau,$$

$r$  - расстояние от места взрыва до точки наблюдения,  $\mu_0$  - магнитная постоянная,  $M_0$  - амплитуда возмущения в точке взрыва,  $\sigma$  - скалярная проводимость грунта. Ось  $Z$  сферической системы координат  $r, \phi, \theta$  направлена вдоль вектора  $B$ , начало координат помещено в центр взрыва.

В результате поисков однопараметрической аппроксимации оказалось удобным использовать монотонную функцию, представляющую собой интеграл от функции  $E_\phi(t)$ . Для того, чтобы избавиться от ошибок калибровки по амплитуде на разных станциях, интеграл был нормирован на единицу. Построим такой интеграл для теоретической функции  $E_\phi(t)$ :

$$s(t, \lambda_0) = \frac{1}{s_0(t_1, \lambda_0)} \int_{t_0}^t f_T(\alpha, \lambda_0) d\alpha \quad (13)$$

где  $f_T(\alpha, \lambda_0)$  - теоретическая функция при фиксированном  $\lambda_0$ ,

$$s_0(t_1, \lambda_0) = \int_{t_0}^{t_1} f_T(\alpha, \lambda_0) d\alpha, \quad (t_0 \leq t \leq t_1),$$

$s(t_1, \lambda_0) = 1$ ,  $t_0$  - момент начала сигнала,  $t_1$  - момент времени, когда функция  $f_m$  в первый раз обращается в нуль. Время также приводим к единичному интервалу, разделив его на  $(t_1 - t_0)$ .

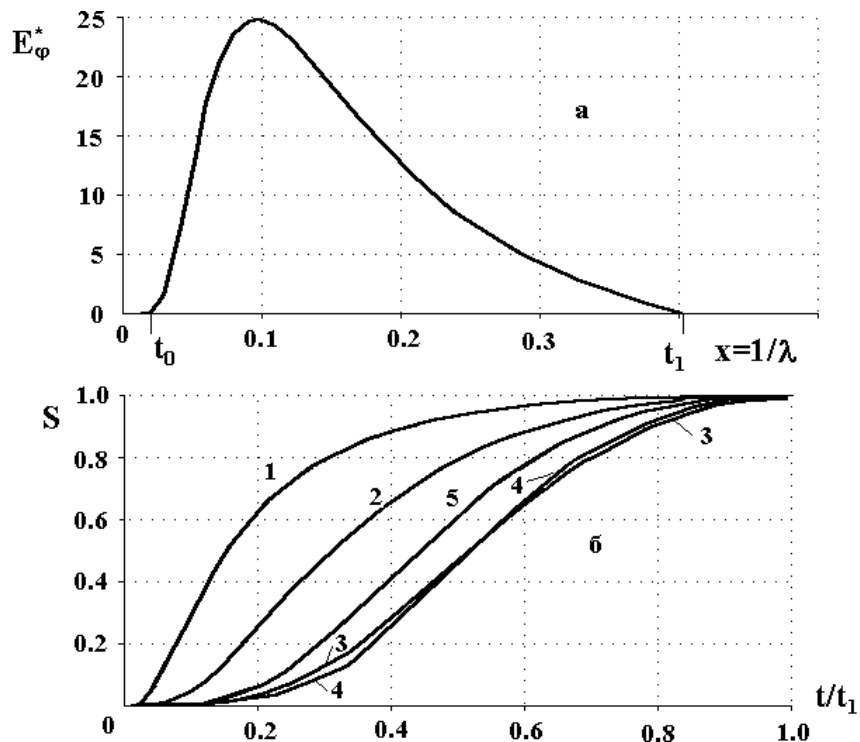


Рис. 12. Начальный “полупериод” геомагнитного возмущения: а) - теоретическая кривая для  $\lambda_0 = 1$ ; б) - зависимость нормированной площади от времени: 1 -  $\lambda_0 = 0,1$ ; 2 -  $\lambda_0 = 1$ ; 3 -  $\lambda_0 = 10$ ; 4 -  $\lambda_0 = 20$ ; 5 - 100

На рис. 12а представлен первый "полупериод" теоретического сигнала при  $\lambda_0 = 1$ , а на рис. 12б изображены графики зависимости интеграла  $s(t, \lambda_0)$  от времени для нескольких значений  $\lambda_0$ .

