

ЭКСПРЕСС-МЕТОДИКА

Работа в РФА с системными многопараметрическими процессами в области медицины на примере использования научно-практического аналитического программного комплекса «ФАКТОР-ПРОГНОЗ» («PLANETA»)

«Фактор-Прогноз» – система РФА, 2006 год,
свидетельство об официальной регистрации в реестре программ для ЭВМ РФ
№ 2006613748,
«PLANETA – Медицинская информация в клинической и экспериментальной радиологии», 2019 год, свидетельство об официальной регистрации в реестре программ для ЭВМ РФ № 2019617905

Соавторы: Колосов Е.А., Долгушин Б.И., Стилиди И.С., Лактионов К.К., Черных М.В., Подвизников С.О., Юдин Д.И., Стрижаков А.Н., Сельчук В.Ю., Игнатко И.В., Мирющенко М.М., Шахламова М.Н., Исаева Э.А., Колосов В.А., Жордания К.И., Романцова О.М., Виршке Э.Р., Черкасов В.А., Трофимов И.А., Никулин М.П., Горбунова В.А., Лаптева М.Г., Перегудов Н.А., Петров В.Б., Киреев А.А.

Технические редакторы: Петров В.Б., Киреев А.А.

Медицинские редакторы: Подвизников С.О., Никулин М.П., Лаптева М.Г.

Москва 2024 год

Оглавление

1. Видеоуроки по работе в программе
2. Введение
3. Работа в ПО «ФАКТОР-ПРОГНОЗ» («PLANETA»)
 - 3.1. Новизна РФА в ПО «ФАКТОР-ПРОГНОЗ» («PLANETA»).
 - 3.2 Системные многопараметрические процессы с точки зрения многопараметрического анализа
 - 3.3 Общая схема информационно–аналитического программного комплекса «ФАКТОР-ПРОГНОЗ» («PLANETA»)
 - 3.4. Шаги работы с РФА в ПО «ФАКТОР-ПРОГНОЗ» («PLANETA»)
 - 3.5. Программный научно-практический аналитический комплекс «ФАКТОР-ПРОГНОЗ» («PLANETA»)
4. СУБД «ФАКТОР-ПРОГНОЗ» («PLANETA») с готовыми опциями
 - 4.1. СУБД «PLANETA» с готовыми опциями
 - 4.2. Варианты сетевых структур сеть «ФАКТОР-ПРОГНОЗ» (PLANETA) и сеть SQL
5. Традиционная однопараметрическая доказательная статистика «ФАКТОР-ПРОГНОЗ»
 - 5.1. Возможности традиционной однопараметрической доказательной статистики «ФАКТОР-ПРОГНОЗ»
 - 5.1.1. Таблицы общей статистики по выборкам по всем параметрам и справочная информация по математическим методам
 - 5.1.2. Расчёт всех корреляций по выборкам с оценкой достоверности в таблицу с графиками и таблицами расчётов
 - 5.1.3. Таблица расчёта достоверности разности между качественными состояниями качественных параметров по параметрам с гистограммами доверительных интервалов и таблицами результатов
 - 5.1.4. Таблица достоверности разности между двумя выборками
 - 5.1.5. Расчёт статистических характеристик между параметрами
 - 5.1.6. Вероятностные таблицы (выживаемость, вероятность отказа) и функции Каплан – Мейера (с расчётом достоверной разности)
 - 5.1.7. Количественные таблицы по выборкам
 - 5.2. Детальная (не табличная) доказательная статистика 2
 - 5.2.1. Детальное исследование корреляции
 - 5.2.2. Детальное исследование достоверности разности
 - 5.2.3. Многопараметрические гистограммы
6. Многопараметрическая доказательная статистика. РФА на примере аналитического ПО «ФАКТОР-ПРОГНОЗ»
 - 6.1. Многопараметрическая доказательная статистика. РА и его развитие. РФА
 - 6.2. Принцип работы с РФА
 - 6.3. Таблица сравнения методов однопараметрической доказательной статистики и многопараметрического РФА
 - 6.4. Возможности РФА в медицине
 - 6.5. Импорт и подготовка данных СУБД «ФАКТОР-ПРОГНОЗ» («PLANETA»)
 - 6.5.1. Разложение качественные параметры как факторы в конец анкеты (таблицы) данных для работы в РФА
 - 6.5.2. Окно быстрого формирования выборок данных

7. РФА в специализированном научно-практическом ПО «ФАКТОР-ПРОГНОЗ»
 - 7.1. Проблемы использования РА при выполнении научно-исследовательской работы
 - 7.2. Получения весовых функций РФА
 - 7.3 Особенности работы РФА для медицины
 - 7.4. Возможное смешение в одной исследуемой группе разных процессов и способы выделения однородных групп
 - 7.5. Создание функций РФА с автоматическим прогнозом
 - 7.6. Различия РФА для числового и качественного критерия
 - 7.7. Создание линейных весовых функций РФА
 - 7.8. Создание нелинейной функции РФА n-го порядка
 - 7.9. Одновременная оптимизация по двум и более нелинейным параметрам – функции РФА n-го порядка
 - 7.10. Мониторинг результатов РФА
8. Прогноз. Создание систем высокоточного прогноза на весовых функциях РФА
 - 8.1. Создание систем высокоточного прогноза на весовых функциях РФА
 - 8.2. Два этапа прогноза для потоков и выборок данных с более чем 1 процессом.
 - 8.3. Прогноз по критерию «Качество» (0–1).
 - 8.4. Создание систем прогноза по рабочим базам данных
 - 8.5. Демонстрационные видео по созданию систем прогноза по линейным и нелинейным функциям РФА
 - 8.6. Основополагающие условия оптимизации и прогнозирования
 - 8.7. Необходимый объем информации для решения поставленных задач
9. Примеры научных работ в РФА в аналитическом программном комплексе «ФАКТОР-ПРОГНОЗ»
 - 9.1. Пример оформления научной статьи с РФА по оценке значимости параметров и созданию системы прогноза
 - 9.2. Пример работы РФА с переходом от менее точных дискретных оценок состояния больного к более точным аналоговым оценкам в диагностике раком желудка
 - 9.3. РФА в области лучевой диагностики в онкологии
 - 9.4. Пример выделения однородных групп по разным критериям
 - 9.4.1. Пример подобия и различия решений РФА по критерию наличия метастазов для общего потока пациентов
 - 9.4.2. Пример разделения общего потока пациентов по критерию
 - 9.4.3. Сравнение результатов исследований по детской БД (НИИ ДОиГ) и взрослой БД (НИИ КО) пациентов по критерию «Наличие метастазов»
 - 9.5. Дифференцированный подход к диагностике и тактике ведения больных с гиперпластическими процессами эндометрия в постменопаузе
 - 9.6. Пример оценки эффективности схем лечения
 - 9.7. Пример определения оптиума
 - 9.8. РФА в области научной стоматологии
 - 9.9. Сравнение расчётных и опытных значений
 - 9.10. Работа с нелинейными функциями РФА порядка > 1
10. Математический аппарат РФА. на примере аналитического программного комплекса «ФАКТОР-ПРОГНОЗ»
 - 10.1. Программная реализация РФА

- 10.2. Правила работы в РФА
- 10.3. Использование качественных параметров в расчётах РФА
- 10.4. Математический аппарат РФА
- 10.5. Краткая схема математического аппарата РФА
- 11. РА и планирование экспериментов с использованием ПО «ФАКТОР-ПРОГНОЗ»
 - 11.1. РА
 - 11.2. Планирование экспериментов
- 12. Используемые материалы

1. Видеоуроки по работе в программе

(для запуска видео перейдите по ссылке, необходимо подключение к сети Интернет)

- 1) Аналитический комплекс «ФАКТОР-ПРОГНОЗ» («PLANETA»)
 - Создание систем прогноза только по графическим характеристикам
 - Создание систем прогноза по лабораторным анализам для предотвращения ошибочных результатов и прогноз дорогостоящих анализов по имеющимся
 - Создание систем прогноза по научным базам данных
 - Создание и аналитика нелинейных систем прогноза степени > 2 для оптимизации и высокоточного прогноза
 - Подготовка данных: импорт – модификация – выборки данных
- 2) Бесплатная СУБД «PLANETA»
 - Сокращенный вариант: основные функции с автоматической настройкой
 - Полная версия + импорт таблиц данных из Excel + настройка ячеек и форм
- 3) Одномерная доказательная статистика (Стьюдент, Фишер...)
 - Сокращенный вариант: основные функции с автоматической настройкой
 - Полная версия + импорт таблиц данных из Excel
- 4) Многопараметрический анализ – оценка эффективности, оптимизация, высокоточный прогноз.
 - Пример в медицине

2. Введение

Основной недостаток однопараметрической статистики (критерии Фишера, Стьюдента) для исследовательской работы с системными процессами, особенно в биологических системных процессах (особенно в медицине), – это низкая информативность для выявления системных взаимосвязей, значимостей и приоритетов влияний параметров многопараметрических процессов на их результат и невозможность обеспечить доказанность результатов без учёта влияний остальных параметров системных процессов в единой математической схеме.

В медицине, где все процессы не только носят системный характер, но и индивидуальны, проблематично создавать идентичные группы для сравнительных доказательных оценок по достоверной разности групп в однопараметрических методах.

Поэтому для исследований и доказательств в системных процессах используется инструмент регрессионного анализа (РА), позволяющий получать оценки и решения с учётом всех параметров системного процесса и индивидуальности каждого процесса в единой математической схеме.

В настоящее время компьютерных технологий возможности анализа данных зависят не столько от математического аппарата, сколько от возможностей программного обеспечения: полноценная аналитическая работа в РФА возможна в специализированном аналитическом комплексе «PLANETA».

Не специализированные статистические пакеты с регрессионным анализом типа SPSS для числовых таблиц получают те же результаты весовых коэффициентов решения систем уравнений регрессионного анализа, но не имеют главных опций для процесса исследований:

- нет опций быстрых выборок данных;
- нет возможности работы со всем набором данных (и с числовыми, и с качественными параметрами) в единой математической схеме;
- нет возможности оценки эффективности параметров (нет преобразований результатов в размерности критериев);
- нет специализированных форм представления результатов РФА в виде гистограмм вкладов параметров в результаты процесса в размерности критериев;
- нет возможности быстрого проведения поисковых исследований по выборкам данных с целью выявления однородных групп в общем потоке данных;
- нет возможности работать с функциями РФА степени > 1 для получения более точных функций РФА и проведения работ по оптимизации воздействий;
- нет возможности создания систем высокоточного прогноза по весовым функциям РФА.

В научном практическом комплексе «PLANETA» («ФАКТОР-ПРОГНОЗ») создан весь цикл научно-практической работы – от ведения регистрационных и дополнительных научных параметров и создания автоматического прогноза для регистрационной мультиформы ведения регистрационных данных.

Классический РА подразумевает работу только с числовыми параметрами системных процессов). РФА (дальнейшее развитие РА) позволяет работать со всем набором параметров (числовым и качественным) в единой математической схеме РА.

Математические инструменты для аналитической работы с системными процессами – методы РА и РФА позволяют:

- 1) оценивать эффективность параметров;
- 2) создавать системы высокоточного прямого прогноза;
- 3) регулярно и быстро обновлять информацию во времени;
- 4) легко развивать и уточнять системы высокоточного прогноза во времени с постоянным мониторингом изменений процессов во времени.

Данная методика отработывалась на самых сложных многопараметрических системных процессах – процессах онкологической медицины.

В методике специально приведены самые простые и понятные примеры из онкологической и неонкологической медицины, экономики.

В медицинских (биологических) системных процессах рекомендуется использовать многопараметрические методы РФА. Эти методы более информативны и проще в понимании и трактовке результата анализа системных процессов и обладают свойством доказанности полученного результата на международном уровне.

РФА имеет ряд существенных преимуществ:

- работает на текущем рабочем материале и не требует дополнительных затрат на проведение исследований – не требует создания специальных групп для ведения исследований и доказательств;
- позволяет сразу на текущей базе данных (БД) оценивать эффективность новых параметров новых процессов и точно управлять процессами на уровне свойства доказанности полученных решений;
- позволяет создавать системы высокоточного прогноза по всему перечню параметров – критериев процесса со свойством доказанности результата, что очень актуально в таких сложных областях медицины, как онкология.

Реализация методов РФА возможна на специализированных системах РФА типа «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА») и её модификации для медицины.

3. Работа в ПО «ФАКТОР-ПРОГНОЗ» («PLANETA»)

Цель методики – показать по этапам работу с инструментом РФА на примере научного программного аналитического комплекса «ФАКТОР-ПРОГНОЗ» («PLANETA»).

Аналитический комплекс «ФАКТОР-ПРОГНОЗ» («PLANETA») – это программное развитие математического инструмента РА до РФА.

РФА, в отличие от РА и однопараметрической статистики, позволяет определять истинную роль фактора в процессе в единой математической схеме на фоне и с учётом всех остальных (числовых и качественных) параметров процесса.

РФА, как математическое развитие РА, имеет свойство доказанности полученного результата на международном уровне, и весовые функции решения систем уравнений РФА оцениваются ВАК как решения, имеющие свойство «открытие» при проведении научных работ на защиту ученой степени доктора наук (для кандидатских работ достаточно свойства «новизна»).

РФА позволяет выделять однородные группы процессов по картинам значимостей параметров процесса на критерии (параметры результата процесса) по всему комплексу критериев – это важно, если вы работаете с группой, в которую входит > 1 (несколько) разных процессов.

Смещение разных процессов в одной исследовательской группе может встречаться в медицинских базах данных – такое смещение не позволяет получать весовые функции РФА с минимальными ошибками и 100-процентным прогнозом. Обычно РФА – очень точный инструмент, и если у вас не получаются прогностические функции со 100-процентным прогнозом, то это может означать присутствие в одной группе разных процессов.

Есть два способа выделять однородные группы с одним процессом при проведении научных работ в области медицинских онкологических процессов:

- 1) по всему потоку: определять однородные группы по сравнению картин влияния параметров процесса на весь набор критериев по различным выборкам данных;
- 2) в небольших группах по максимальным ошибкам: если у вас небольшая исследовательская группа, то отсеять другие процессы можно по максимальным ошибкам прогноза для случаев, которые попадают в зону неопределенности; обычно это 3–5% от всего количества опытов в исследовательской группе, и их можно удалять (до 5%) из группы прогноза.

РФА позволяет перейти от дискретных (качественных) критериев оценки процессов к более точным аналоговым вероятностным оценкам критериев.

РФА позволяет создавать системы высокоточного прогноза на всём объёме данных (и качественном, и числовом) в единой математической схеме.

3.1. Новизна РФА в ПО «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА»).

1. Математический инструмент РФА впервые позволяет проводить *анализ всей информации* – числовой и качественной – в единой математической схеме РА.
2. Рассмотрение всего набора параметров систем в единой математической схеме РФА позволило получать точные окончательные решения и оценки по всему объёму данных процесса *со свойством доказанности результата*.
3. Математический инструмент РФА в ПО «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА») позволяет *приводить результаты расчётов к единой размерности критериев* и давать полную и наглядную картину влияний параметров процесса на критерии в виде гистограмм РФА.
4. Решения по *всему* объёму данных (качественных и числовых) в единой математической схеме позволили получать *высокоточные весовые функции прогноза по всему перечню результирующих параметров критериев* – по числовым и по качественным параметрам-критериям оценки.
5. РФА как инструмент дальнейшего развития РА обладает свойством доказанности полученных решений, что в мировой практике принимается как свойство «открытия» для проведения научных работ на защиту учёной степени доктора наук в различных областях исследований.
6. Важный аспект математического инструмента РФА – возможность проводить исследовательские работы на текущем рабочем материале без дополнительных затрат и издержек на создание исследовательских экспериментальных групп с ограничениями на рабочий процесс – например, как это делается в области онкологии на использовании достоверности разности групп при меньших свойствах доказанности результатов.
7. Для получения первичной оценки по новой схеме лечения на фоне всех параметров организма в составе уже имеющейся текущей группы пациентов – достаточно уже двух-трёх результатов, и в случае отрицательного эффекта прекратить лечение этими препаратами, а в случае положительного эффекта продолжить уточнение окончательной оценки эффективности этого препарата до 10–15 случаев. Для первичной оценки традиционными однопараметрическими методами достоверности разности групп потребовались бы в минимуме две группы по 10–20 человек с ограничениями на проверенные воздействия.
8. В области сложных биологических систем, таких как онкология, РФА позволяет выявлять наличие зависимостей параметров организма с динамикой критериев процессов развития новообразований. Такие зависимости не выявляются никакими другими методами однопараметрического анализа, так как являются результатом комплекса многопараметрических взаимодействий.

3.2. Системные многопараметрические процессы с точки зрения многопараметрического анализа

Прежде чем начать анализ системных многопараметрических процессов, нужно ясно представлять, к каким видам многопараметрических процессов они относятся.

Примеры возможных видов процессов по основным областям:

Медицина – биологические процессы – самая сложная категория процессов, каждый пациент представляет собой индивидуальный многопараметрический процесс с индивидуальными нарушениями.

Экономика и финансы – более простые процессы, работающие по заданным определённым внешним правилам, связаны между собой и различаются между собой по набору правил функционирования процессов; каждый процесс – многопараметрический и индивидуальный только по категориям деятельности и внешним условиям среды; процессы подчиняются заданным правилам, единым и индивидуальным, и установленным внешним и внутренним законам функционирования.

Технические процессы – самые простые не индивидуальные многопараметрические процессы с различиями по группам изделий, видам, категориям, типам и сериям. К техническим процессам применим принцип соблюдения идентичности групп для однопараметрических доказательных критериев (Фишера, Стьюдента...).

Но в макроисследованиях рынка техники и для сравнительных оценок конкурентной способности и функционирования в различных внешних условиях пригоден только инструмент РФА.

3.3 Общая схема информационно-аналитического программного комплекса «ФАКТОР-ПРОГНОЗ» («PLANETA»)



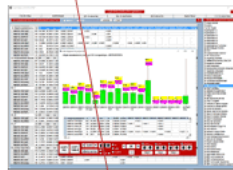
система формирования выборок ... все операции проводятся для сформированной выборки данных

Однопараметрическая аналитика – дост.разн.групп, корреляции, (Стьюдент, Фишер ...)

Анкетная форма с графиками



Доказательная однопараметрическая статистика: Стьюдент, Фишер ... ВСЕ расчёты за одну операцию – 4 ВАРИАНТА



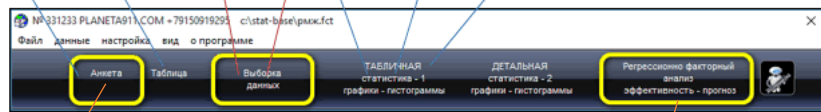
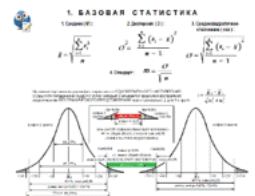
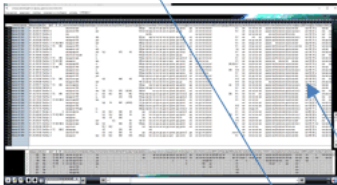
Достоверность корреляций ВСЕ расчёты и графики за одну операцию (графики по щелчку на ячейке)



Вероятностные таблицы, функции, графики: надёжность, выживаемость, Калл-Мейер...