По мере увеличения параметра  $\lambda_0$  вплоть до значений  $\lambda_0 = 10$  каждая следующая кривая лежит ниже предыдущей. При  $\lambda_0 = 20$  понижение прекращается, кривые начинают пересекаться, а при последующем увеличении  $\lambda_0$  кривые  $s(t, \lambda_0)$  снова поднимаются (кривая 5 рис. 12б). Следовательно, монотонность зависимости  $s(\lambda_0)$  сохраняется лишь в области значений  $\lambda_0 < 20$ . Далее мы ограничимся только этой областью значений  $\lambda_0$ .

Поиск подходящих переменных для однопараметрической аппроксимации привел нас к линейной функции вида

$$\ln(-\ln(s)) = b + u \ln(-\ln(q)) , \quad (14)$$

где  $q = (t-t_0)/(t_1-t_0)$  с константами  $b$  и  $u$ .

Применим функцию (14) для исследования теоретически рассчитанных кривых (12). Для каждой кривой (12), соответствующей фиксированному значению  $\lambda_0$ , было проведено интегрирование в пределах первой "полуволны" типа рис. 12а согласно формуле (13) и получена зависимость  $s(q, \lambda_0)$ . Затем были рассчитаны двойные логарифмы величин  $s$  и  $q$  и построена по ним линейная регрессия.

На рис. 13 точками представлены результаты описанной процедуры, а сплошными линиями проведены регрессионные прямые. Кривая 2 построена для сигнала ГМВ, изображенного на рис. 12а при  $\lambda_0 = 1$ . Кривая 1 построена аналогичным образом путем интегрирования начальной стадии теоретической кривой (12) при  $\lambda_0 = 0,1$ , а кривая 3 - при  $\lambda_0 = 10$ .

Из рис. 13 видно, что регрессионные прямые очень хорошо ложатся практически на все рассчитанные точки. Таким образом, каждой кривой (12), т.е. каждому значению  $\lambda_0$  может быть поставлена в соответствие функция вида (14). Исследование показывает, что с изменением  $\lambda_0$  на два порядка (от 0,1 до 10) очень мало меняется константа  $u$  и гораздо заметнее – константа  $\beta$ . Поэтому далее мы исследуем только связь между  $\beta$  и  $\lambda_0$ .

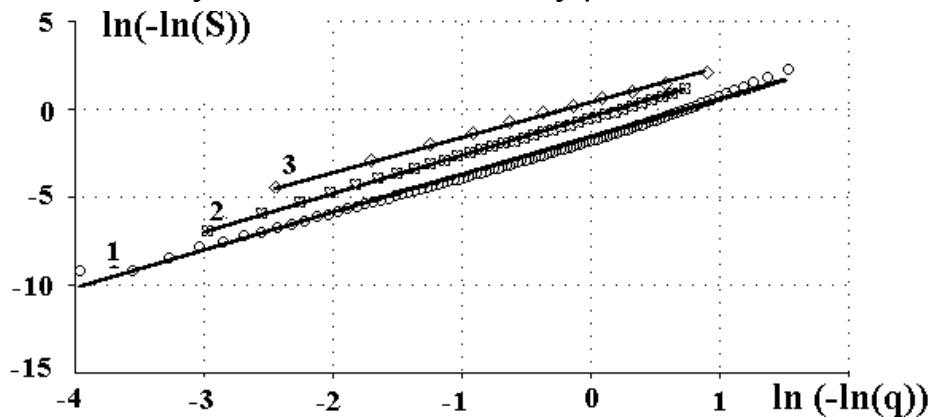


Рис. 13. Зависимость двойного логарифма площади под кривой от двойного логарифма времени

Варьируя параметр  $\lambda_0$  и используя соотношение (14) для теоретической модели (12), получаем зависимость между константой  $\beta$  и параметром  $\lambda_0$ :

$$\beta = 0.545 \ln(\lambda_0) - 0.438 \quad (15)$$

Хорошая линейная зависимость двойного логарифма площади от двойного логарифма времени заставила нас проверить, не присуща ли подобная связь между этими величинами и экспериментальным записям. Такая проверка была проведена на эксперименте Билби (штат Невада, 1963). Начальная стадия сигнала ГМВ для Билби представлена на рис. 14а для азимутальной (кривая 1) и радиальной (кривая 2) компонент соответственно.

На рис. 14б представлены полученные линейные регрессии по записям азимутальной (кривая 1) и радиальной (кривая 2) компонент. Видно, что все точки очень хорошо лежат на прямую линию. Кривая 2 нанесена пунктиром и почти на всем своем протяжении совпадает с кривой 1.

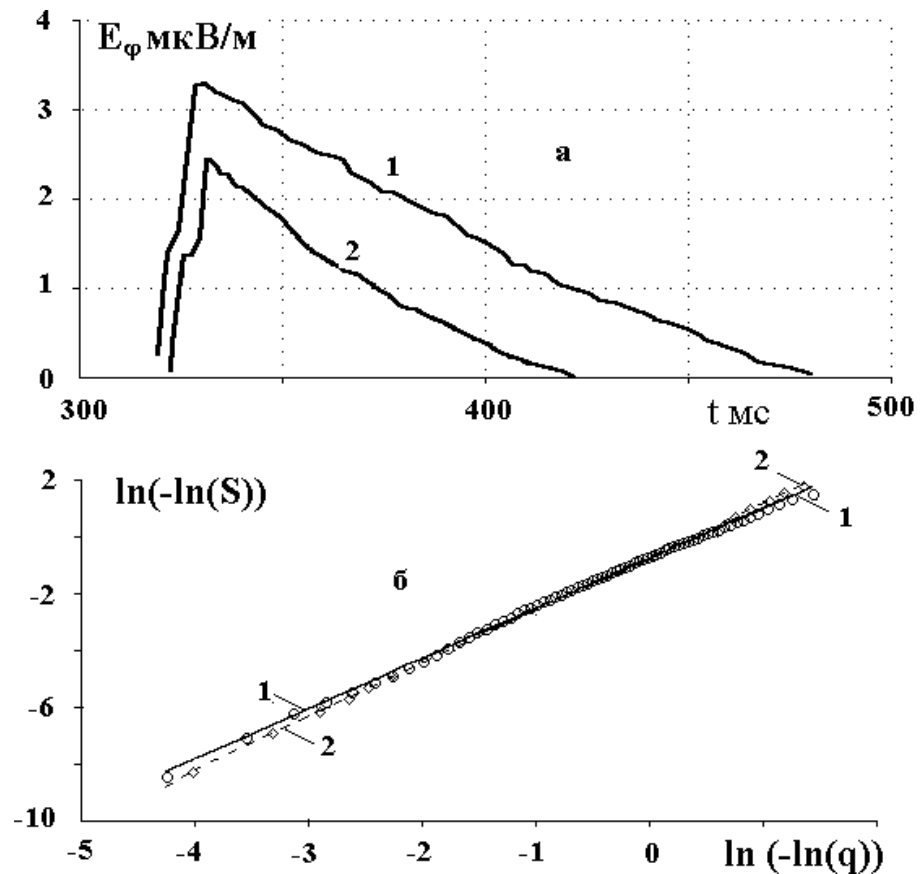


Рис. 14. Экспериментальный сигнал и его обработка: а) - начальная стадия геомагнитного возмущения от взрыва Билби; б) - зависимость двойного логарифма площади от двойного логарифма времени для Билби; 1 - азимутальная компонента, 2 - радиальная

Итак, построена формальная модель скользящего норматива процесса. На основе модели выявлены информативные параметры сигнала. Получена связь этих параметров с мощностью взрыва (рис.15).

Введено преобразование (13), позволяющее превратить график зарегистрированного импульса (первого отклонения магнитограммы) в монотонную нормированную кривую. С помощью этого преобразования выявлены свойства сигнала, получаемого на основе теоретической модели.