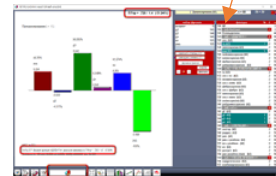
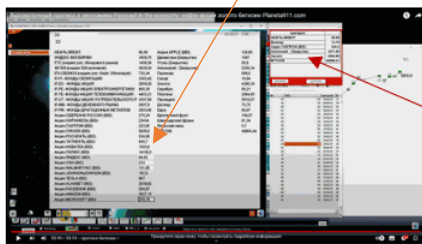


Табличная форма со ВСЕМИ опциями преобразования столбцов



Многопараметрическая аналитика «РФА и прогноз»

Создание систем высокоточного прогноза по всему перечню критериев



Функции РФА порядка > 1 поиск оптимумов 3D и 2D графика

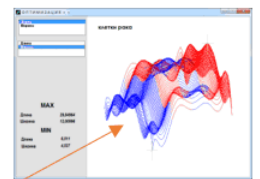
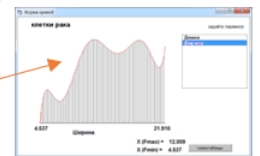
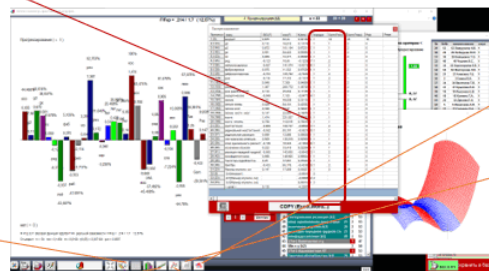
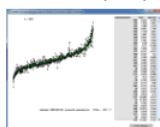


Таблица весовых вкладов, весовых коэффициентов и диапазонов изменения параметров

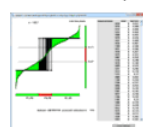


Графики адекватности функций РФА – соответствие расчётных и опытных значений



числовой критерий

качественный критерий



Проверка реальной точности прогноза функций РФА по схеме «минус 1» (поочередно 1 опыт убирается из расчёта и используется в прогнозе как прогнозируемый)

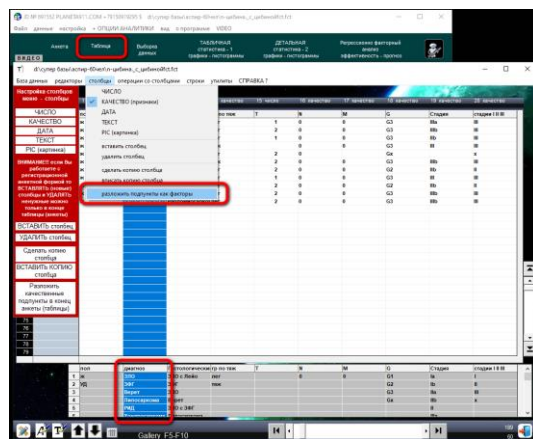
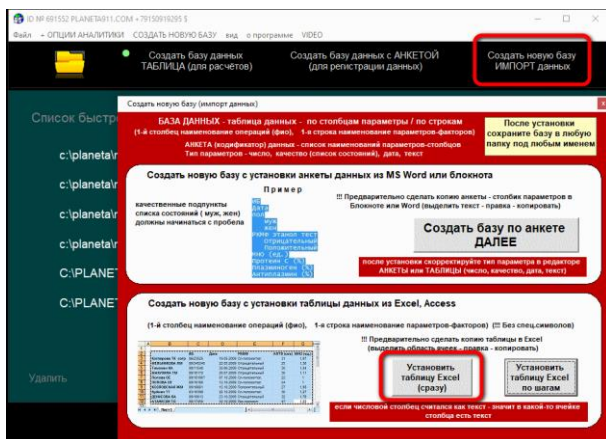


Импорт таблиц данных из любых приложений Windows.

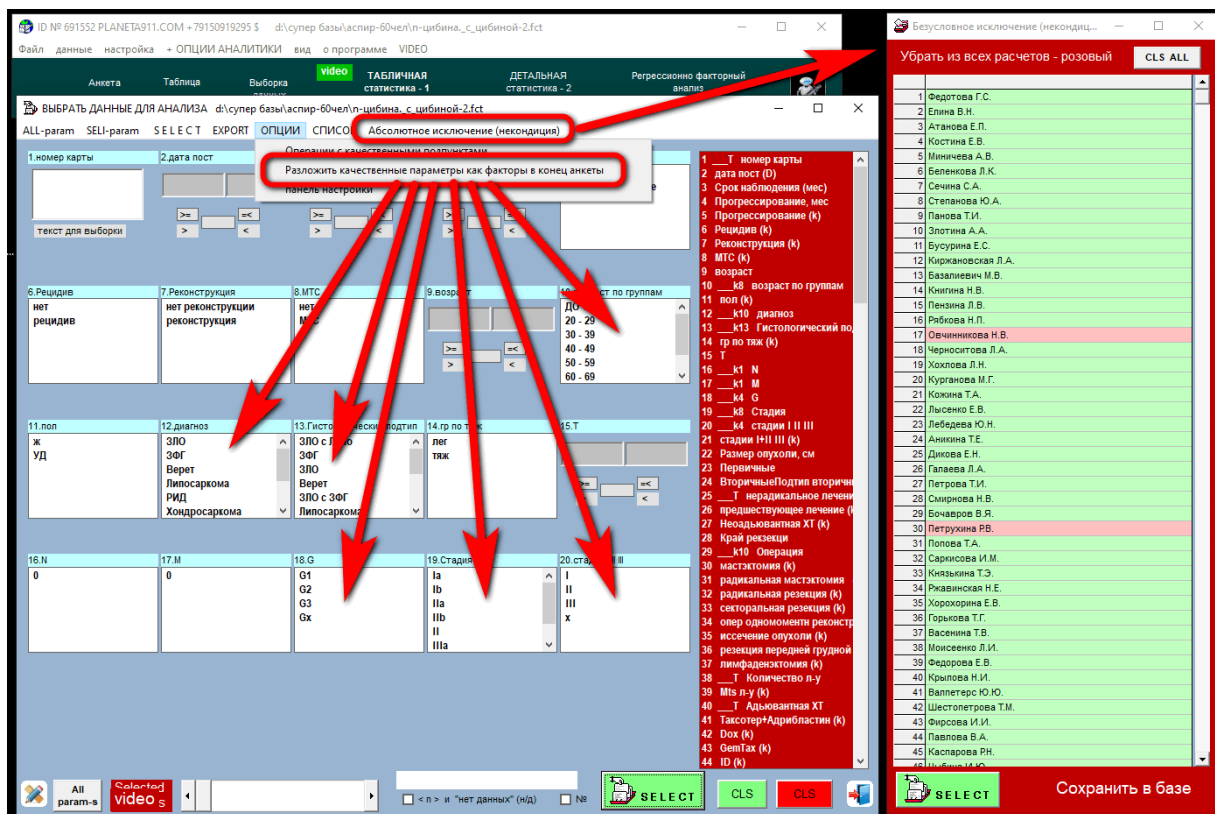
Текстовые ячейки автоматически сформируются в качественные параметры

Для качественных состояний больше двух.

Опция – разложить «подпункты как факторы» в двузначные факторы (1 – нет, 2 – есть фактор) в конец таблицы (анкеты) для работы в РФА



В форме «ВЫБОРКИ ДАННЫХ» тоже есть опция разложения параметров с $n > 2$ в виде дополнительных двузначных параметров-столбцов в конце таблицы (анкеты) для работы с РФА – для всех качественных параметров с $n > 2$.

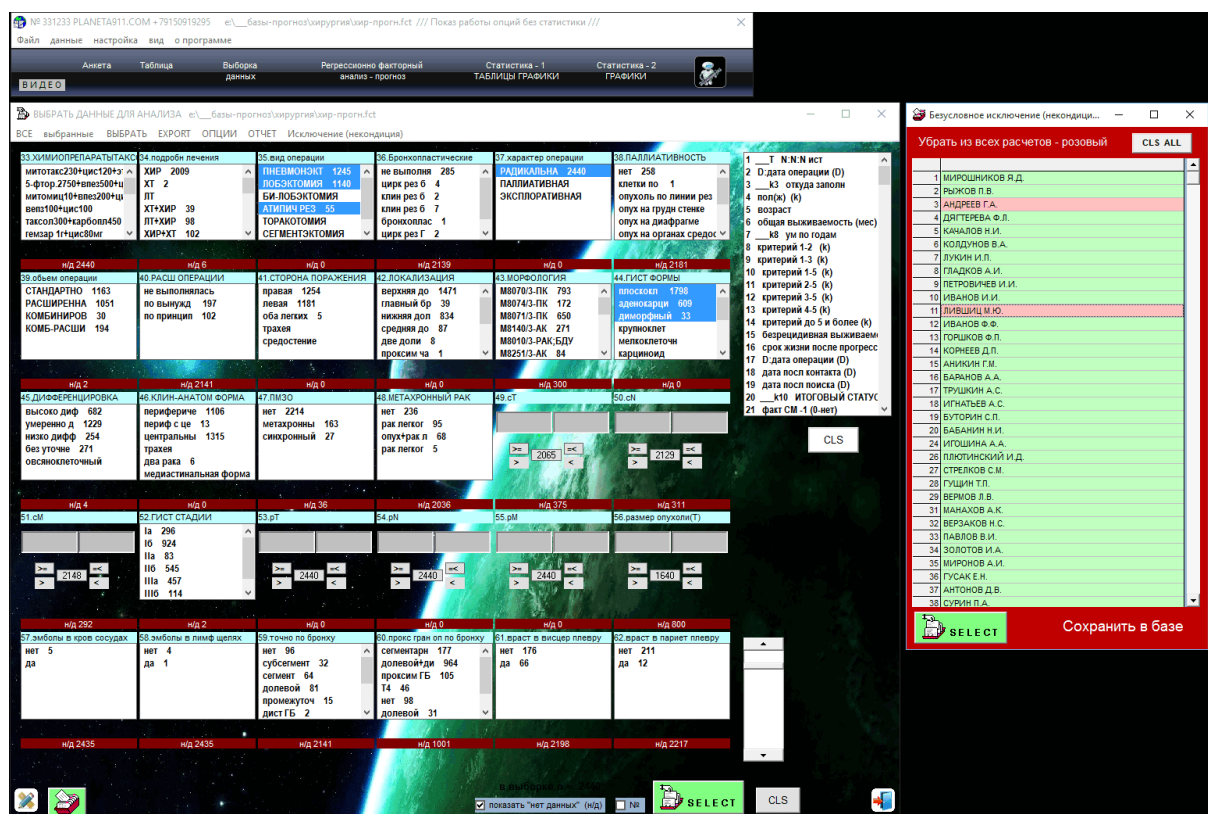


Шаг 2 – в окне «Выборки данных» можно сделать любую выборку данных для расчётов по необходимости.

Выборки данных для РФА очень важны – особенно когда есть наличие разных процессов в базе данных.

Разные процессы имеют разные картины влияния параметров процесса на его результаты (критерии) и должны рассматриваться отдельно.

Разделять процессы можно по параметрам и по картинам влияния параметров процесса на его результаты (критерии)



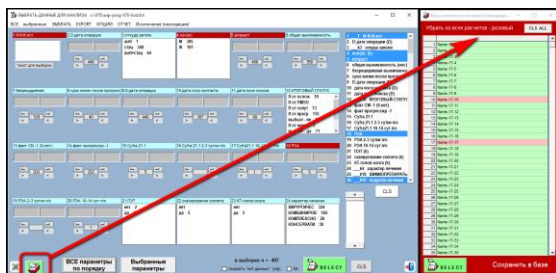
Дополнительная красная форма позволяет безусловно исключать любые опыты из любых расчётов.

Безусловно исключаются обычно явно неадекватные или ошибочные данные опытов, явно не вписывающиеся в процесс, – для уменьшения ошибок прогноза.

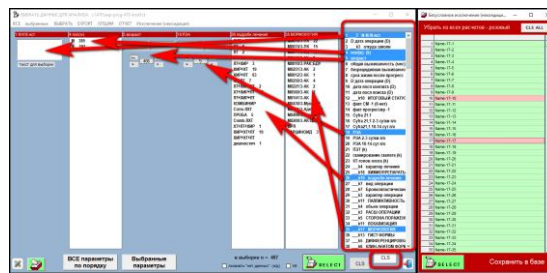
В медицинских биологических базах можно исключать до 5% данных, при этом адекватность и точность весовых функций значительно возрастает – это также способ удалять из расчётов данные других процессов, которые явно не вписываются в весовую схему исследуемого процесса и вносят большие ошибки в адекватность реальной зависимости

Безусловное исключение и параметры из разных мест


Задание ограничений по параметрам.
Безусловное исключение – красное окно



Для удобства работы с выборками можно выбрать только нужные параметры из всей базы



Шаг 3.1 – получение весовых функций РФА

Получение весовых функций РФА осуществляется в окне «Регрессионно-факторный анализ» по кнопке 

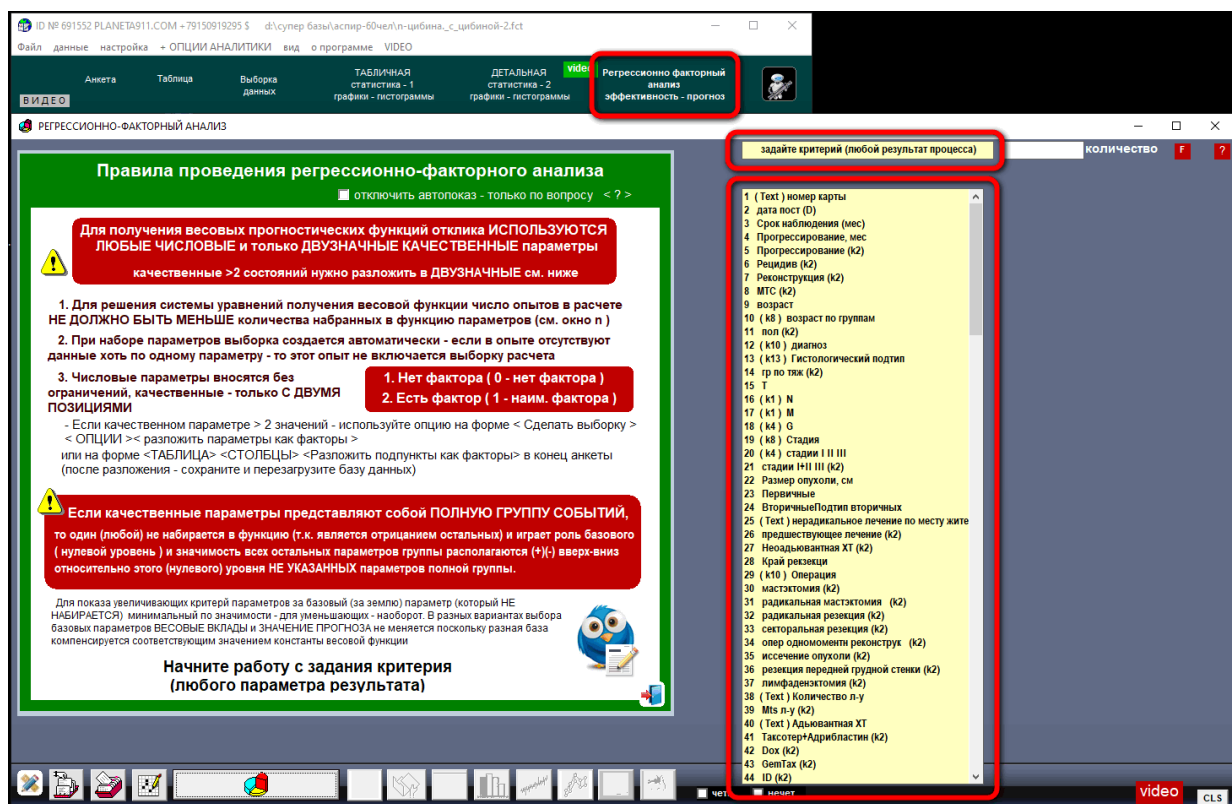
1. Сделайте выборку данных, если это необходимо.

Выборки данных в области медицины необходимы, чтобы избежать смешения разных биологических процессов в расчётах: при смешении разных процессов в расчётах вы не получите высокоточных оценок значимости параметров на результаты (критерии) этих процессов и не сможете создать точные системы прогноза параметров результата (критериев).

Большие значения ошибок в расчётах свидетельствуют о присутствии в расчётах опытов из разных процессов, а малые ошибки свидетельствуют о высокой степени однородности процесса и готовности весовой функции к использованию для прогноза параметров результатов (критериев).

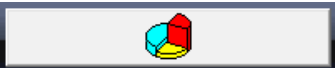
2. Перейдите в окно «Регрессионно-факторный анализ» и задайте параметр критерия оценки из списка параметров (жёлтый список) базы данных (верхнее поле критерия) – обычно это числовые или качественные (двузначные) параметры результата (критерии оценки).

Задание критерия для расчёта весовой функции РФА (на форме РФА есть подсказка (зелёная панель) по правилам получения расчётов функций РФА).



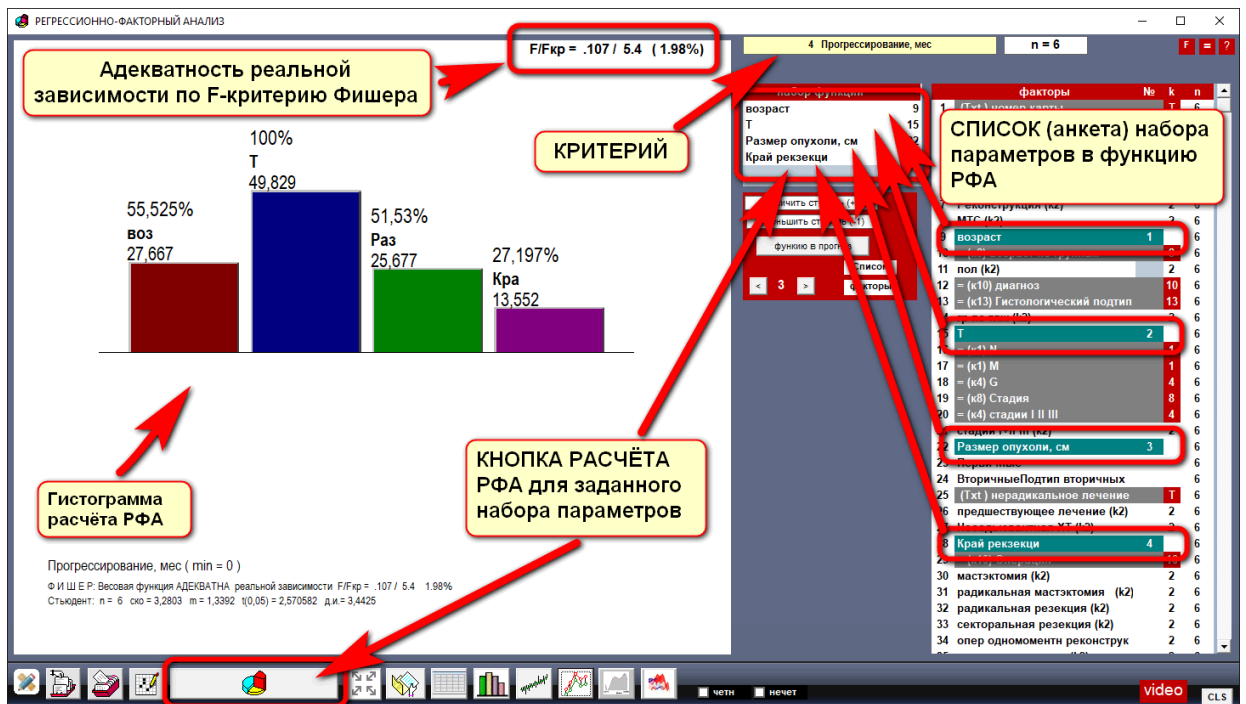
Использование числовых параметров в качестве критерия даёт более точные результаты и меньшие ошибки, чем использование качественных параметров.

3. После задания критерия из правого списка параметров базы данных мышкой набираются параметры в левый список набора весовой функции РФА и нажимается кнопка расчёта. Если адекватность весовой функции РФА реальной зависимости увеличивается (ошибка между расчётными и опытными значениями уменьшается – см. % ошибки от критического значения), то параметр оставляется; если адекватность весовой функции уменьшается (ошибка увеличивается), то параметр повторным щелчком мыши убирается из набора весовой функции.

После задания критерия и параметров РФА для получения весовой функции с её характеристик адекватности нажимаете кнопку 

Простой пример по анализу хирургической операции в РФА в группе РМЖ:

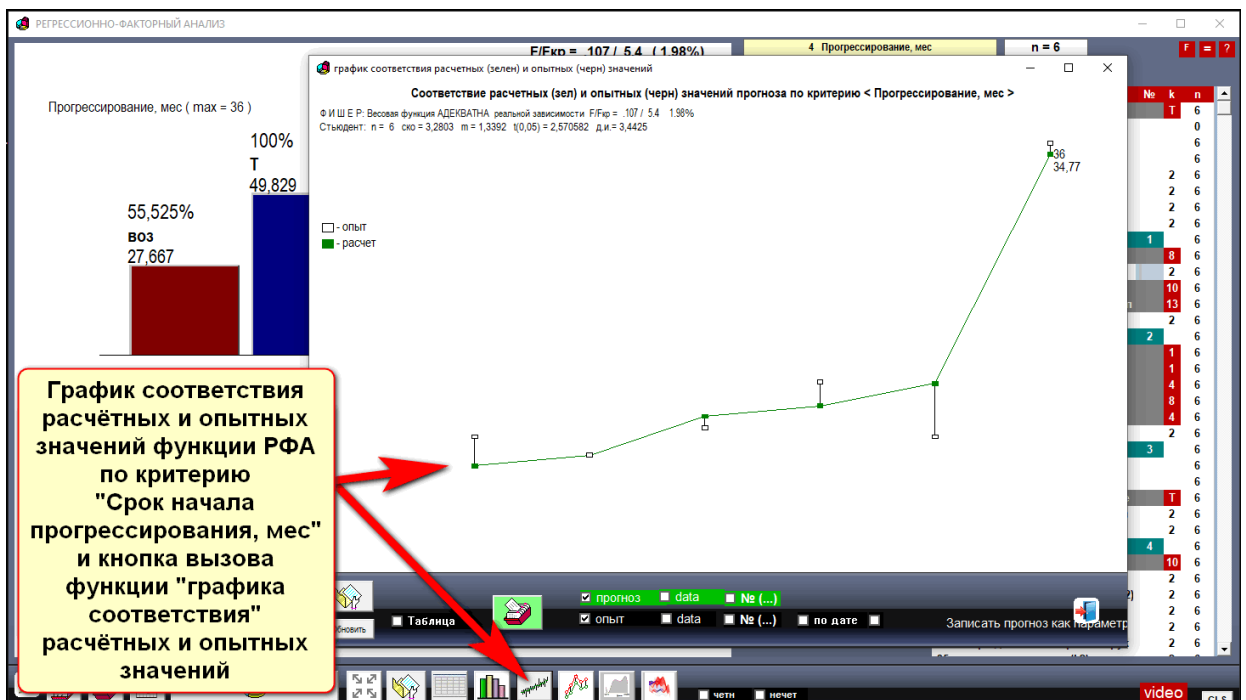
По числовому критерию «Срок начала прогрессирования, мес.»



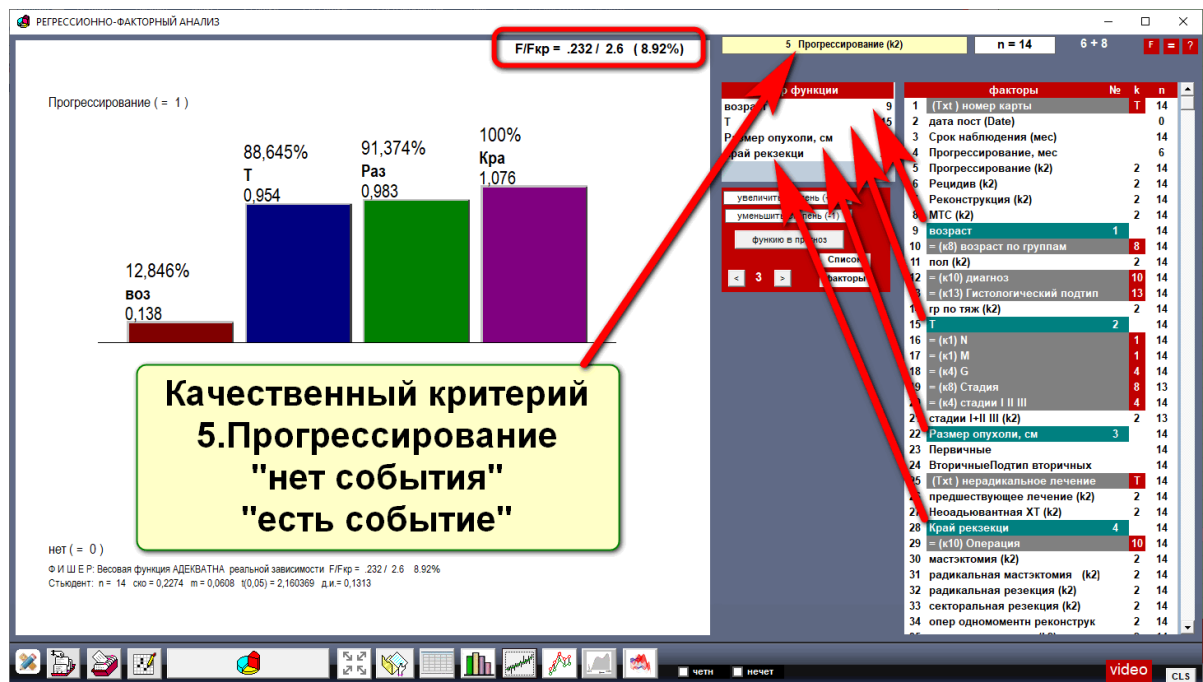
По этому расчёту можно сделать выводы: параметры «возраст», «Т», «размер опухоли, см» – не регулируемые параметры диагноза, но параметр «край резекции» регулируется хирургом, и чем больше край резекции, тем позднее возникает рецидив.



Графики соответствия расчётных и опытных значений по числовому критерию «Срок начала прогрессирования, мес.»



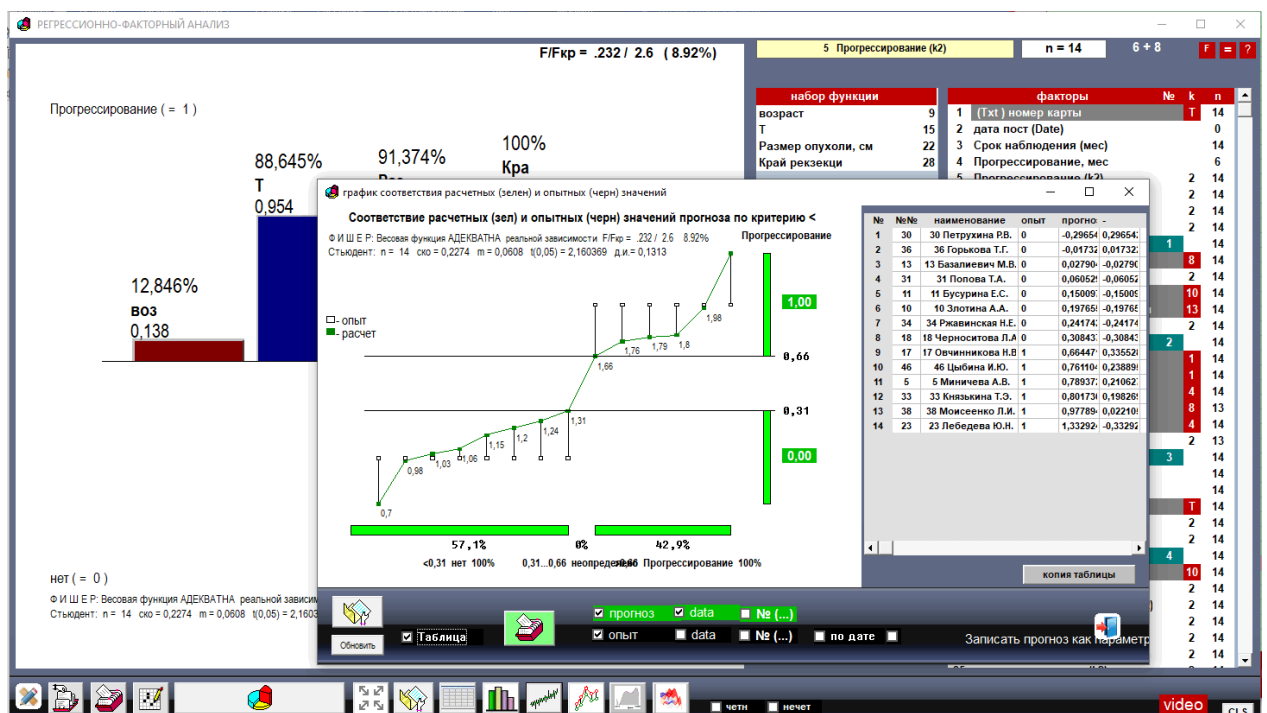
Графики соответствия расчётных и опытных значений по качественному критерию «Событие прогрессирувания»



По этой картине величин и приоритетов значимостей можно сделать выводы, что на событие возникновения рецидива «возраст» не существенно не влияет – значимы не регулируемые параметры диагноза и наиболее значим на событие возникновения рецидива регулируемый хирургом параметр «Край резекции».



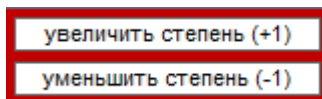
Графики соответствия расчётных и опытных значений по качественному критерию «Прогрессирование»



По данным расчётов и числового и качественного критериев наблюдается высокая информативность, но числовые критерии имеют более высокую точность и адекватность реальной зависимости, чем качественные критерии при меньшем объеме данных.

Но, в отличие от качественных критериев, числовые критерии при точных определениях сроков события не отвечают на вопрос по факту возникновения или отсутствия возникновения данного события «прогрессирования» – у числовых и качественных критериев разные функциональные задачи.

Шаг 3.2 – оптимизация



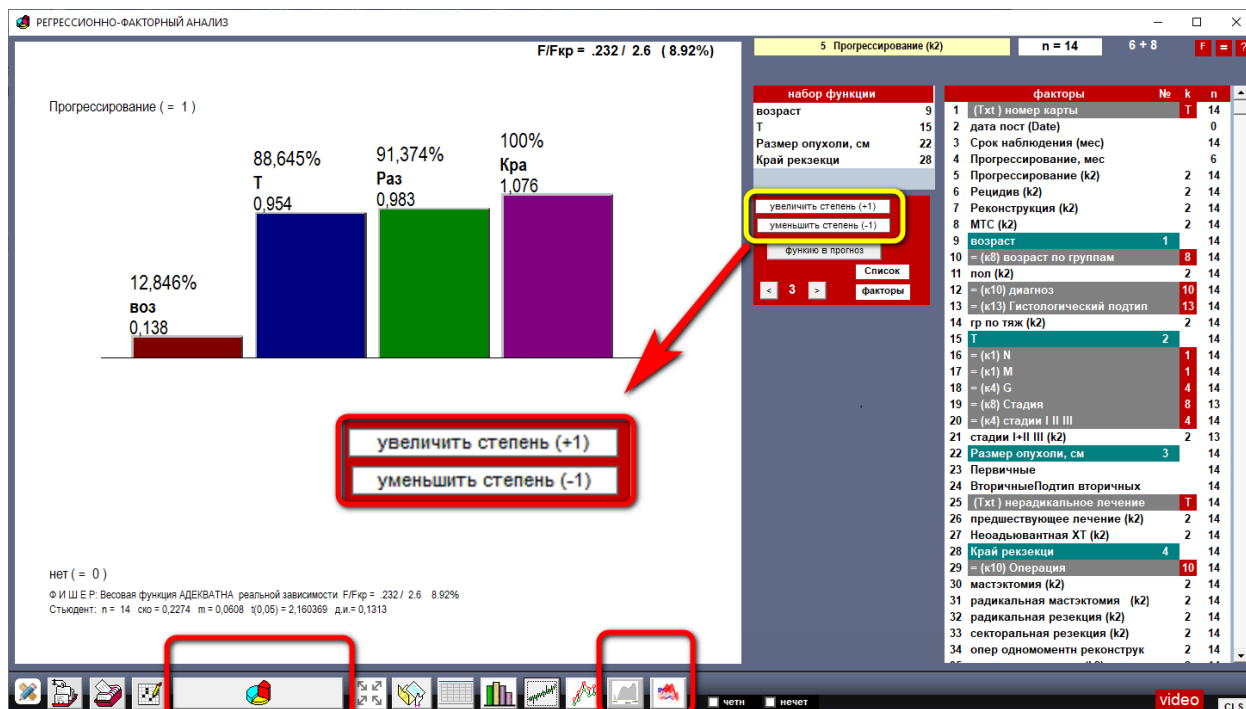
Получение нелинейных весовых функций

РФА порядка > 1

Получение нелинейных весовых функций РФА в виде полиномов степени параметров > 1 возможно, если позволяет набор данных процесса.

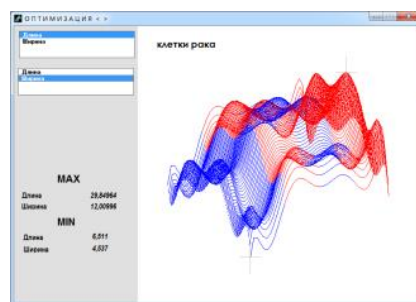
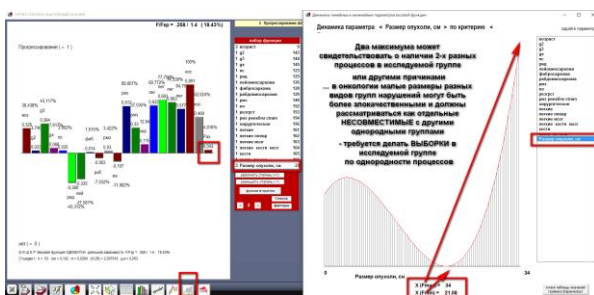
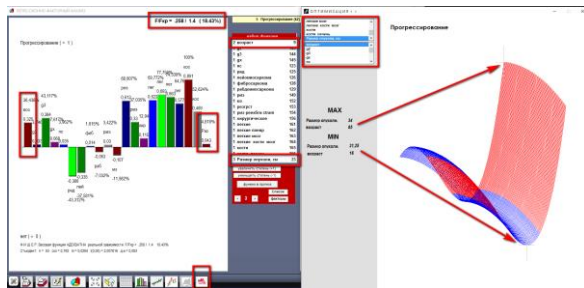
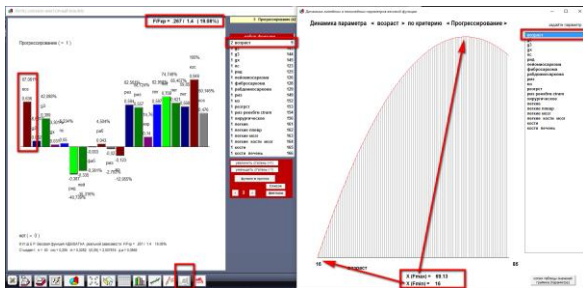
Поднимать степень параметров можно только у числовых параметров при условии, что адекватность реальной зависимости функции улучшается.

Технически в программе «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА») это делается очень просто двумя кнопками – степень повышается и оставляется до значения, при котором адекватность функции улучшается (уменьшается ошибка между расчётными и опытными значениями).





Посмотреть вид нелинейности параметра полинома со степенью > 1 и выявить оптимумы можно на формах 2D и 3D измерений по кнопкам на панели РФА.



Шаг 4 – таблица расчёта РФА

РЕГРЕССИОННО-ФАКТОРНЫЙ АНАЛИЗ

F/Fкр = .232 / 2.6 (8.92%)

5 Прогрессирование (k2) n = 14 6 + 8

Прогрессирование (= 1)

Возраст: 12,846%, вес: 0,138

T: 88,645%, 0,954

Раз: 91,374%, 0,983

Край резекции: 1,08

1 (Прогрессирование (= 1))

возраст: 0,14

размер опухоли, см: 0,98

Край резекции: 1,08

+ Const 2,395

Прогресс	наим	век(F)	изм(F)	К(век)	порядок	Хотт(Fmin)	Хотт(Fmax)	Fmin	Fmax
1 (9)	возраст	0,138	11,79	0,0031251	1	18	57	18	1
2 (15)	T	0,954	116,179	0,73968				1	1
3 (22)	Размер опухоли, см	0,983	166,475	-0,23968				1,5	1,5
4 (28)	Край резекции	1,076	87,916	-0,83968				0	0
	< const >	2,395							

Таблица расчётов РФА

Готовая форма оформления гистограмм РФА для переноса в приложения Windows (для документов) Power Point, Word, Excel...

Ф И Ш Е Р. Весовая функция АДЕКВАТНА реальной зависимости F/Fкр = .232 / 2.6 8.92%

Студент: n = 14 ско = 0,2274 m = 0,0608 t(0,05) = 2,160369 д.и. = 0,1313

Файлы: name, data, %

Оформление: Align, < .00, > .00, < >, < >

video CLS



Таблица расчёта РФА



Гистограмма РФА для слайдов

Прогрессирование F/Фкр = .232 / 2.6 n = 14	Наименование	ВЕС (F)	Изм (F)	К (вес)	Порядок	X опт (F min)	X опт (F max)	P min	P max
1 (9)	Возраст	0,138	11,78	0,0031251	1	18	57	18	57
2 (15)	T	0,954	118,178	0,7852366	1	1	2	1	2
3 (22)	Размер опухоли, см	0,983	166,475	-0,2278143	1	13	1,5	1,5	13
4 (28)	Край резекции	1,076	87,916	-0,8704583	1	2	0	0	2
	< const >	2,395		2,39539					

1 столбец – в скобках порядковый номер параметра в базе данных;

2 столбец – наименование параметров;

3 столбец – ВЕС (F) – весовой вклад параметра в критерий в размерности критерия:

для числового параметра это среднее значение по группе;

для качественного параметра (0 или 1) – 1 (есть фактор);

4 столбец – изм (F) – значение диапазона изменения критерия между P min и P max параметра;

5 столбец – К (вес) – весовые коэффициенты РФА и свободный член < const > внизу столбца;

6 столбец – порядок – максимально заданная степень параметра в весовой функции;

7, 8 столбцы – X опт (F min), X опт (F max) – значение параметра при котором критерий принимает минимальное и максимальное значение;

9, 10 столбцы – P min, P max – минимальное и максимальное значение параметра в группе – необходимая информация при прогнозе; значения параметров при прогнозе не должны выходить за эти границы, иначе прогноз будет считаться условным.

ВНИМАНИЕ! – эти значения нужно обязательно указывать в инструкции к прогнозу.

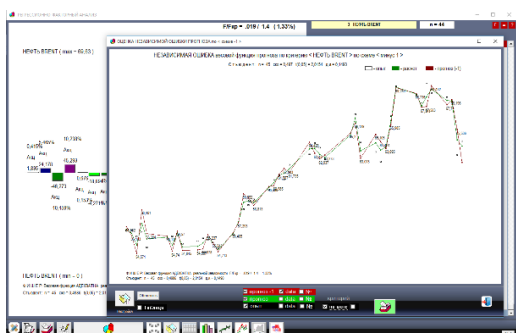
При выходе параметров за пределы диапазонов прогноз нужно считать условным.