Преобразованный сигнал в широкой области параметров можно аппроксимировать соотношением (14).

Для практического использования написана программа автоматизированной оценки мощности камуфлетного взрыва по первой полуволне магнитограммы. Программа получает на вход оцифрованную магнитограмму и, в зависимости от того, какая компонента (радиальная или азимутальная) использована, выдает оценку мощности. При обработке магнитограммы используется вся кривая, а не отдельные ее точки. Это позволяет заметно повысить точность оценки параметров сигнала.

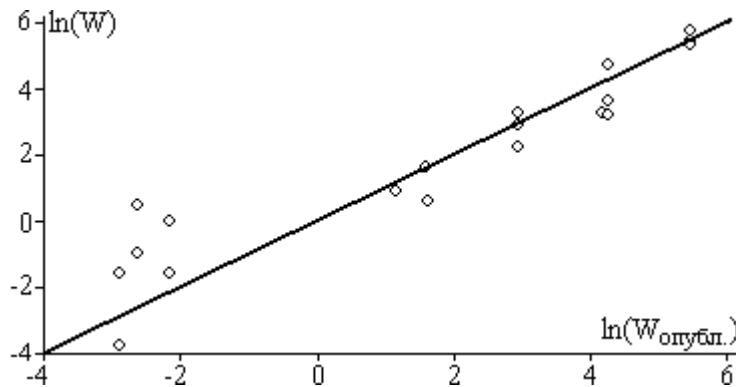


Рис. 15. Сравнение рассчитанной мощности с опубликованной

Разработанная теория описывает только азимутальную компоненту электрического поля  $E_\phi$  в системе координат, связанной с локальным направлением магнитного поля Земли. В высоких широтах это направление близко к вертикали, поэтому  $E_\phi$  воспринимается обычным горизонтальным датчиком. Для оценки радиальной компоненты  $E_r$  требуются сведения о пространственной ориентации магнитного поля и анизотропии характеристик горных пород в точке взрыва. То, что радиальная компонента поля ведет себя аналогично азимутальной, является интересным наблюдением. Примененный метод обработки позволил выявить эту аналогию.

Разумеется, малое количество опубликованных кривых затрудняет уточнение полученных оценок мощности взрыва по геомагнитному сигналу. Тем не менее, полученные оценки максимально используют информацию, заключенную в геомагнитном сигнале. Получение дополнительных наблюдений взрывов должно еще улучшить результат.

Решение этой задачи приобретает известное значение в связи с проблемой сокрытия проводимых ядерных испытаний. Главным методом контроля ядерных испытаний остается сейсмический. К сожалению, при энергиях взрыва с тротильным эквивалентом 1-10 кт сейсмические сигналы ядерных взрывов не-

отличимы от сигналов химических взрывов аналогичной мощности. Надежным отличием ядерного взрыва от химического является геомагнитный сигнал специфической формы. Геомагнитный сигнал приходит в пункт наблюдения на несколько секунд раньше сейсмического. Это позволяет уверенно регистрировать геомагнитный сигнал и, как было показано выше, по его форме оценивать мощность произведенного взрыва.

Эта глава иллюстрирует возможное использование разработанных методов и программ в другой предметной области естествознания.

### **Результаты.**

1. В диссертации разработаны формализованные методы структуризации и вербализации интуитивного знания. Методы были развиты в процессе анализа нескольких видов лечебной деятельности врача.
2. Построены модели профессиональных решений на основе языка логических симптомов. Создан также язык сценариев, который представляет процесс, развивающийся во времени, в виде системы фактов, принадлежащих определенным моментам времени и составляющих события. Созданы алгоритмы, выявляющие структуру реальных действий специалиста. Разработаны программы, составляющие технологическую базу анализа наблюдений и заключений врача.
3. На основе разработанных методов формализации получены решения 7 практических задач медицинской диагностики. Модели профессионального знания в этих задачах учитывают особенности профессиональных групп врачей (хирурги, акушеры, педиатры, и др.). Решение задачи формулируется для врача в виде небольшого набора рекомендаций, ориентированных на особенности состояния конкретного больного.
4. Созданы методы формализации профессионального знания, являющиеся основой алгоритма организации работы группы экспертов над новой проблемой.