Шаг 5 – создание систем автоматического прогноза

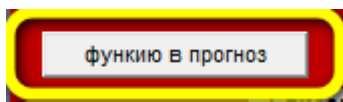
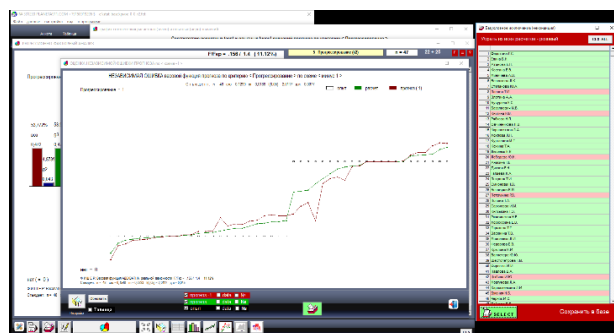


1. Функции и ошибки адекватности по схеме «Минус 1» показывают, какие ошибки будут в прогнозе по вновь прибывшим – поочередно 1 случай убирается из расчёта и рассчитывается как вновь прибывший для расчёта ошибки прогноза по вновь прибывшим.

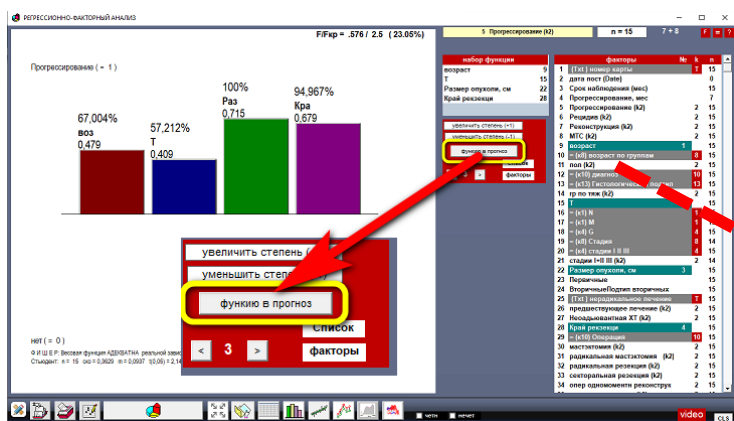
Соответствие расчетных и опытных значений для числового критерия по схеме «Минус один»



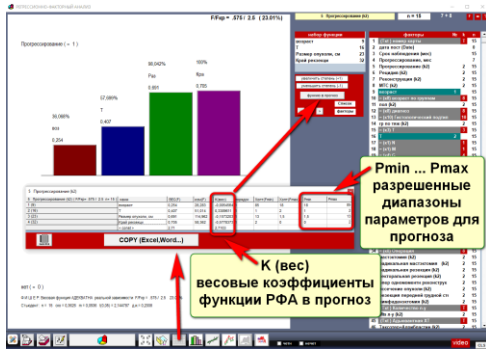
Соответствие расчетных и опытных значений для качественного критерия по схеме «Минус один»



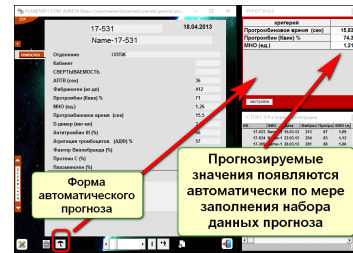
Функция РФА записывается в форму «Прогноз» по кнопке «Функцию в прогноз» и «Дописать»: для автоматического прогноза при заполнении в регистрационной анкете данных.



В форму «Прогноз» устанавливается столбец весовых коэффициентов К (вес) из таблицы результатов РФА. Р min ... Р max – разрешённый диапазон параметров при прогнозе.



Автоматический прогноз по нескольким критериям



ПРИМЕР 2. Сравнение работы инструмента РФА в области медицины и экономики.

Математический аппарат РФА универсален для работы с любыми системными процессами. Медицинские (биологические) процессы – самые сложные, поэтому предлагается для лучшего понимания математического инструмента РФА рассмотреть примеры по шагам РФА в экономике и медицине для сравнения:

1 и **2** ШАГ – разложить качественные параметры с $n > 2$ как факторы в конец таблицы и сделать выборку (по необходимости).

Экономическая база.
Критерий – число.
Стоимость "Нефть Brent"

Медицинская база.
Критерий – качество.
Событие "Прогрессирование"

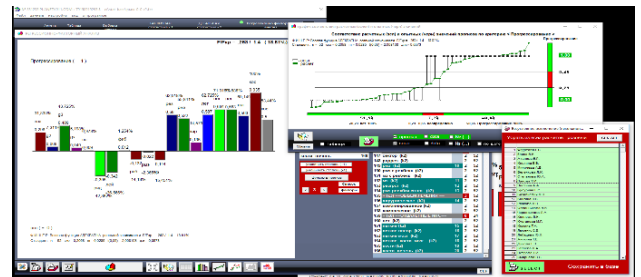
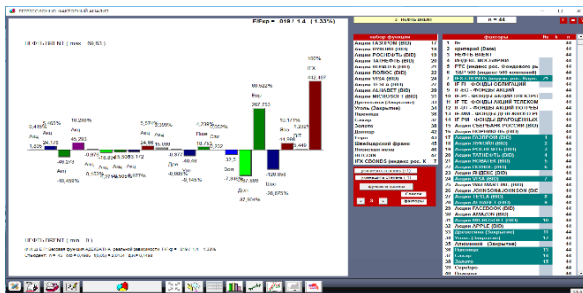
ВИДЕО Экономика

ВИДЕО Медицина

3 и **4** шаги  – получение функций РФА

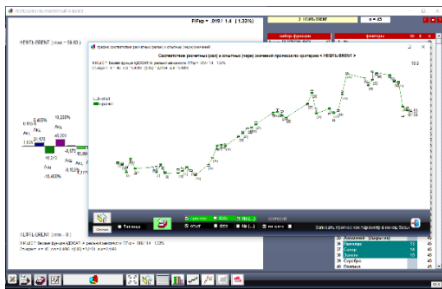
Задание критерия «Нефть Brent» (экономика) и набор параметров в весовую функцию. Числовой критерий

Задание качественного критерия «Прогрессирование» (медицина) и набор параметров в весовую функцию. Качественный критерий

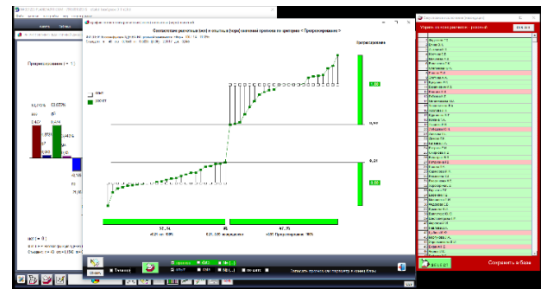




Соответствие расчетных и опытных значений (экономика) для числового критерия



Соответствие расчетных и опытных значений (медицина) для качественного критерия

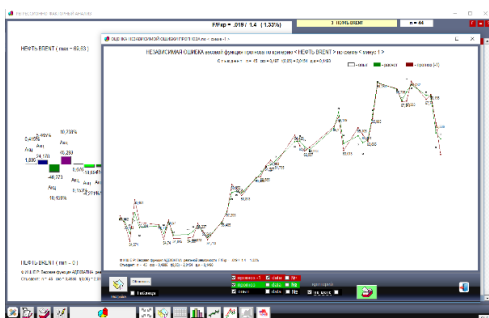


5 шаг – записать функцию РФА в автоматический прогноз

Проверка ошибки прогноза по схеме «Минус 1»

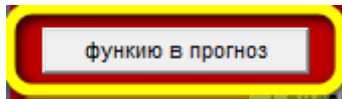


Соответствие расчетных и опытных значений (экономика) для числового критерия по схеме «Минус один»

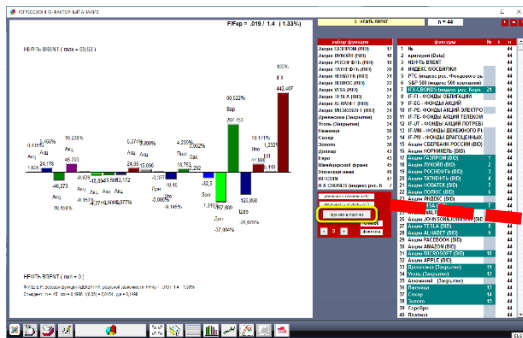


Соответствие расчетных и опытных значений (медицина) для качественного критерия по схеме «Минус один»





Записываем окончательную функцию РФА в форму «Прогноз» по кнопкам «Функцию в прогноз» и «Дописать»



Золото	38
Доллар	42
Евро	43
Швейцарский франк	45
Японская иена	46
БИТСОИН	47
IFX.CBONDS (индекс рос. К	7

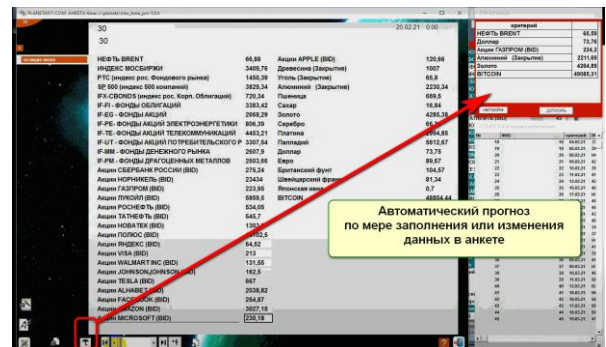
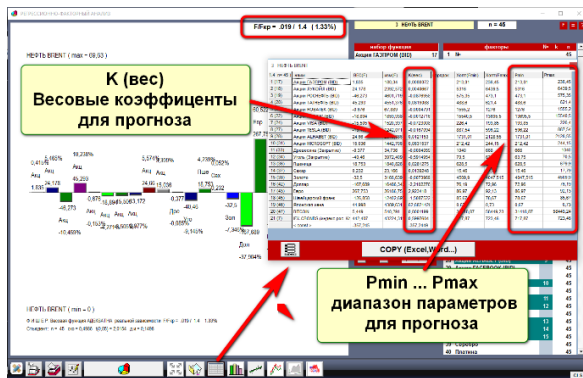
увеличить степень (+1)
уменьшить степень (-1)
функцию в прогноз
СПИСОК факторы

ПРОГНОЗ	
критерий	
НЕФТЬ BRENT	65,59
Доллар	73,76
Акции ГАЗПРОМ (БИД)	224,2
Алюминий (Закрытие)	2211,89
Золото	4264,89
БИТСОИН	49085,31

настройка **ДОПИСАТЬ**

Автоматический прогноз в экономике по перечню критериев

Автоматический прогноз в экономике по перечню критериев



ВНИМАНИЕ! В таблице расчётов за зеленым столбцом коэффициентов идут столбцы диапазонов допустимых минимальных и максимальных значений параметров – эти значения нужно обязательно указывать в инструкции к прогнозу. При выходе параметров за пределы диапазонов прогноз нужно считать условным.

3.5 Программный научно-практический аналитический комплекс «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА»)

Эффективная научная работа в медицине возможна только при наличии всего вспомогательного обеспечения научной работы непосредственно на текущем рабочем материале:

I – наличие практической регистрационной СУБД со стандартным регистрационным набором данных и возможностью вести дополнительный набор данных с возможностью добавления параметров самими пользователями;

II – наличие традиционного инструмента традиционной однопараметрической доказательной статистики расчётов по выборкам и по всем параметрам сразу за одно

нажатие кнопки расчёта в виде таблиц – огромная экономия времени на традиционные научные расчёты;

III – наличие многопараметрического доказательного инструмента РФА для многопараметрических процессов – оценки влияния всех параметров процесса на параметры результатов (критерии) процесса в единой математической схеме РФА и создания систем высокоточного прогноза на весовых функциях РФА.

В научном аналитическом программном комплексе «PLANETA» присутствуют все этапы научной работы:

Импорт-экспорт данных + СУБД (регистрационная мультианкетная и табличная формы + выборки данных + однопараметрическая доказательная статистика + многопараметрический РФА – оценка эффективности – величин значимости параметров на весь перечень критериев (параметров результатов) + создание систем высокоточного прогноза).

Этапы научного анализа:

СУБД «ФАКТОР-ПРОГНОЗ» («PLANETA») с готовыми опциями

- разработан с целью возможности ведения всей научной работы на текущем рабочем материале – без необходимости создания баз для исследований;
- меню абзацев сокращает время на создание документов с 1–2 часов до 1–5 минут;
- огромная экономия времени и ресурсов на ведение данных и научные исследования.

Традиционная однопараметрическая доказательная статистика:

- все возможные варианты расчётов по всем параметрам получаются в таблицах по одному нажатию кнопки;
- огромная экономия времени на ведение расчётов по сравнению с другими статистическими пакетами типа SPSS, Statistica, где нет возможности быстрых выборок данных и все расчёты нужно делать, перебирая попарные комбинации.

6. Многопараметрическая доказательная статистика – РФА:

- развитие РА;
- позволяет решить главную задачу любого анализа – получить исчерпывающую картину влияния всех параметров процесса (и числовых, и качественных) на результат (критерий) в единой математической схеме РА;
- оценка эффективности и оптимизация параметров воздействия;
- создание систем высокоточного прогноза.

4. СУБД «ФАКТОР-ПРОГНОЗ» («PLANETA») с готовыми опциями

Импорт-экспорт данных, выборки данных, возможность делать научную работу на текущем рабочем материале (без траты времени на создание баз данных для научных исследований и проверки эффективности новых препаратов)

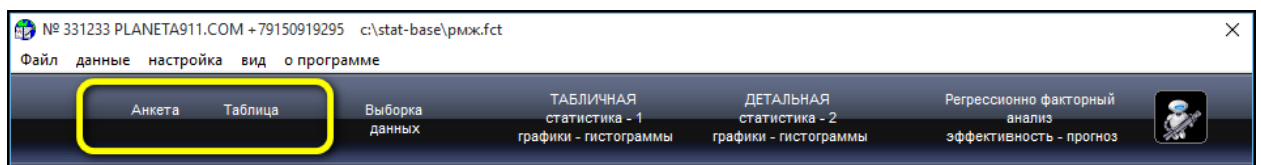
Готовая СУБД «PLANETA» разработана как приложение к аналитике РФА для того, чтобы научную работу по РФА можно было делать на текущем рабочем материале без траты времени на создание баз данных – в медицине эта работа занимает несколько месяцев поиска и выписывания данных.

СУБД «PLANETA» имеет интеграцию с медицинскими СУБД на Oracle и SQL.

4.1. СУБД «PLANETA» с готовыми опциями

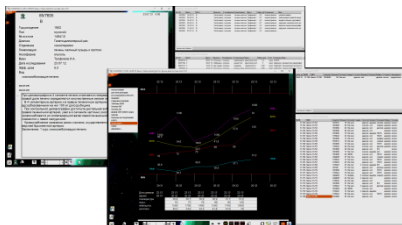
[Короткое видео](#)

[Полное видео](#)



Основной готовый функционал СУБД «PLANETA»:

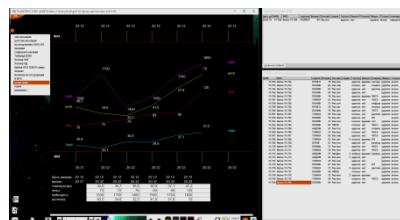
Мультианкетная форма
(+ графики)



VID-DEMO

VID-DEMO

Ведение графиков внутри
анкеты



VID-DEMO

VID-FULL

Показ графиков между
анкетами с одним ID №



VID-DEMO

VID-FULL

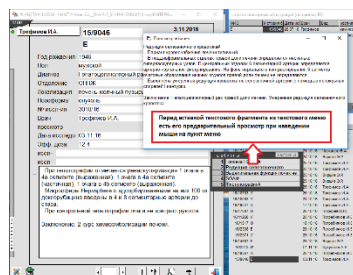
Галереи картинок с подписями
для каждой ячейки



VID-DEMO

VID-FULL

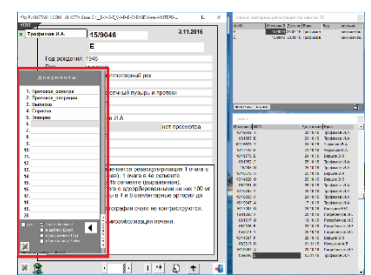
Вставка готовых абзацев,
добавляемых пользователем



VID-DEMO

VID-FULL

Два вида печати документов:
Planeta и Excel

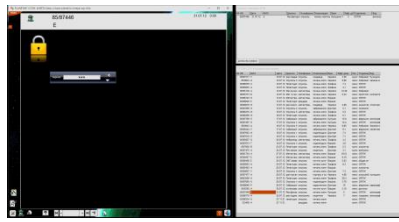


VIDEO

Табличная форма (научная работа)



Защита паролем отдельной записи исполнителем



Шифрование записи данных



VIDEO

VID-DEMO

VID-FULL

VID-DEMO

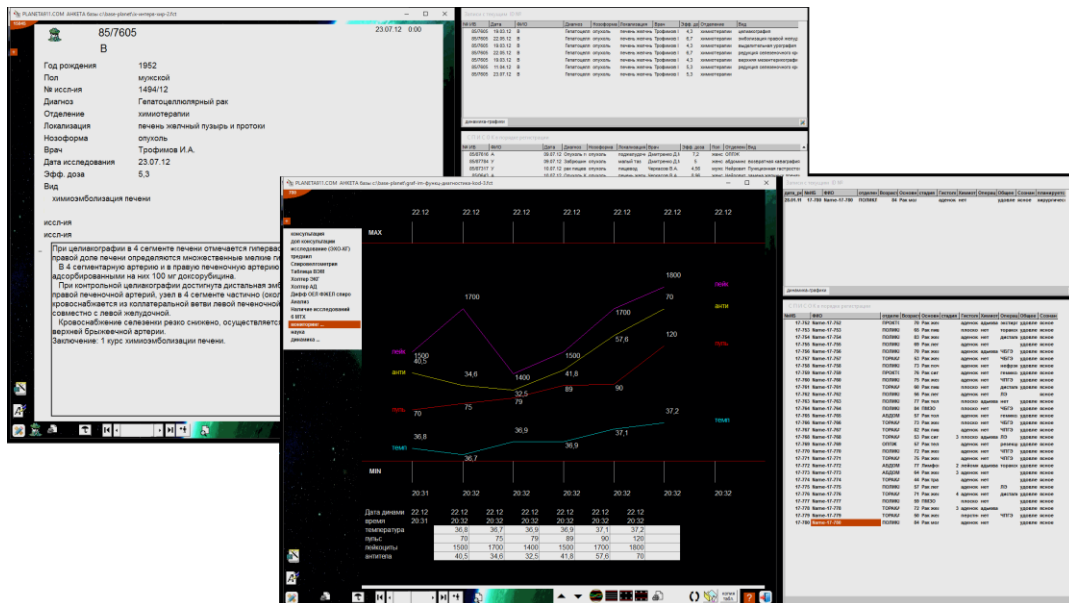
VID-FULL

Автоматическая мультиформа позволяет вести неограниченное количество статических и графических форм представления данных, а также в окне автоматического прогноза позволяет получать автоматические прогностические оценки по весовым прогностическим функциям РФА.

Разработана с целью возможности ведения всей научной работы на текущем рабочем материале – без необходимости создания баз для исследований.

Меню абзацев сокращает время на создание документов с 1–2 часов до 1–5 минут.

Огромная экономия времени и ресурсов на ведение данных и научные исследования.



Автоматический прогноз параметров гемостаза при заполнении комплекта данных (на форме прогноза по кнопке «Прогноз» на мультиформе регистрации данных)

ПРОГНОЗ

критерий	значение
Протромбиновое время (сек)	15,83
Протромбин (Квик) %	74,3
МНО (ед.)	1,21

настройка

СПИСОК в порядке регистрации

ИД	ФИО	Дата	Фибрин	Протро	МНО (ед.)
17-523	Иванова	15.03.13	213	87	1,09
17-524	Иванова-1	22.03.13	254	83	1,12
17-395	Иванова-1	28.03.13	281	88	1,08

Форма автоматического прогноза

Прогнозируемые значения появляются автоматически по мере заполнения набора данных прогноза

Неограниченные галереи картинок для каждого параметра с одним номером ID.

Галерея картинок и сканы документов для каждого параметра. В базе, например, 200 параметров – можно вести 200 галерей под одним ID-номером.

ГАЛЕРЕЯ 00/86770 А

№ 00/86770 А

Name: Ivannikova, E. N.
ID: 21 / 227
DOB: 01.1965
Acquisition Date: 25.05.2021
Acquisition Time: 15:37:30
Image Number: 17
A: 1.57

Acquisition Number: 9
Slice Location: 881.00
Slice Thickness: 4.00
CS: 2.00
Exposure Time: 800
KVP: 130.00
X-ray Tube Current: 10
Feed: 12.00

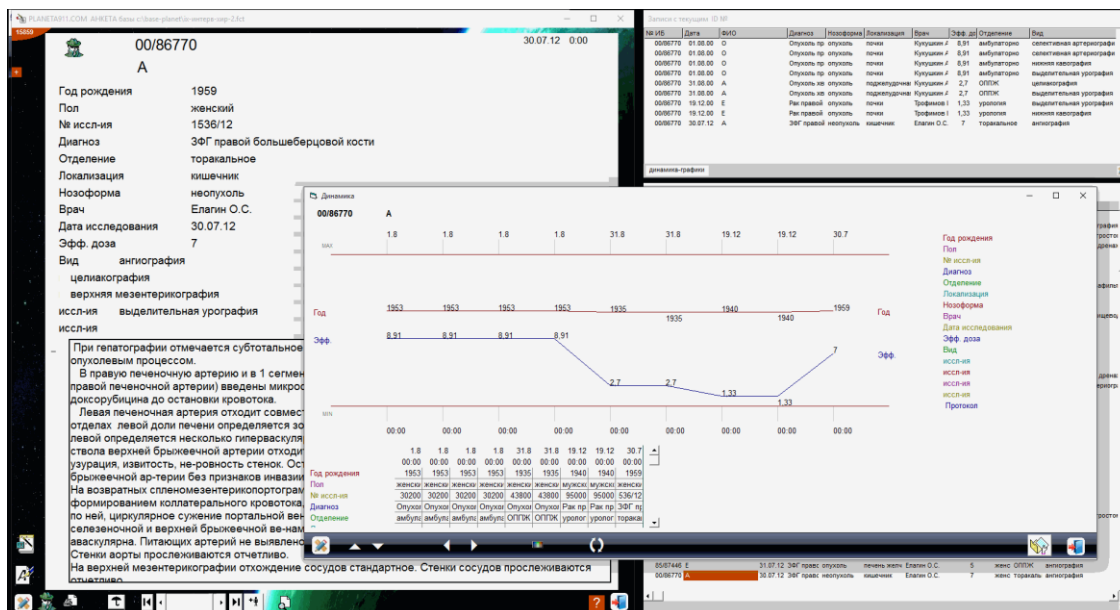
Institution Name: 80115
Manufacturer's Model Name: 37

Contrast/Bolus Agent: APPLIED
Gain/Detector Fil.: 0.00
Image Comment:
: 00:00
C : 00010

На цезиографии отхождение сосудов стандартное. Стенки артерий прослеживаются отчетливо. Непосредственно от черепного свода отходит сальниковая артерия, которая является одним из источников кровоснабжения опухоли, и ветви верхней брыжечной артерии отнесены влево. Отхождение сосудов стандартное. Стенки артерий прослеживаются отчетливо.

На возвратных спленомезентеро-портальных кровотоках по спленической, верхней брыжечной, а также по правой и левой ветвям портальной вены не нарушен. Стенки казанных сосудов прослеживаются отчетливо. Воротная вена растянута опухолью однако без нарушения кровотока по ней. возвратная портография катеризация и дренирование брюшной полости. Контроль нефростомы.

Междуанкетная динамика для записей с одним номером ID позволяет проследить динамику параметров между записями с одним регистрационным номером для последовательности обращений или сдачи анализов в лабораториях – можно переносить графики в отчётные протоколы.



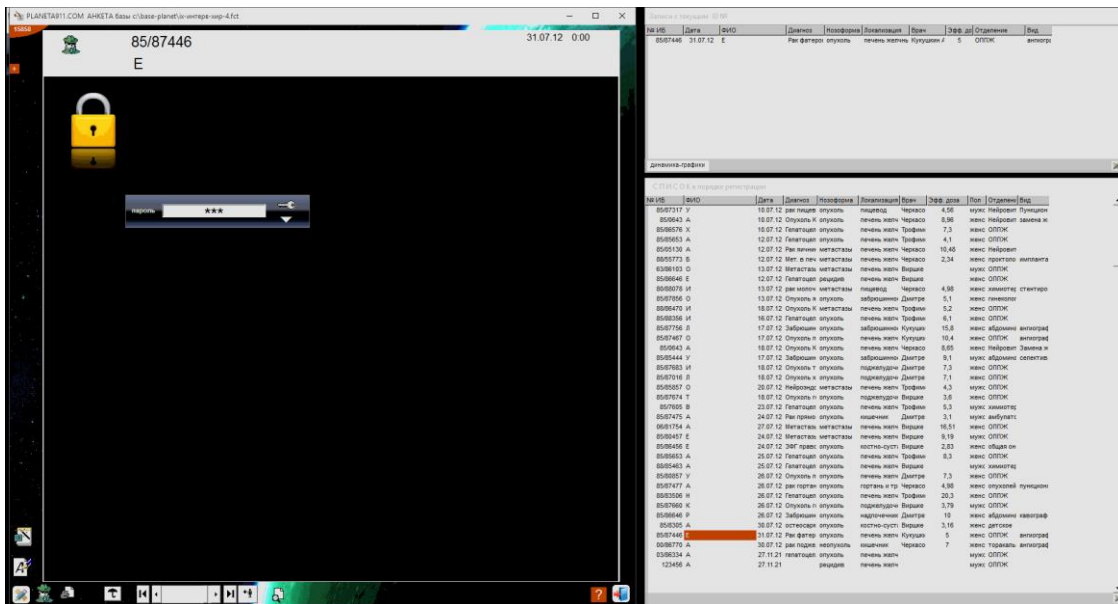
Меню готовых абзацев

Меню абзацев для вставки может добавлять сам пользователь, это позволяет уменьшить время создания документов (с юридически и технически отработанными шаблонами) с часов до минут.

The screenshot shows a medical software interface with the following components:

- Top Left:** Patient ID: 15/9046, Name: Трофимов И.А., Date of study: 3.11.2016, Doctor: Трофимов И.А.
- Center:** A text block containing a medical report in Russian, mentioning "Редукция селезеночного кровотока#" and "При гепатографии отмечается реваascularизация 1 очага в 4а сегменте".
- Right:** A menu titled "Просмотр абзацев" (View paragraphs) with a list of text fragments for insertion, including "Редукция селезеночного кровотока#", "Выделительная функция почек не...", and "Фистулография#".
- Bottom Right:** A table of laboratory results with columns: ФИО, История б, Дата ис, Врач, Вид, ИССЛ-Я.

В любой базе данных есть список исполнителей (врачей в медицине) – любой исполнитель может задать себе пароль и закрывать (открывать) данные своих записей (где он указан как исполнитель) от постороннего доступа до момента готовности материала или по другим причинам, если в этом есть необходимость (настройка и работа, см. видео).



Табличная форма

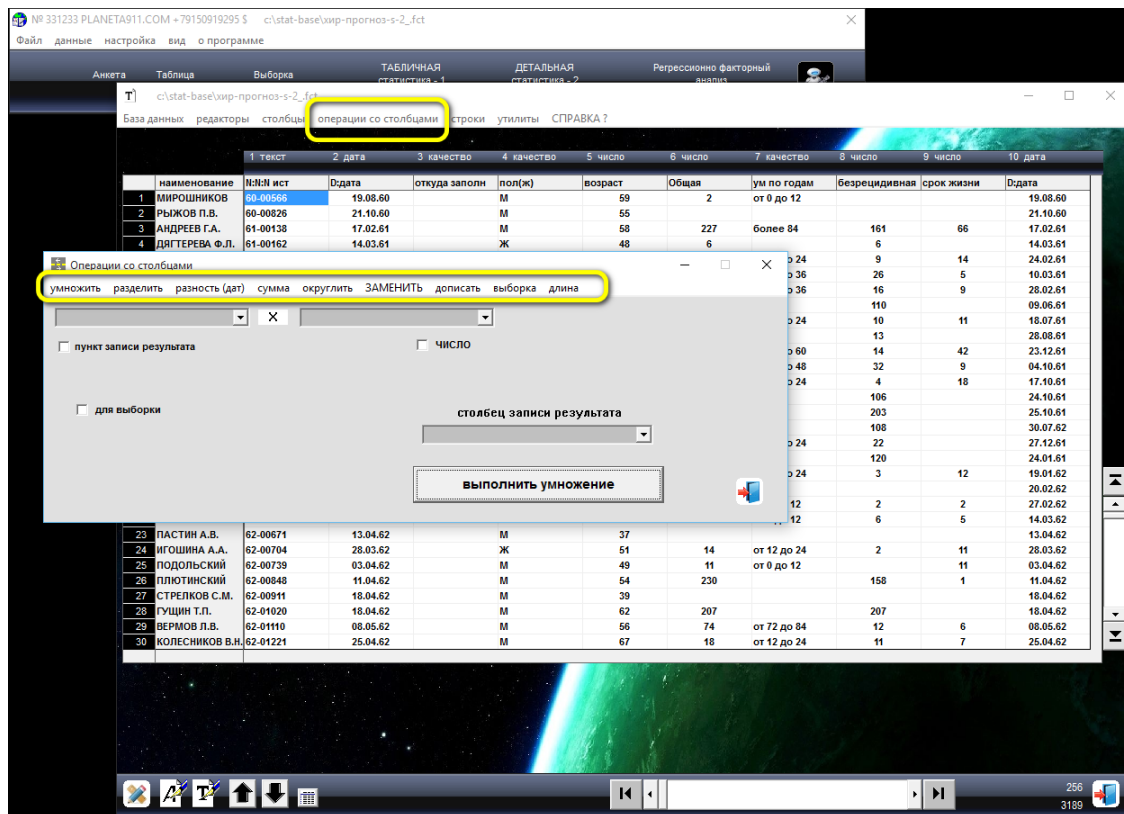
с анкетой данных создана для обеспечения процесса анализа данных.

The screenshot shows a data editor application with a table and a sidebar menu. Red arrows point from text boxes to specific parts of the interface:

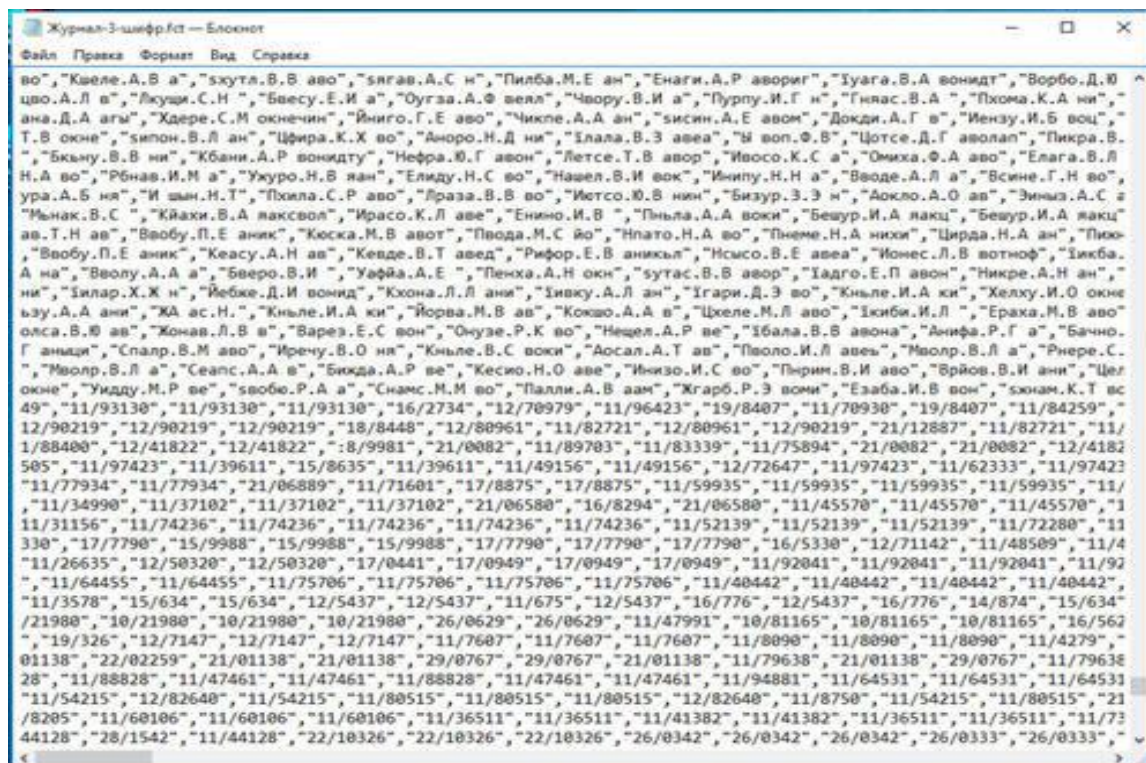
- Форма "Таблица данных"** (Data Table Form): points to the main data table.
- Анкета данных для качественных параметров - список качественных состояний** (Data Form for qualitative parameters - list of qualitative states): points to the table's columns.
- Операции со столбцами** (Operations with columns): points to the sidebar menu.
- Редактор анкеты** (Form Editor): points to the bottom of the sidebar menu.

The sidebar menu includes options like: ЧИСЛО, КАЧЕСТВО, ДАТА, ТЕКСТ, РИС (картинка), ВСТАВИТЬ столбец, УДАЛИТЬ столбец, Сделать копию столбца, ВСТАВИТЬ КОПИЮ столбца, Разложить качественные подпункты в конец анкеты (таблицы).

Форма «Операции со столбцами» позволяет делать любые операции преобразования и создания новых столбцов по выборкам данных (при наличии регистрационных анкет новые столбцы добавлять только в конец всей анкеты (таблицы)).



Защита информации – шифрование данных на физических носителях – нестандартное персональное динамическое шифрование.



Система полного 100-процентного автоматического восстановления данных при физических сбоях оборудования.

Database recovery - восстановление базы данных

При любом восстановлении будет сделана копия сбойной базы данных под именем даты и времени восстановления в папке базы данных

Archive copy database - архивные копии БД (БД - база данных)

1. FULL AUTOMATIC RECOVERY (name_av.fct - in folder DB)
Файл полного автоматического восстановления - name_av.fct - в папке БД

2. Дополнительная автоматическая копия базы на ВНЕШНИХ НОСИТЕЛЯХ (USB,...) для случая когда диск компьютера восстановлению не подлежит - укажите путь и включите опцию
C:\PLANETA\COPY_BASE
Задайте папку копий БД для восстановления

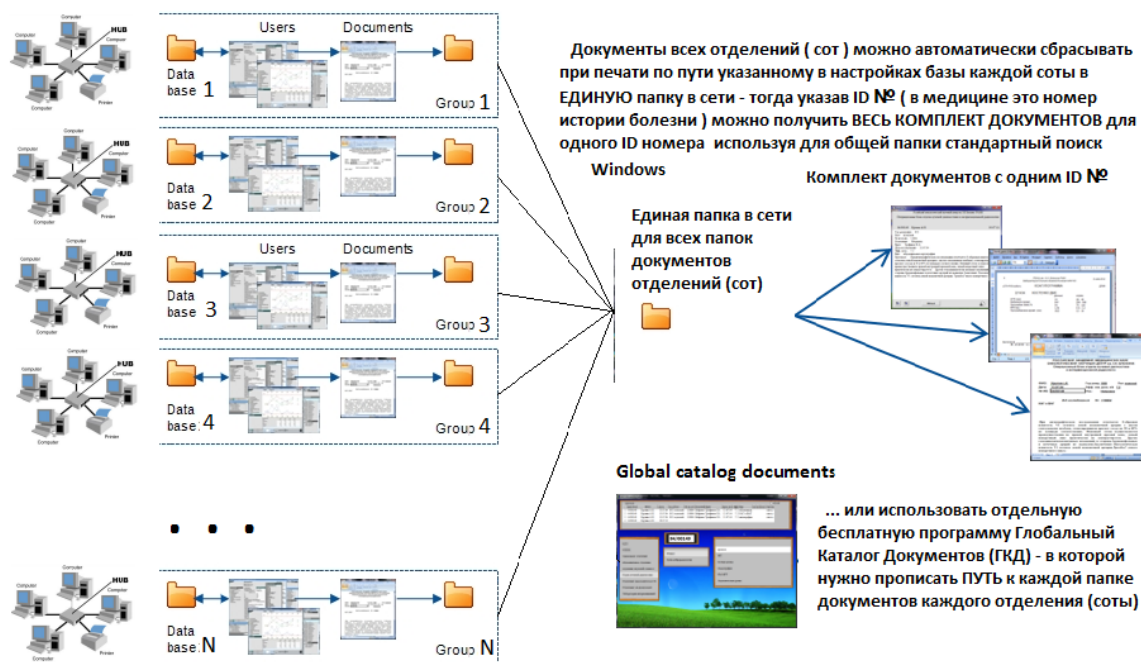
ВКЛ - Создание копии при завершении работы Date_name_EXIT.fct
за сколько дней сохранять копии

Примечание: При использовании опции < SAVE AS > настроек создается КОПИЯ старого варианта БД с именем БД+дата изменения (name_date_save.fct) - В ПАПКЕ БАЗЫ ДАННЫХ (в меню <Open>)

OFF(Выкл) автоматическое создание копии старой БД при исп. опции < SAVE AS >

4.3 Варианты сетевых структур сеть «ФАКТОР-ПРОГНОЗ» (PLANETA) + сеть SQL

1 уровень – «сотовый», только сеть СУБД «ФАКТОР-ПРОГНОЗ» (без SQL, Oracle) – лучший вариант работы по возможностям, скорости, надёжности и защищённости.



Документы (в формате PLANETA или в формате Excel) создаются автоматически при печати и сбрасываются в папку, указанную в настройках с именем ID № (в медицине это номер истории болезни – 2-й столбец базы данных) + приставка идентификатора отделения, которая тоже указывается в настройках сети.

Если вы в сети сделаете папку для всех отделений (баз-сот) и в ней зададите папки отделений (баз-сот), то, имея ярлык этой папки, вы сможете собирать весь комплект документов из этой папки в поиске проводника по ID № записи (истории болезни).