### **Внедрение.**

Результаты диссертации внедрены в клиническую и научную практику в Московском областном научно-исследовательском клиническом институте (МОНИКИ), в Московском НИИ акушерства и гинекологии (МОНИИАГ), в 23-м роддоме, во 2 ММИ, в МИФИ. Программные инструменты, разработанные для реализации описанных методов, используются также в НИИ молекулярной биологии им. А.Н. Белозерского при МГУ, в Российском научном центре хирургии, в ИПУ РАН.

### Основные публикации.

1. Алексеевский А.В., Гришин В.Л., Клименко П.А., Котов Ю.Б., Лукашенко С.Ю., Сахнина Э.И., Сичинава Л.Г. Федорова М.В. Структура клинических данных в проблеме диагностики и лечения плацентарной недостаточности. // Препринт АН СССР, Научный совет по комплексной проблеме «Кибернетика», – М.,1988, 35 с.
2. Горбачев Л.П., Глушков А.И., Котов Ю.Б., Семенова Т.А., Скрыльник А.А. Оценка мощности камуфлетного взрыва по начальной стадии геомагнитного возмущения. // Геомагнетизм и аэрономия.1999, т.39, N5, с.77-82.
3. Горбачев Л.П., Котов Ю.Б., Семенова Т.А. Анализ записей геомагнитных возмущений (на примере эксперимента "Старфиш"). // Геомагнетизм и аэрономия, 1998, т.38, N 2, с.85-93.
4. Горбачев Л.П., Котов Ю.Б., Семенова Т.А. Генерация геомагнитных колебаний на поздней стадии камуфлетного взрыва. // ПМТФ, 1999, Т.40, N 4, с.16-24
5. Гришин В.Л., Диканбаева С.А., Ефимов В.С., Котов Ю.Б., Наумова В.И., Розкин М.Я.. Выработка и использование формального правила для оценки состояния свертывания крови у детей с гломерулонефритом. // Морфофункциональные диагностические подходы в эксперименте и клинике. М., 1986, с. 86-94.
6. Ефимов В.С., Гришин В.Л., Котов Ю.Б., Наумова В.И., Диканбаева С.А., Розкин М.Я.. Диагностическая ценность показателей коагулограммы у детей с хроническим гломерулонефритом. // Педиатрия 1987, N 5, с. 21-22.
7. Ефимов В.С., Котов Ю.Б. Оценка скорости реакции в тонком слое по экспериментальным данным. // Препринт Института прикладной математики им. М.В. Келдыша АН СССР, 1973, N 74, 11 с.
8. Ефимов В.С., Котов Ю.Б., Рачицкая Н.Г., Севастьянов О.И., Лакин К.М., Исследование динамики фибринолиза в тонком слое под влиянием стрептокиназы. // Фармакология и токсикология 1975 , т. 38, № 4, с. 436.
9. Котов Ю.Б. Модели процесса фибринолиза и их использование при автоматизированной обработке измерений. // Препринт Института прикладной математики им. М.В. Келдыша АН СССР, 1984, N 108, 15 с.
- 10.Котов Ю.Б. О формализации структуры решений врача. // Препринт ИПМ РАН, 2001г., № 14, 20 с.
- 11.Котов Ю.Б. Оценка динамики объекта относительно популяции. // Препринт Института прикладной математики им. М.В. Келдыша АН СССР, 1989, N 90, 18 с.
- 12.Котов Ю.Б. Построение шкалы оценки состояния организма новорожденных у матерей с сахарным диабетом. // Инженерная физика 2000, N 2, с.60-65.
- 13.Котов Ю.Б. Программа визуализации и анализа совокупности наблюдений многомерного процесса. // Препринт ИПМ им. М.В. Келдыша АН СССР, 1992, N 54, 19 с.