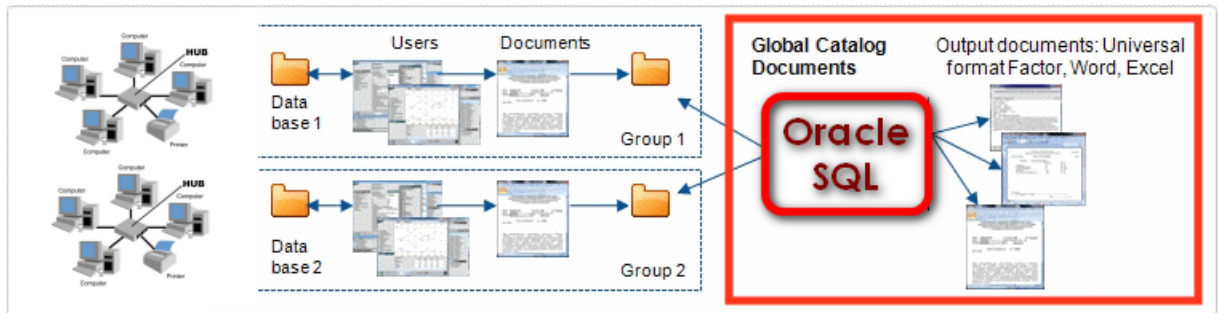
Но есть и отдельная программа как дополнение к «PLANETA» – Глобальный Каталог Документов (ГКД); в этом случае папки документов могут быть где угодно, но в настройках программы ГКД нужно будет прописывать названия отделений (сот) и сетевые пути к папкам документов отделений (сот).

Недостатки: необходимо прописывать настройки – названия отделений разделов и подразделов и пути к папкам отделений (сот).

Достоинства: папки документов могут находиться на компьютере в отделении (соты) и не зависеть от работы всей сети, и достигаемость на запись – 1 сетевой коммутатор отделения (соты) – не зависит от сбоев работы при превышении критической нагрузки файлового сервера всей сети (при сбоях одновременного доступа всех отделений которого нарушается работа с документами всех отделений).

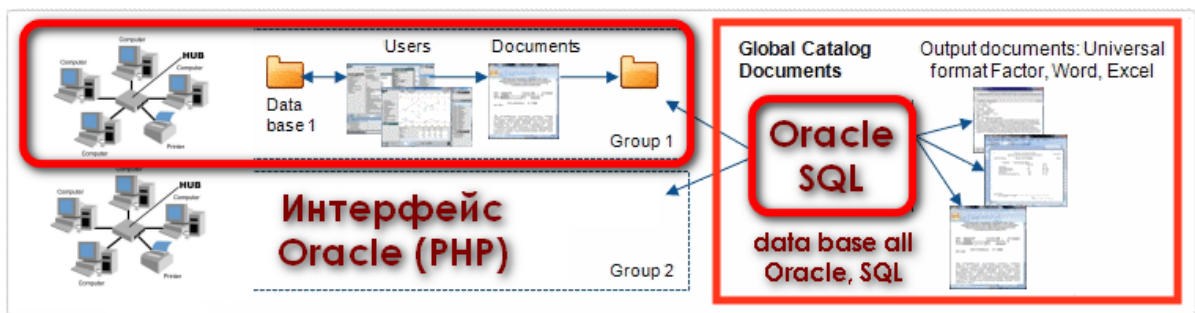
2 уровень – «глобальный» – «ФАКТОР-ПРОГНОЗ» + сервер SQL, Oracle:

– файловый сервер SQL только для просмотра (нет ввода SQL).



3 уровень – «смешанный» – «ФАКТОР-ПРОГНОЗ» + сервер SQL, Oracle:

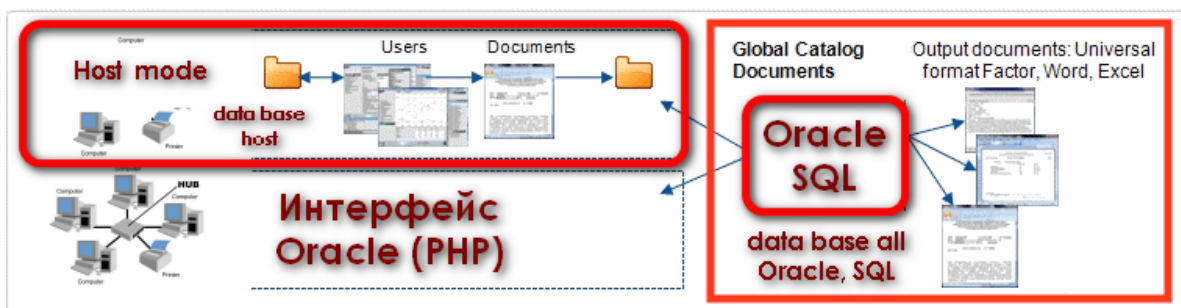
– с интерфейсом ведения данных SQL, Oracle (PHP).



4 уровень – «модуль» – сервер SQL, Oracle с интерфейсом ведения данных (PHP) + модуль «ФАКТОР-ПРОГНОЗ» (ведение данных и аналитика).

Дополнительный аналитический научный модуль «PLANETA» – подключённый как дополнительный модуль на 1 компьютер (для аналитической научной работы) к любой СУБД (SQL, Oracle...).

Есть кнопка сброса регистрационных данных в любую СУБД (SQL...).



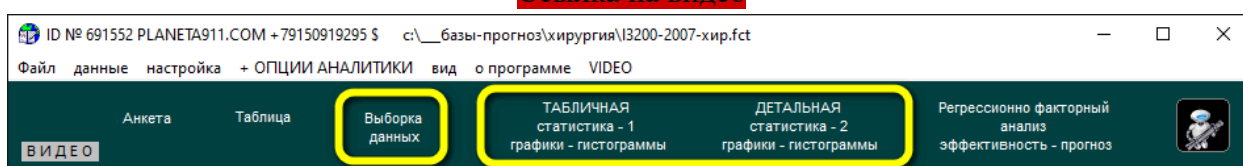
5. Традиционная однопараметрическая доказательная статистика «ФАКТОР-ПРОГНОЗ»

Может использоваться как дополнительное традиционное подтверждение решений РФА многопараметрических системных процессов.

Основное преимущество ПО «ФАКТОР-ПРОГНОЗ»- все расчёты для выборки данных можно получить сразу в единой таблице по нажатию одной кнопки, остаётся только выбрать нужные значения (не требуется тратить время на попарный перебор параметров для получения результатов, как в ПО «Statistica» или «SPSS»).

Общая схема однопараметрической доказательной статистики «ФАКТОР-ПРОГНОЗ»

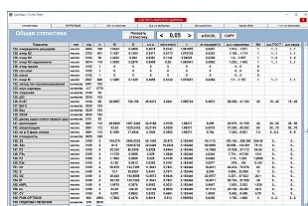
[Ссылка на видео](#)



Доказательная однопараметрическая статистика в таблицах и графиках
«Статистика 1» [Ссылка на видео](#)

Расчёты всех возможных комбинаций параметров по одной кнопке

1.1. Общая статистика



[Ссылка на видео](#)

1.2. Достоверность корреляции (графики)



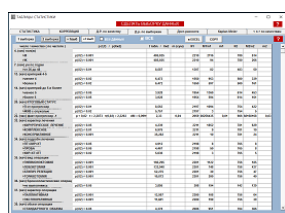
[Ссылка на видео](#)

1.3. Достоверность разности групп по качествам (гистограммы)



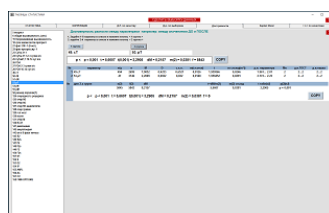
[Ссылка на видео](#)

1.4. Достоверной разности групп между двумя выборками



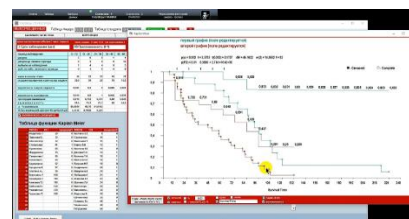
[Ссылка на видео](#)

1.5. Достоверная разность между двумя параметрами анкеты



[Ссылка на видео](#)

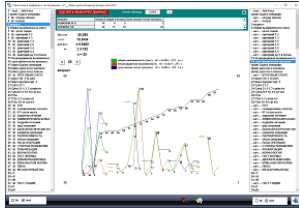
1.6. Таблицы вероятности и функции Каплан-Мейера



[Ссылка на видео](#)

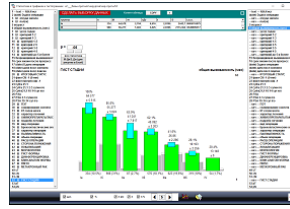
Детальная статистика для исследования «Статистика 2»

2.1. Многопараметрические графики корреляций



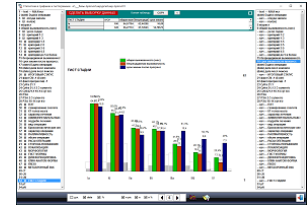
[Ссылка на видео](#)

2.2. Достоверность разности по качеству



[Ссылка на видео](#)

2.3. Многопараметрические гистограммы



[Ссылка на видео](#)

5.1. Возможности традиционной однопараметрической доказательной статистики «ФАКТОР-ПРОГНОЗ»

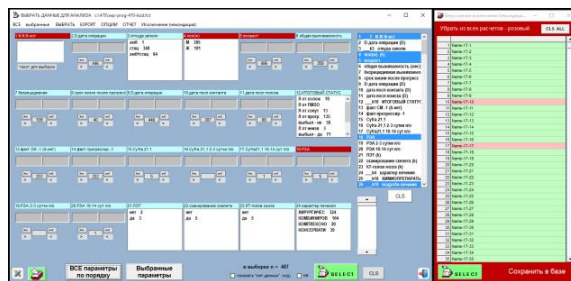
В аналитике комплекса планета есть два преимущества по отношению к другим системам аналитики типа SPSS, Statistica:

1. Система быстрых выборок данных по любым ограничениям;
2. Табличная форма результатов по всем возможным вариантам расчётов сразу по нажатию одной кнопки – не нужно тратить время на попарный перебор параметров, как в других системах типа SPSS, Statistica – максимальная экономия времени анализа.

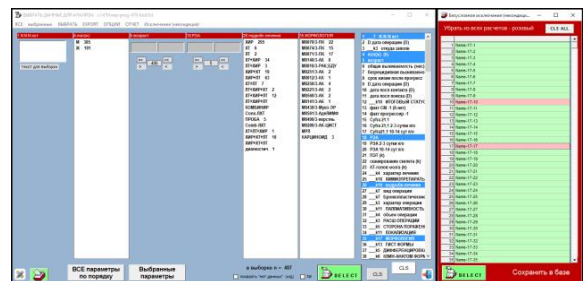
Если в «PLANETA» нет нужной математической опции, можно сделать копию таблицы любой выборки и перенести её в любое другое приложение Windows (Excel, SPSS, Statistica) – и в нём получить нужное решение.

Задание ограничений по параметрам.

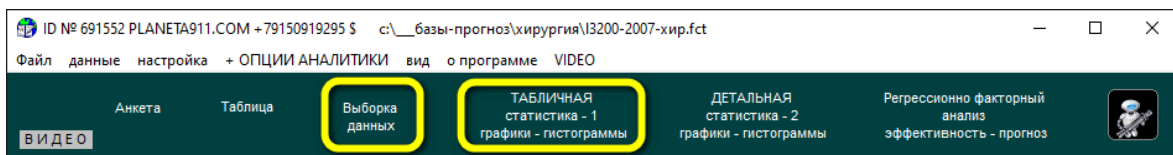
Безусловное исключение – красное окно



Можно выбрать только нужные параметры из всей базы для удобства работы с выборками



Ссылки с подробным описанием всех методов



ТАБЛИЦЫ СТАТИСТИКИ

ВЫБОРКА ДАННЫХ Таблицы Фишера 0,05 0,01 Таблица t-студента 0,9 0,001 Контроль ?...? ? video

БАЗОВАЯ СТАТИСТИКА Показать статистику ?

Параметры

Параметр	Мо	ди.
114. характеристика		
115. очередность	1	
116. очер N2	1	
117. очер N3	1	
118. очер N3 над	1	
119. очер график	0	
120. гол мозг	0	
121. кости	0	
122. очер M	1	
123. метод леч п		
124. опух маркер		
125. смирова		
126. p53	30	1
127. Ki-67		
128. Bcl-2		
129. Bax		
130. VEGF		
131. делец корог	40	3
132. диплоидия	60	5
133. анеуплоидия	7,5	6
134. кл в S-фазе		
135. плоидность		
136. S2	204,45	
137. S2a	75,15	
138. P2	54,65	
139. P2a	10,6	
140. F2	1,209	
141. F2a	0,135	
142. G1	62	
143. S	20,1	
144. G2	38	
145. IP	1,935	
146. ANPL	43,5	
147. AN	6,75	
148. CV	2	
149. FNM-OPTION		

Статистика комплекса PLANETA - это доказательная статистика на традиционных методах критериев Стьюдента и Фишера и новых возможностей регрессионно-факторного анализа по оптимизации и высокоточного прогноза

Ниже указан перечень используемых МИРОВЫХ КРИТЕРИЕВ в части доказательства значимости достоверности разности исследуемых групп по критериям Стьюдента и Фишера в науке и практике

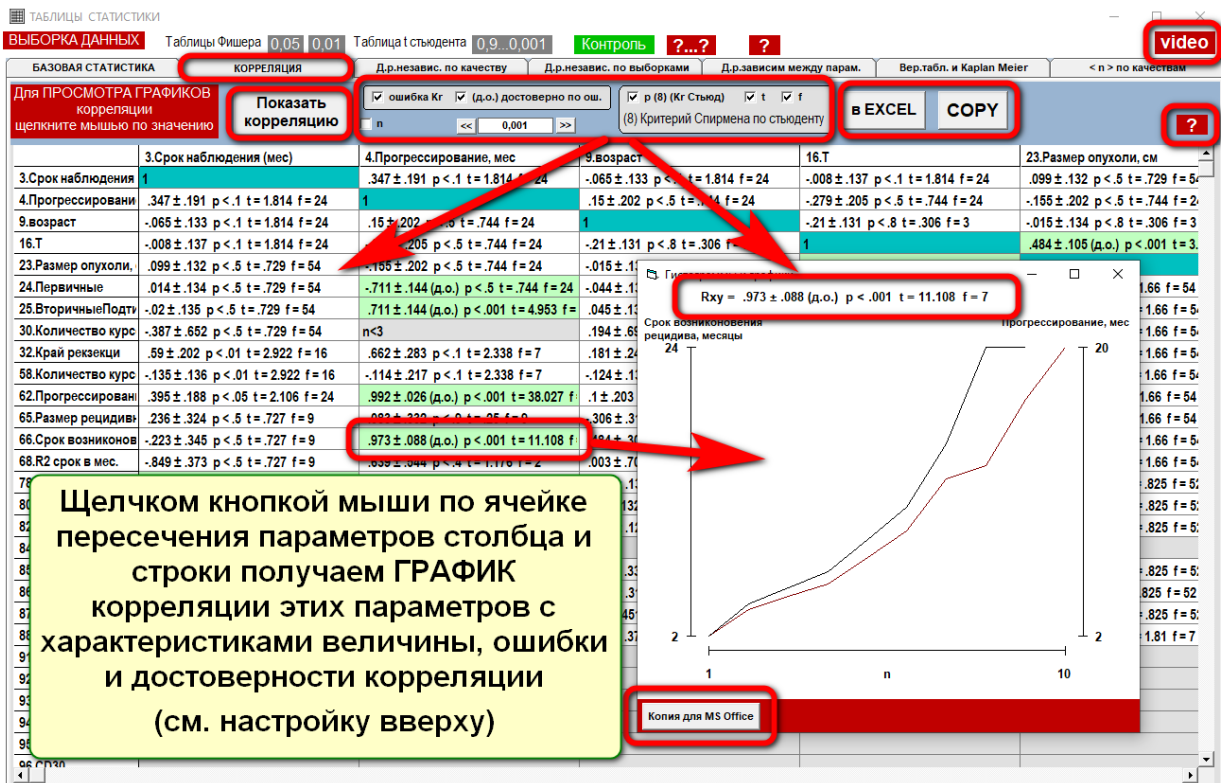
Критерии Стьюдента и Фишера - это МАКСИМУМ ВОЗМОЖНОСТЕЙ СТАТИСТИКИ по доказательству достоверной разности групп - если не используются весовые функции регрессионно-факторного анализа комплекса PLANETA

Ссылки в интернете на описание используемых математических оценочных критериев

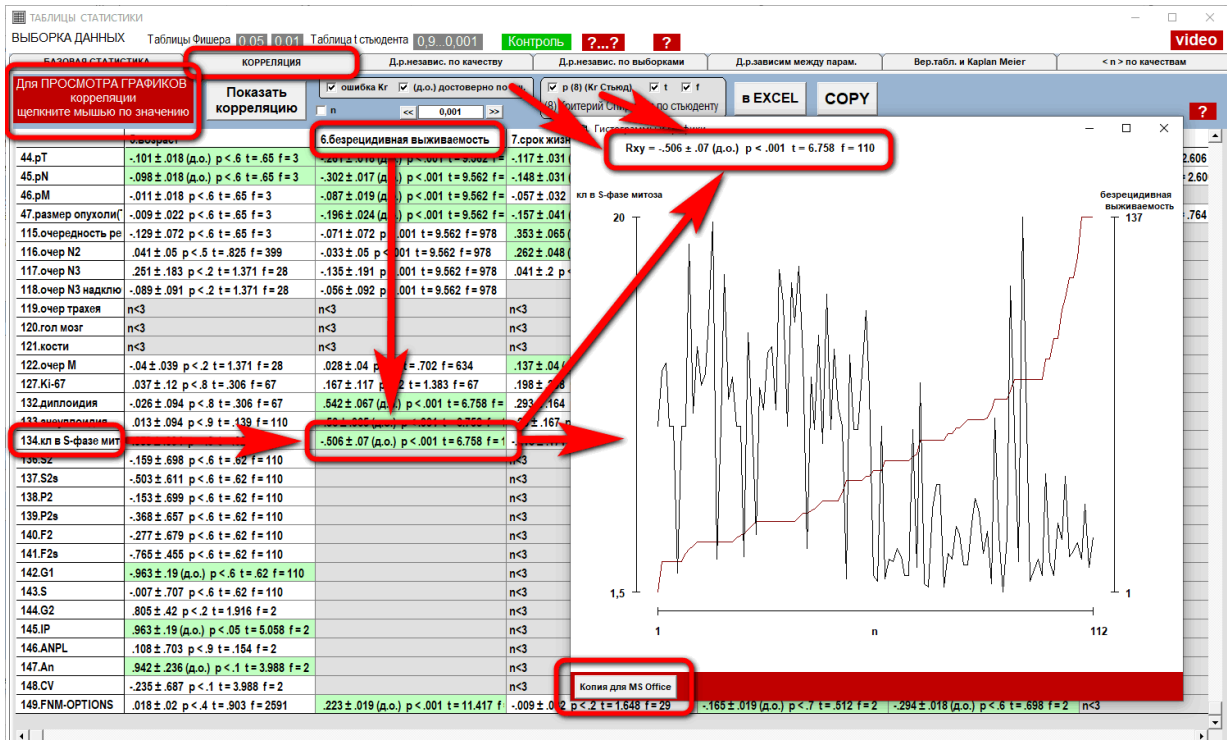
Критерий	Ссылка
ДОСТОВЕРНОСТЬ РАЗНОСТИ ГРУПП (1...6)	
по < СРЕДНИМ > (1,2)	
по < n > (3,4,5)	
(1) t- критерий p(c) 0,05 ... 0,001 для достоверной разности НЕЗАВИСИМЫХ совокупностей ГОССЕТА (Стьюдент):	medatstatistic.ru/methods/methods.html medatstatistic.ru/methods/methods14.html
(2) t- парный критерий Стьюдента для достоверной разности ЗАВИСИМЫХ совокупностей ГОССЕТА (Стьюдент):	medatstatistic.ru/methods/methods1.html lektsii.org/7-81924.html
(3) Достоверность разности долей p(n) НЕЗАВИСИМЫХ групп: по t- критерий Стьюдента	lektsii.org/7-44030.html
(4) Критерий Хи-квадрат ПИРСОНА НЕЗАВИСИМЫХ групп:	medatstatistic.ru/methods/method4.html
(5) Точный критерий Фишера НЕЗАВИСИМЫХ групп: точен для малых чисел наблюдений - точен для N групп < 10 (для групп с N < 10 критерий Хи-квадрат ПИРСОНА считается неточным)	medatstatistic.ru/methods/method5.html
(6) (1,2,3,4,5,6) Лекция АНАЛИЗ двух выборок	
(6) F- критерий Фишера для определения достоверности разности дисперсий НЕЗАВИСИМЫХ групп (по таблице 0,05...0,01)	taput.ru/res/informat/mop/lectons/lection_6.htm
по < ДИСПЕРСИЯМ СРЕДНИХ > Фишер (6)	
КОРРЕЛЯЦИЯ (7,8)	
(7) Критерий КОРРЕЛЯЦИИ ПИРСОНА:	medatstatistic.ru/methods/method8.html
(8) Критерий СПИРМЕНА статистической значимости коэффициента корреляции по t критерию Стьюдента:	medatstatistic.ru/methods/method9.html

5.1.2. Расчёт всех корреляций по выборкам с оценкой достоверности в таблицу с графиками и таблицами расчётов

Таблица корреляций с расчётами и готовыми графиками



Готовые таблицы и графики можно переносить в любые приложения Windows



5.1.3. Таблица расчёта достоверности разности между качественными состояниями качественных параметров по параметрам с гистограммами доверительных интервалов и таблицами результатов

При выборе качественного параметра в левом списке – получаем все расчёты.

При выборе кнопкой мыши параметра в таблице результатов расчёта – получаем готовую графическую картину гистограммы с доверительными интервалами для переноса в другие приложения (Word, Excel) для оформления.

Щелчком кнопкой мыши по наименованию параметра получаем гистограмму состояний с границами доверительных интервалов и таблицу расчётов

Можно подкорректировать параметры гистограммы и перенести гистограмму и таблицу в любое приложение Windows

Есть слайдовый вариант

Достоверность разности
Срок наблюдения (мес) $p = 0,05$ по параметру <G>
149,4
 $\pm 86,02$

73,67 $\pm 21,33$ 92,2 $\pm 50,3$

G1 G2 G3 Gx

3.6259 > 2.6896 k1=t4 k2=t1 / p(c) < 0,02 = 2.3266 (0,002) = 3.2526 dM = 104,9659 m(2) = 31,7078 f = 37
t = 0,5531 (0,6) = 0,5459 dM = 3,9

11,9275 > 5,9644 k1=t11 k2=t4 / p(c) < 0,1
t = 1,1772 (0,3) = 1,0512 dM = 1,2559
11,1956 > 6,3882 k1=t4 k2=t4 / p(c) < 0,4
t = 0,1453 (0,9) = 0,1278 dM = 0,0333
16,4077 > 4,018 k1=t4 k2=t3 / p(c) < 0,6
t = 0,1453 (0,9) = 0,1278 dM = 0,0333
16,4077 > 4,018 k1=t4 k2=t3 / p(c) < 0,6
t = 0,374 (0,8) = 0,2582 dM = 0,5273

Срок наблюдения (мес) | n | M | ско | m | t(0,05) | m*(0,05) | G1 | G2 | G3 | Gx | p(FD) / p(FD) | p(c) | p(c) | стьюдента средних | в EXCEL | COPY

Можно создавать дополнительные качественные состояния по выборкам для быстрого получения гистограмм многовариантных расчётов.

бессредственная выжимка: $p = 0,05$ по параметру <ГИСТ СТАДИИ>

62,44 $\pm 6,01$ 51,72 $\pm 3,12$ 39,64 $\pm 10,08$ 37,29 $\pm 3,35$ 23,28 $\pm 2,42$ 16,05 $\pm 2,41$ 8,08 $\pm 2,78$

Ia Ib IIa IIb IIIa IIIb IV

бессредственная выжимка | n | M | ско | m | t(0,05) | m*(0,05) | Ia | Ib | IIa | IIb | IIIa | IIIb | IV | p(FD) / p(FD) | p(c) | p(c) | стьюдента средних | в EXCEL | COPY

5.1.4. Таблица достоверности разности между двумя выборками

ТАБЛИЦЫ СТАТИСТИКИ
ВЫБОРКА ДАННЫХ Таблицы Фишера 0,05 0,01 Таблица t-студента 0,9 0,001 Контроль ?...? ?

БАЗОВАЯ СТАТИСТИКА КОРРЕЛЯЦИЯ Др.независ. по качеству Др.независ. по выборкам Др.зависим между парам. Вер.табл. и Kaplan Meier < n > по качеству

1 выборка 2 выборка << 0,001 >> в EXCEL COPY

число случаев (по частоте) (1) p(n) (L1) (3) p(n) долей p(n) Фишера точный крит.(5) / p(FD) Фиш дисп (6) (4) Хи2 Пирсона N1 M1 n1 m1 N2 M2 n2 m2

3. (кач) откуда зашли
 >стац p(n) < 0,005 t = 2,8477 (0,005) = 2,8133 dM = 0,1639 m(2) = 0,6365 f = 393
 >амб+стац p(n) < 0,005 t = 3,011 (0,005) = 2,8133 dM = 0,169 m(2) = 0,6362 f = 393

4. (кач) пол(ка)
 >М p(n) < 0,001 t = 14,3393 (0,001) = 3,2905 dM = 0,1737 m(2) = 0,0121 f = 3160
 >Ж p(n) < 0,01 t = 14,3393 (0,001) = 3,2905 dM = 0,1737 m(2) = 0,0121 f = 3160

5. (чис) ВОЗРАСТ
 p < 0,001 t = 7,7787 (0,001) = 3,2905 dM = 2,2826 m(2) = 0,2934 f = 3161

11. (кач) ИГОНОВЫЙ СТАТУС
 >л от ослож леч p(n) < 0,01 t = 2,7151 (0,01) = 2,5758 dM = 0,0217 m(2) = 0,008 f = 3117
 >л от ПМЗО p(n) < 0,5 t = 0,715 (0,5) = 1,0364 dM = 0,0064 m(2) = 0,0059 f = 3117
 >л от солуг забол p(n) < 0,5 t = 0,333 (0,5) = 0,1257 dM = 0,0019 m(2) = 0,0003 f = 3117
 >л от прогрессир p(n) < 0,05 t = 2,711 (0,05) = 1,96 dM = 0,037 m(2) = 0,018 f = 3117
 >выбыв - не прослежен p(n) < 0,05 t = 0,375 (0,5) = 0,2534 dM = 0,0029 m(2) = 0,0076 f = 3117
 >л от неизв причины p(n) < 0,1 t = 1,7866 (0,1) = 1,6449 dM = 0,0136 m(2) = 0,0076 f = 3117
 >выбыв - далее не просп p(n) < 0,02 t = 0,6751 (0,2) = 2,3263 dM = 0,0332 m(2) = 0,0129 f = 3117
 >ЖИВ без опухоли p(n) < 0,8 t = 0,2534 (0,8) = 0,2534 dM = 0,0029 m(2) = 0,0097 f = 3117
 >ЖИВ с опухолью p(n) < 0,05 t = 2,077 (0,5) = 1,96 dM = 0,0045 m(2) = 0,0022 f = 3117
 >ЖИВ без уточнений p(n) < 0,3 t = 1,2056 (0,3) = 0,364 dM = 0,0066 m(2) = 0,0054 f = 3117

20. (кач) ПЭТ
 >нет p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3
 >да p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3

21. (кач) сканирование скелета
 >да p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3

22. (кач) КТ-голов мозга
 >да p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3

23. (кач) характер лечения
 >ХИРУРГИЧЕСКОЕ ЛЕЧЕНИЕ p(n) < 0,2 t = 1,3479 (0,2) = 1,2816 dM = 0,0105 m(2) = 0,0044 f = 3129
 >КОМБИНИРОВАННОЕ ХИРУРГИЧЕСКОЕ ЛЕЧЕНИЕ p(n) < 0,4 t = 0,9627 (0,4) = 0,8416 dM = 0,0105 m(2) = 0,0044 f = 3129
 >КОМПЛЕКСНОЕ ЛЕЧЕНИЕ p(n) < 0,01 t = 2,6037 (0,01) = 2,5758 dM = 0,0105 m(2) = 0,0044 f = 3129
 >ХОНСЕРВАТИВНОЕ ЛЕЧЕНИЕ p(n) < 0,9 t = 0,2343 (0,9) = 0,1257 dM = 0,0019 m(2) = 0,0003 f = 3129

24. (кач) ХИМИОТЕРАПИЯ/ТАКСОЛ
 >евез100+нцс100 p(n) < 0,05 t = 2,0637 (0,05) = 2,0322 dM = 0,0105 m(2) = 0,0044 f = 3129
 >гемзар f+нцс100 p(n) < 0,5 t = 0,7408 (0,5) = 0,6818 dM = 0,0105 m(2) = 0,0044 f = 3129

25. (кач) подробн лечения
 >ХИР p(n) < 0,01 t = 2,7848 (0,01) = 2,5758 dM = 0,0098 m(2) = 0,0022 f = 3129
 >ХТ p(n) < 0,8 t = 0,3294 (0,8) = 0,2534 dM = 0,0098 m(2) = 0,0024 f = 3129
 >ЛТ p(n) < 0,6 t = 0,6219 (0,6) = 0,5244 dM = 0,0099 m(2) = 0,0015 f = 3129
 >ХТ+ХИР p(n) < 0,2 t = 1,5728 (0,2) = 1,2816 dM = 0,0095 m(2) = 0,0044 f = 3129
 >ХИР+ХТ p(n) < 0,2 t = 1,4608 (0,2) = 1,2816 dM = 0,0105 m(2) = 0,0072 f = 3129
 >ХИР+ЛТ p(n) < 0,6 t = 0,651 (0,6) = 0,5244 dM = 0,0051 m(2) = 0,0079 f = 3129
 >ЛТ+ХИР+ХТ p(n) < 0,6 t = 0,6034 (0,6) = 0,5244 dM = 0,0016 m(2) = 0,0027 f = 3129
 >ХТ+ХИР+ЛТ p(n) < 0,05 t = 2,3195 (0,05) = 1,96 dM = 0,0051 m(2) = 0,0022 f = 3129
 >ЛТ+ХИР+ЛТ p(n) < 0,05 t = 2,3955 (0,05) = 2,2807 dM = 0,0074 m(2) = 0,0025 f = 3129
 >КОМБИНИР p(n) < 0,05 t = 1,9974 (0,05) = 1,96 dM = 0,0036 m(2) = 0,0018 f = 3129
 >Сопв+ЛХТ p(n) < 0,9 t = 0,1763 (0,9) = 0,1257 dM = 0,0002 m(2) = 0,0013 f = 3129
 >ПРОБА p(n) < 0,001 t = 5,3998 (0,001) = 3,2905 dM = 0,0249 m(2) = 0,0046 f = 3129
 >Сопв+ЛХТ p(n) < 0,001 t = 4,4071 (0,001) = 3,2905 dM = 0,0164 m(2) = 0,0037 f = 3129
 >ХИР+ХТ+ЛТ p(n) < 0,7 t = 0,4659 (0,7) = 0,3853 dM = 0,0008 m(2) = 0,0018 f = 3129

26. (кач) вид операции
 >ПНЕВМОЭКТОМИЯ p(n) < 0,001 t = 5,8247 (0,001) = 3,3003 dM = 0,256 m(2) = 0,0439 f = 306
 >ХТ+ХИР p(n) < 0,8 t = 0,3177 (0,8) = 0,2534 dM = 0,0051 m(2) = 0,0162 f = 306
 >ЛТ+ХИР p(n) < 0,2 t = 1,5291 (0,2) = 1,2824 dM = 0,0241 m(2) = 0,0158 f = 306
 >ХИР+ЛТ p(n) < 0,001 t = 4,8552 (0,001) = 3,3003 dM = 0,1584 m(2) = 0,0326 f = 306

26. (кач) вид операции
 >ПНЕВМОЭКТОМИЯ p(n) < 0,001 t = 5,8247 (0,001) = 3,3003 dM = 0,256 m(2) = 0,0439 f = 306
 >ЛОБЭКТОМИЯ p(n) < 0,001 t = 8,6803 (0,001) = 3,3003 dM = 0,4475 m(2) = 0,0516 f = 307
 >БИ-ЛОБЭКТОМИЯ p(n) < 0,1 t = 1,6628 (0,1) = 1,6464 dM = 0,052 m(2) = 0,0313 f = 307
 >АТЛРИЯ РЕЗЕКЦИЯ p(n) < 0,5 t = 0,8217 (0,5) = 0,6747 dM = 0,0253 m(2) = 0,0308 f = 307
 >СЕКТОМТОМИЯ p(n) < 0,2 t = 1,4505 (0,2) = 1,2824 dM = 0,0213 m(2) = 0,0147 f = 307

27. (кач) Бронхопластические операц
 >не выполнялись p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3
 >радикальная p(n) < 0,001 t = 8,8783 (0,001) = 3,3003 dM = 0,3704 m(2) = 0,0417 f = 308

28. (кач) РАДИКАЛЬНАЯ ПАЛЛАТИВНОСТЬ
 >нет p(n) < 0,2 t = 1,4491 (0,2) = 1,2901 dM = 0,0909 m(2) = 0,0627 f = 56

29. (кач) объем операции
 >СТАНДАРТНОГО ОБЪЕМА p(n) < 0,001 t = 5,2297 (0,001) = 3,3003 dM = 0,3004 m(2) = 0,0574 f = 272
 >РАСШИРЕННАЯ p(n) < 0,02 t = 2,5663 (0,02) = 2,3301 dM = 0,159 m(2) = 0,062 f = 272

31. (кач) РАСШ ОПЕРАЦИИ
 >по вынужд показанием p(n) < 0,005 t = 2,9407 (0,005) = 2,8707 dM = 0,346 m(2) = 0,1177 f = 63
 >по принцип показанием p(n) < 0,005 t = 2,9407 (0,005) = 2,8707 dM = 0,346 m(2) = 0,1177 f = 63

32. (кач) СТОРОНА ПОРАЖЕНИЯ
 >правая p(n) < 0,3 t = 1,2733 (0,3) = 1,037 dM = 0,072 m(2) = 0,0565 f = 308
 >левая p(n) < 0,3 t = 1,2733 (0,3) = 1,037 dM = 0,072 m(2) = 0,0565 f = 308

33. (кач) ЛОКАЛИЗАЦИЯ
 >нижняя доля p(n) < 0,8 t = 0,291 (0,8) = 0,2534 dM = 0,0161 m(2) = 0,0553 f = 308
 >средняя доля p(n) < 0,2 t = 1,4723 (0,2) = 1,2824 dM = 0,0476 m(2) = 0,0323 f = 308

Достоверность разности между выборкам
 ТАБЛИЦА ПО ВСЕМ ПАРАМЕТРАМ
 Делаете 1 выборку и нажимаете кнопку "1 выборка"
 Делаете 2 выборку и нажимаете кнопку "2 выборка"

Достоверность разности независимых групп (д. р.) p (n) по Стьюденту, Точный критерий Фишера, Хи-квадрат Пирсона.

Для качественных параметров достоверность разности по n

ТАБЛИЦЫ СТАТИСТИКИ
ВЫБОРКА ДАННЫХ Таблицы Фишера 0,05 0,01 Таблица t-студента 0,9 0,001 Контроль ?...? ?

БАЗОВАЯ СТАТИСТИКА КОРРЕЛЯЦИЯ Др.независ. по качеству Др.независ. по выборкам Др.зависим между парам. Вер.табл. и Kaplan Meier < n > по качеству

1 выборка 2 выборка << 0,001 >> в EXCEL COPY

число случаев (по частоте) (1) p(n) (L1) (3) p(n) долей p(n) Фишера точный крит.(5) / p(FD) Фиш дисп (6) (4) Хи2 Пирсона N1 M1 n1 m1 N2 M2 n2 m2

3. (кач) откуда зашли
 >стац p(n) < 0,05 t = 2,0848 (0,05) = 1,984 dM = 0,2047 m(2) = 0,9982 f = 63
 >амб+стац p(n) < 0,05 t = 2,0848 (0,05) = 1,984 dM = 0,2047 m(2) = 0,9982 f = 63

4. (кач) пол(ка)
 >Ж p(n) < 0,001 t = 7,9759 (0,001) = 3,3003 dM = 41,5443 m(2) = 5,1837 f = 245

6. (чис) безрецидивная выживаемость
 p < 0,001 t = 7,9759 (0,001) = 3,3003 dM = 41,5443 m(2) = 5,1837 f = 245

11. (кач) ИГОНОВЫЙ СТАТУС
 >л от ослож леч p(n) < 0,05 t = 2,3082 (0,05) = 1,9823 dM = 0,0461 m(2) = 0,02 f = 308
 >л от ПМЗО p(n) < 0,2 t = 1,4455 (0,2) = 1,2824 dM = 0,0212 m(2) = 0,0146 f = 308
 >л от солуг забол p(n) < 0,05 t = 2,2712 (0,05) = 1,9823 dM = 0,048 m(2) = 0,0212 f = 308
 >л от прогрессир p(n) < 0,001 t = 7,1304 (0,001) = 3,3003 dM = 0,375 m(2) = 0,0526 f = 308
 >выбыв - не прослежен p(n) < 0,01 t = 2,6487 (0,01) = 2,5908 dM = 0,0595 m(2) = 0,0225 f = 308
 >выбыв - далее не просп p(n) < 0,001 t = 4,0891 (0,001) = 3,3003 dM = 0,1562 m(2) = 0,0382 f = 308
 >ЖИВ без опухоли p(n) < 0,001 t = 3,4763 (0,001) = 3,3003 dM = 0,1196 m(2) = 0,0344 f = 308
 >ЖИВ с опухолью p(n) < 0,2 t = 1,521 (0,2) = 1,2824 dM = 0,0239 m(2) = 0,0157 f = 308
 >ЖИВ без уточнений p(n) < 0,1 t = 1,7166 (0,1) = 1,6464 dM = 0,0269 m(2) = 0,0157 f = 308

12. (чис) факт см -1 (0-нет)
 p < 0,001 t = 6,9142 (0,001) = 3,3003 dM = 0,3284 m(2) = 0,0475 f = 263

13. (чис) факт прогрессир -1
 p < 0,001 t = 6,7341 (0,001) = 3,3003 dM = 0,3573 m(2) = 0,0531 f = 265

21. (кач) характер лечения
 >ХИРУРГИЧЕСКОЕ ЛЕЧЕНИЕ p(n) < 0,001 t = 4,7393 (0,001) = 3,3003 dM = 0,193 m(2) = 0,0407 f = 308
 >КОМБИНИРОВАННОЕ ХИРУРГИЧЕСКОЕ ЛЕЧЕНИЕ p(n) < 0,001 t = 4,3102 (0,001) = 3,3003 dM = 0,1708 m(2) = 0,0396 f = 308

25. (кач) подробн лечения
 >ХИР p(n) < 0,001 t = 5,8247 (0,001) = 3,3003 dM = 0,256 m(2) = 0,0439 f = 306
 >ХТ+ХИР p(n) < 0,8 t = 0,3177 (0,8) = 0,2534 dM = 0,0051 m(2) = 0,0162 f = 306
 >ЛТ+ХИР p(n) < 0,2 t = 1,5291 (0,2) = 1,2824 dM = 0,0241 m(2) = 0,0158 f = 306
 >ХИР+ЛТ p(n) < 0,001 t = 4,8552 (0,001) = 3,3003 dM = 0,1584 m(2) = 0,0326 f = 306