14. Котов Ю.Б. Программа наглядного анализа числовых таблиц. // Препринт ИПМ, 1988, N 123, 29 с.
15. Котов Ю.Б. Программа получения скользящих нормативов по набору реализаций процесса. // М.: Препринт ИПМ, 1993, N 27, 23 с.
16. Котов Ю.Б. Программа симптомного анализа // Препринт ИПМ им. М.В. Келдыша АН СССР, 1990, N 58, 26 с.
17. Котов Ю.Б. Программы статистических критериев для микрокалькулятора БЗ-34. // Препринт Института прикладной математики им. М.В. Келдыша АН СССР, 1985, N 26, 30 с.
18. Котов Ю.Б. Технология построения шкалы оценки состояния организма. // Препринт Института прикладной математики им. М.В. Келдыша РАН, 2000, N 26, 19 с.
19. Котов Ю.Б. Федорова М.В. Шалаев О.Н. Построение решающего правила прогноза исхода самопроизвольных родов после кесарева сечения. // Научная сессия МИФИ-99. Сб. науч.труд.т.1. М: МИФИ, 1999, с.88-89.
20. Котов Ю.Б., Федорова М.В. Шалаев О.Н. Минимизация размерности описания объекта при выборе тактики в медицинской задаче. // V Всероссийская конференция "Нейрокомпьютеры и их применение", сб. докл., М.: 1999, с.310-313.
21. Котов Ю.Б., Федорова М.В., Троицкая М.В. Оценка состояния новорожденного у матери с сахарным диабетом. // Научная сессия МИФИ-2000. Сб. научных трудов М., 2000, т. 5, с.20-21.
22. Краснопольский В.И., Федорова М.В., Котов Ю.Б., Шалаев О.Н., Мареева Л.С., Левашова И.И., Капустина М.В., Буреева В.Л. Прогнозирование исхода самопроизвольных родов на основании математической оценки партограммы у женщин с рубцом на матке после кесарева сечения. // Проблемы акушерства и гинекологии в исследованиях МОНИИАГ (1929-1994). М., Academia, 1995, с. 38-42.
23. Краснопольский В.И., Федорова М.В., Ларичева И.П., Князев Ю.А., Петрухин В.А., Титченко Л.И., Левашова И.И., Гришин В.Л., Витушко С.А., Бурумкулова Ф.Ф., Котов Ю.Б., Гурьева В.М., Коваленко Т.С. Мониторинг беременных, страдающих сахарным диабетом. // Методические указания N99/01, МЗ РФ, МОНИИАГ, М., 1999, с 22.
24. Петрухин В.А., Федорова М.В., Гришин В.Л., Ефимов В.С., Котов Ю.Б. Нетрадиционный способ введения гепарина в комплексном лечении осложненной беременности у женщин с сахарным диабетом. "Современные проблемы диагностики и лечения нарушений репродуктивного здоровья женщины" // Сб. трудов 1-го Северо-Кавказского съезда акушеров-гинекологов", 7-8 сент. 1994 г, Ростов, с.100.
25. Петрухин В.А., Федорова М.В., Левашова И.И., Князев Ю.А., Тареева Т.Г., Микаэлян Н.П., Витушко С.А., Гришин В.Л., Котов Ю.Б. Оптимизация ведения беременности и родов у женщин с инсулинзависимым сахарным диабетом (ИЗСД). // Проблемы акушерства и гинекологии в исследованиях МОНИИАГ (1929-1994). М., Academia, 1995, с. 26 - 31.

26. Петрухин В.А., Федорова М.В., Левашова И.И., Ларичева И.П., Титченко Л.И., Гришин В.Л., Котов Ю.Б., Витушко С.А. Плацентарная недостаточность при беременности, осложненной инсулинзависимым сахарным диабетом. // Вестник Российской ассоциации акушеров-гинекологов, 1999, N3, с.19-25.
27. Репродуктивное здоровье женщины и потомство в регионах с радиоактивным загрязнением (последствия аварии на ЧАЭС). – Медицина, М. 1997., 393с. п/ред. М.В. Федоровой В.И. Краснопольского А.М. Лягинской. Разд.1, гл.1,4; Разд.3, гл.3; Разд.4, гл.2.
28. Селиваненко В.Т., Котов Ю.Б. Возможности прогнозирования развития ранней артериальной гипертензии после устранения коарктации аорты. // Грудная хирургия 1985, N6, с.36-38.
29. Селиваненко В.Т., Котов Ю.Б. Прогноз развития ранней артериальной гипертензии после хирургической коррекции коарктации аорты. // Тезисы Третьей Всесоюзной научной конференции (12-14 июня 1984, Ереван), с. 270-271.
30. Федорова М.В., Котов Ю.Б. Лукашенко С.Ю., Алексеевский А.В., Дуб Н.В., Сичинава Л.Г., Новикова С.В., Клименко П.А. Возможность прогнозирования задержки внутриутробного развития плода по однократному ультразвуковому исследованию (с использованием скользящих нормативов). // Акушерство и гинекология. N 5, 1991, с.15-19.
31. Федорова М.В., Котов Ю.Б., Лукашенко С.Ю. Опыт использования математических методов при мониторинге беременности и состояния плода. // Вестник Российской ассоциации акушеров-гинекологов 1995, 1, N 2, с. 46-52.
32. Федорова М.В., Краснопольский В.И., Петрухин В.А. Сахарный диабет, беременность и диабетическая фетопатия. – М., Медицина, 2001, 288 с. Гл. 6-7.
33. Федорова М.В., Ларичева И.П., Титченко Л.И., Котов Ю.Б., Лукашенко С.Ю., Петрухин В.А., Антипова И.И., Гришин В.Л. Диагностика и прогнозирование нарушений состояния плода и новорожденного. // Актуальные проблемы перинатологии. Материалы всероссийск. научно-практич. конференции. 15-17 ноября 1994г, г. Чебоксары. Чебоксары 1994, с.244-275
34. Филипович Б.И. Котов Ю.Б. Графический анализ результатов эксперимента. // Приборы и системы управления, 1996, N12, с.40-42 .
35. Филипович Б.И., Котов Ю.Б. Программа графической поддержки исследовательских работ (GRFM). // Препринт ИПУ РАН, М.1997, 28с.
36. Францев В.И., Селиваненко В.Т., Котов Ю.Б. Патогенез, прогноз и профилактика острого гипертензионного синдрома после резекции коарктации аорты. // Тезисы Всероссийского съезда кардиологов Свердловск 22-24 мая 1985, с. 164-165.
37. Krasnopolski V.I., Fedorova M.V., Petrukhin V.A., Grishin V.L., Knyasev Yu.A., Milovanov A.P., Vitushko S.A., Kotov Yu.B. New management method for pregnant women with severe Diabetes Mellitus. // International J. of Gynecology and obstetrics, 1994, v.46, Suppl.2, p.120

38. Fedorova M.V., Kotov Yu.B., Lukashenko S.Yu. The forecast of fetus retardation using ultrasonic data. 10-th Congress of European Association of Gynaecologists and Obstetricians. // Book of abstracts. Monte -Carlo, March 29-Apr. 1, 1995.p 120
39. Petrukhin V.A., Fedorova M.V., Kotov Yu.B., Titchenko L.I., Lukashenko S.Yu. Growth features of newborn in diabetic fetopathy (ultrasound monitoring) // 10-th Congress of European Association of Gynaecologists and Obstetricians. Book of abstracts. Monte -Carlo, March 29-Apr. 1, 1995.p 121
40. Gorbachev L.P., Kotov Yu.B., Semyonova T.A. Geomagnetic variations caused by camouflet explosions: theoretical model and experimental data analysis. // EMC'96 ROMA, International Symposium on Electromagnetic Compatibility, September 17-20, 1996, Rome, Italy p.763-768.
41. Gorbachev L.P., Kotov Yu.B., Semyonova T.A. Generation of geomagnetic oscillations in the late stage of a camouflet explosion. // J.of Applied Mechanics and Technical Physics, 1999, v.40, N 4, pp. 563-570.