26. (кач) вид операции
 >ПНЕВМОЭКТОМИЯ p(n) < 0,001 t = 5,8247 (0,001) = 3,3003 dM = 0,256 m(2) = 0,0439 f = 306
 >ЛОБЭКТОМИЯ p(n) < 0,001 t = 8,6803 (0,001) = 3,3003 dM = 0,4475 m(2) = 0,0516 f = 307
 >БИ-ЛОБЭКТОМИЯ p(n) < 0,1 t = 1,6628 (0,1) = 1,6464 dM = 0,052 m(2) = 0,0313 f = 307
 >АТЛРИЯ РЕЗЕКЦИЯ p(n) < 0,5 t = 0,8217 (0,5) = 0,6747 dM = 0,0253 m(2) = 0,0308 f = 307
 >СЕКТОМТОМИЯ p(n) < 0,2 t = 1,4505 (0,2) = 1,2824 dM = 0,0213 m(2) = 0,0147 f = 307

27. (кач) Бронхопластические операц
 >не выполнялись p(n) < 0,8 t = 0,2774 (0,8) = 0,2774 dM = 0,1667 m(2) = 0,6009 f = 3
 >радикальная p(n) < 0,001 t = 8,8783 (0,001) = 3,3003 dM = 0,3704 m(2) = 0,0417 f = 308

28. (кач) РАДИКАЛЬНАЯ ПАЛЛАТИВНОСТЬ
 >нет p(n) < 0,2 t = 1,4491 (0,2) = 1,2901 dM = 0,0909 m(2) = 0,0627 f = 56

29. (кач) объем операции
 >СТАНДАРТНОГО ОБЪЕМА p(n) < 0,001 t = 5,2297 (0,001) = 3,3003 dM = 0,3004 m(2) = 0,0574 f = 272
 >РАСШИРЕННАЯ p(n) < 0,02 t = 2,5663 (0,02) = 2,3301 dM = 0,159 m(2) = 0,062 f = 272

31. (кач) РАСШ ОПЕРАЦИИ
 >по вынужд показанием p(n) < 0,005 t = 2,9407 (0,005) = 2,8707 dM = 0,346 m(2) = 0,1177 f = 63
 >по принцип показанием p(n) < 0,005 t = 2,9407 (0,005) = 2,8707 dM = 0,346 m(2) = 0,1177 f = 63

32. (кач) СТОРОНА ПОРАЖЕНИЯ
 >правая p(n) < 0,3 t = 1,2733 (0,3) = 1,037 dM = 0,072 m(2) = 0,0565 f = 308
 >левая p(n) < 0,3 t = 1,2733 (0,3) = 1,037 dM = 0,072 m(2) = 0,0565 f = 308

33. (кач) ЛОКАЛИЗАЦИЯ
 >нижняя доля p(n) < 0,8 t = 0,291 (0,8) = 0,2534 dM = 0,0161 m(2) = 0,0553 f = 308
 >средняя доля p(n) < 0,2 t = 1,4723 (0,2) = 1,2824 dM = 0,0476 m(2) = 0,0323 f = 308

5.1.5. Расчёт статистических характеристик между параметрами

Зависимая (и независимая) достоверность разности между значениями параметров до и после воздействия

Достоверность разности по средним между РАЗНЫМИ ПАРАМЕТРАМИ по ЗАВИСИМЫМ (параметры ДО и ПОСЛЕ) и НЕЗАВИСИМЫМ совокупностям

1. Сделайте выборку и задайте 1-й параметр в списке и нажмите кнопку < 1 группа >
2. Сделайте выборку и задайте 2-й параметр в списке и нажмите кнопку < 2 группа >

ЗАВИСИМЫЕ совокупности (2)

1 параметр - совокупность ДО: 40. СТ
2 параметр - совокупность ПОСЛЕ: 44. рТ

$p < .001$ $dM = .212$ $m = .012$ $t = 18.165$ $f = 2661$ $ско = .603$ $n = 2662$

НЕЗАВИСИМЫЕ совокупности (1, 6)

1 группа - выборка: 41. сN
2 группа - выборка: 45. рN

$p(c) < p < 0,001$ $t = 6,9264$ $t(0,001) = 3,2905$ $dM = 0,1533$ $m(2) = 0,0221$ $f = 5950$

№	параметр	н/д	n	M	D	с.к.о.	m	t	m * t	д.и.	Me	д.и. ГОСТ	д.и. наука
1	41.сN	387	2806	0,593	0,6892	0,8302	0,0157	1,959539	0,0307	.562... .624	0	0...0	0...0
2	45.рN	47	3146	0,7463	0,7643	0,8742	0,0156	1,959452	0,0305	.716... .777	0	0...0	0...0

для 2-х групп: $n(2) = 5952$, $f(2) = 5950$, $dM = 0,1533$, $t = 6,9264$, $t(2) = 3,2905$, $m(2) = 0,0221$, $f = 5950$, $p(c) < p < 0,001$

Если в базе данных есть параметры "ДО" и "ПОСЛЕ" воздействия - то можно посчитать достоверность разности для "ЗАВИСИМЫХ" значений параметров

5.1.6. Вероятностные таблицы (выживаемость, вероятность отказа) и функции Каплан – Мейера (с расчётом достоверной разности)

срок наступления события (числ. парам): 6 безрецидивная выживаемость
числ. парам. 0 или 1 (0 - не прослежено): 12 факт CM -1 (0-лет)

Интервал: 12

Каплан-Мейер

$p(c) < p < 0,001$ $t = 3,8771$ $t(0,001) = 3,3003$ $dM = 22,8023$ $m(2) = 5,8813$ $f = 360$
 $p(FD) < 0,01$ $2,7475 > 1,473$ $k1=63$ $k2=297$

Таблица функции Каплан-Мейер

Имя	(1)	продолж(1)	Факт(1)	(2)	продол
ШАДРИН В.	1	1	ХАЙМОВ С.	1	1
БОЙЦОВ Г.С.	1	1	САФАТОВ Е.	2	1
ТИХОНОВ Ф.	2	1	ПОДНЕБЕСИ	2	1
ИГОШИНА Л.	2	1	ТОРОСЯН П.	2	1
ТАРАСОВ И.	2	1	РАБИНОВИЧ	3	1
КОРНЕЦ Н.	3	1	БИРКОВ И.	4	1
МЕДВЕДЕВ	3	1	ГУЗЕВ В.И.	5	1
РУБЕЛЬТ А.	3	1	ПАРХОМЕН	6	1
ШУМАКЕР Л.	3	1	ДАГТЕРЕВА	6	1
БАЛАБАНА И.	4	0	КАМЫШНИ	6	1
СЕДОВ С.П.	4	1	ЛЯДОНОВ И.	7	1
ЕРШОВ Н.И.	4	1	ЛАДЫЖЕНК	7	1
ЛОКТЕВА М.	4	1	ВАСИЛЕНК	10	1
ШУМАКЕР Л.	5	1	ЦУКЕРМАН	11	1
ТРУСОВ Ю.	5	1	КЛОПЕ Г.Л.	12	1
СОУХАНОВ Е.	6	1	БАЛАШОВ	12	1
КОЛЬЦОВ Е.	6	1	КОРЕЛЯН И.	12	1
БАЛАБАНА И.	6	1	БАЛАШОВ	13	1
ТРИКОНОВ	7	1	РЯЗАНДЕМ	13	0
ВОЛКОВ Н.Л.	7	1	АКАТОВА Т.	14	1
МИРЯН Г.П.	8	1	ШУСКИН И.	14	1
СКУРАТОВ	8	1	ЛУЗИН Н.А.	15	1
СИНЕВ П.Х.	8	1	БОХАНОВ В.	17	1
МОРОЗОВ Г.	9	1	ЖЕЛЕНОВ	18	1
ТРИКОНОВ И.	9	0	ИЛЬИН С.С.	19	1
МАЧАЛОВ Н.	9	1	ПЕШЧЕНКО	19	1
АГАПОВ А.Л.	9	1	ЗАМЧЕНКО	19	1

В базе данных нужно указать здесь два числовых столбца: 1 столбец – срок события и 2 числовой столбец – прослеженность события (0 – не прослежено: 1 – прослежено).

5.1.7. Количественные таблицы по выборкам

ТАБЛИЦЫ СТАТИСТИКИ

ВЫБОРКА ДАННЫХ Таблицы Фишера 0,05 0,01 Таблица t студента 0,9...0,001 Контроль ?...? ?

БАЗОВАЯ СТАТИСТИКА КОРРЕЛЯЦИЯ Д.р.независ. по качеству Д.р.независ. по выборками Д.р.зависим между парам. Вер.табл. и Kaplan Meier < p > по качествам

3 откуда заполни
4 пол(ж)

11 ИТОВОГО СТАТУС

20 ПЭТ
21 сканирование скелета
22 КТ-голов мозга
23 характер лечения
25 подробн лечения
27 Бронхопластические операции
28 характер операции
29 ПАЛПАТИВНОСТЬ
30 объем операции
31 РАСШ ОПЕРАЦИИ
32 СТОРОНА ПОРАЖЕНИЯ
33 ЛОКАЛИЗАЦИЯ
34 МОРФОЛОГИЯ
35 ГИСТ ФОРМЫ
36 ДИФФЕРЕНЦИРОВКА

подробн лечения\ви	ПНЕВМОН	ЛОБЭКТС	БИ-ЛОБЭ	АТИПИЧ F	ТОРАКОТ	СЕМГЕНТ	РЕЗКЦИЯ	ВСЕГО	ВСЕГО(%)
ХИР	1066	1077	109	67	155	21		2495	79,43
ХТ	2	1			6			9	0,28
ЛТ				1	2			3	0,09
ХТ+ХИР	29	14	2					45	1,43
ЛТ+ХИР	57	47	6	4	10			124	3,94
ХИР+ХТ	62	54	5	4	2	1		128	4,07
ХИР+ЛТ	75	67	11	6	1	2		162	5,15
ХТ+ЛТ		1			3			4	0,12
ЛТ+ХИР+ХТ	6	11			1			18	0,57
ХТ+ХИР+ЛТ	10	1	1		1			13	0,41
ЛТ+ХИР+ЛТ	9	1			7			17	0,54
КОМБИНИР	3	3	1					7	0,22
Сопз-ЛХТ		1		1	2			4	0,12
ПРОБА		1		1	58			60	1,91
Сomb-ЛХТ	30	3		1	5			39	1,24
ХТ+ЛТ+ХИР	1				2			3	0,09
ХИР+ХТ+ЛТ	8	1	1					10	0,31
ХИР+ХТ+ЛТ								0	
диагностическая опе								0	
ВСЕГО	1358	1283	136	85	255	24		3141	
ВСЕГО(%)	43,23	40,84	4,32	2,7	8,11	0,76			100

11 ИТОВОГО СТАТУС

20 ПЭТ
21 сканирование скелета
22 КТ-голов мозга
23 характер лечения
24 ХИМИОПРЕПАРАТЫТАКСОЛ
26 вид операции
27 Бронхопластические операции
28 характер операции
29 ПАЛПАТИВНОСТЬ
30 объем операции
31 РАСШ ОПЕРАЦИИ
32 СТОРОНА ПОРАЖЕНИЯ
33 ЛОКАЛИЗАЦИЯ
34 МОРФОЛОГИЯ
35 ГИСТ ФОРМЫ
36 ДИФФЕРЕНЦИРОВКА
37 ПЛАН-АНАТОМ ФОРМА
38 ПМЗО
39 МЕТАХРОННЫЙ РАК
43 ГИСТ СТАДИИ

5.2. Детальная статистика

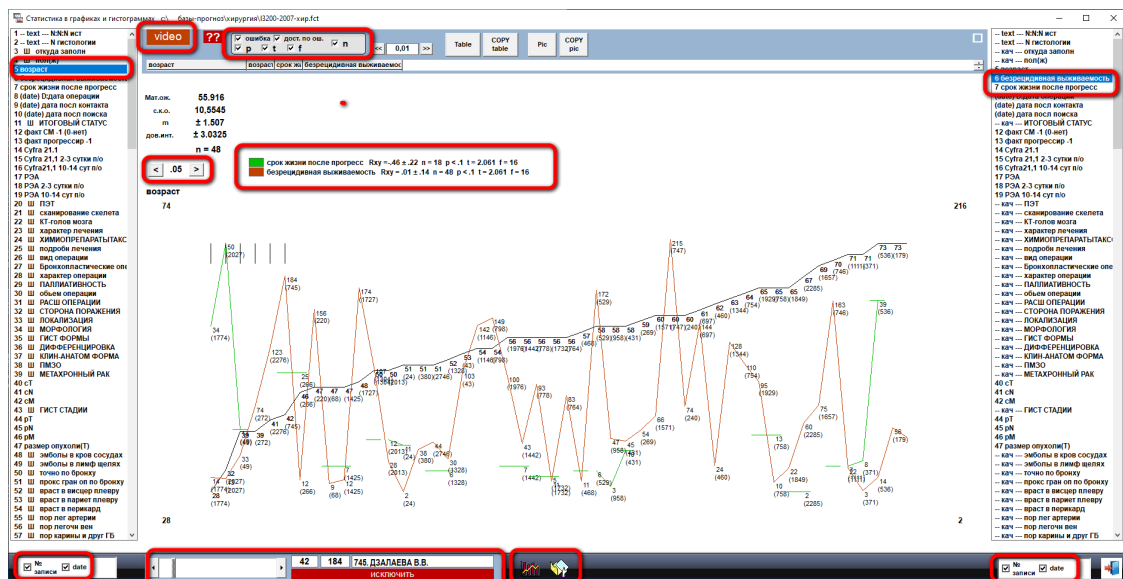
ID № 691552 PLANETA911.COM +79150919295 \$ c:_базы-прогноз\хирургия\3200-2007-хир.фс

Файл данные настройка + ОПЦИИ АНАЛИТИКИ вид о программе VIDEO

Анкета Таблица **Выборка данных** ТАБЛИЧНАЯ статистика - 1 графики - гистограммы **ДЕТАЛЬНАЯ статистика - 2** графики - гистограммы Регрессионно факторный анализ эффективность - прогноз

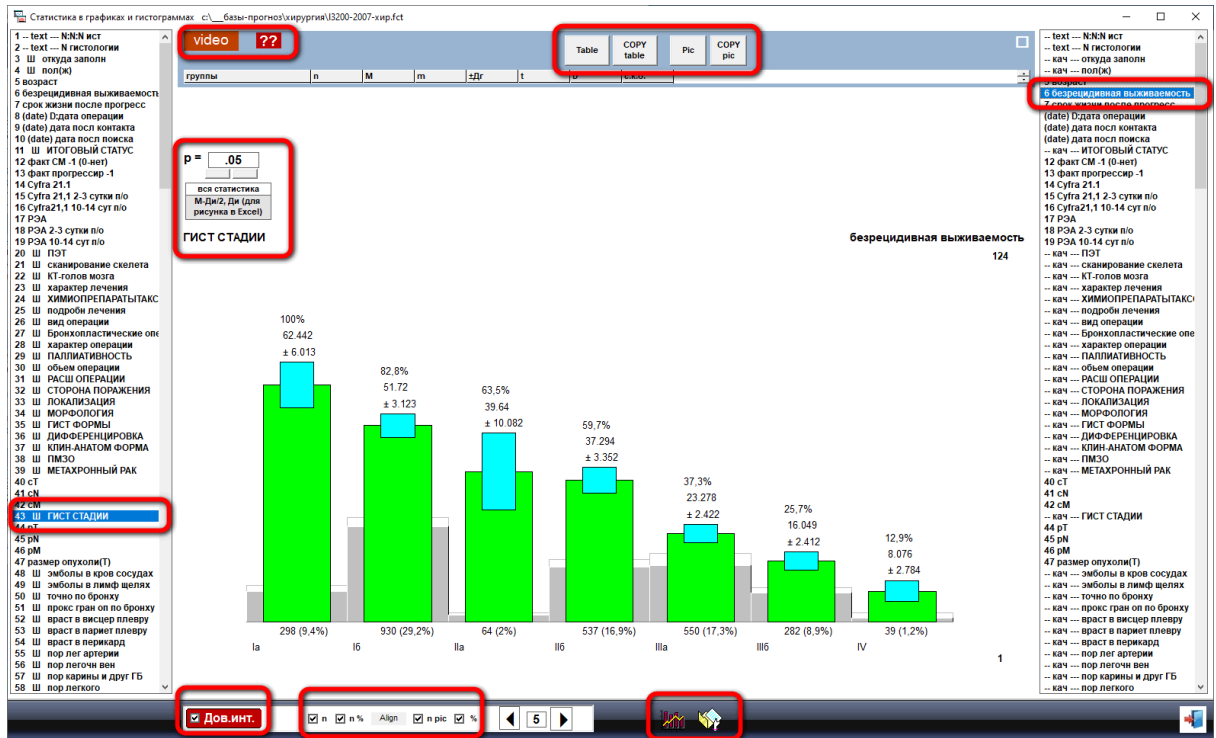
5.2.1. Исследования по выборкам корреляций и графиков

Эту форму с графиками удобно использовать для обнаружения ошибочно набранных значений в базе данных



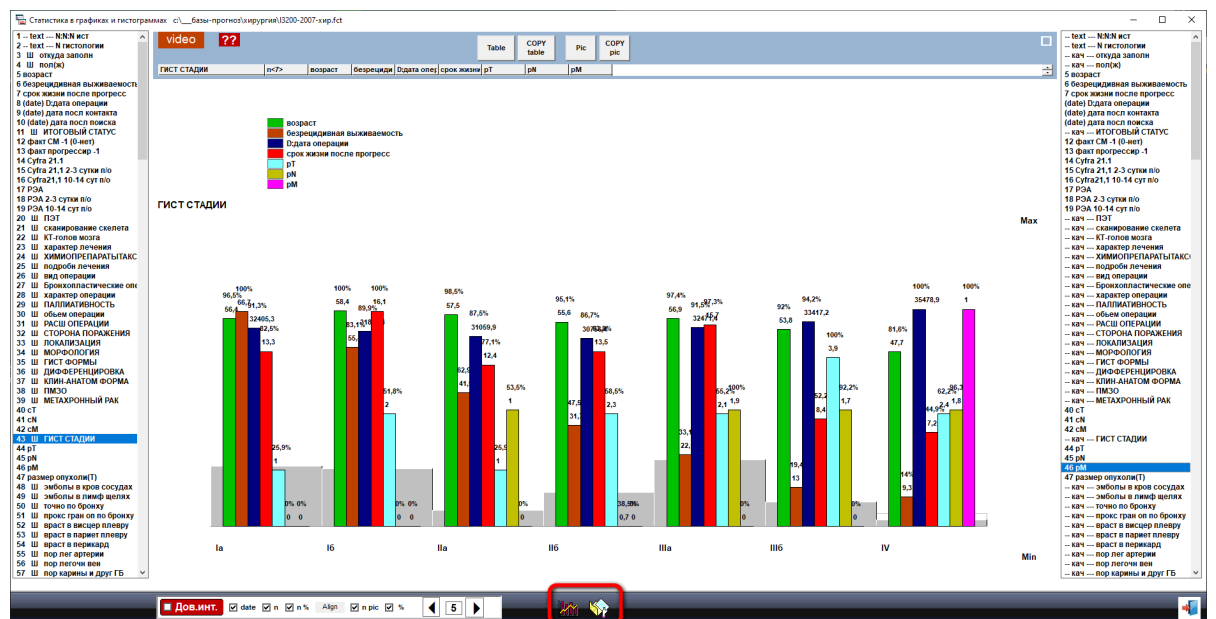
5.2.2. Детальное исследование достоверности разности

Для качественных параметров гистограммы по выборкам удобный режим графической визуализации достоверной разности групп по качественным параметрам имеет графическую форму для построения готовых слайдов:



5.2.3. Многопараметрические гистограммы

Режим построения многопараметрических гистограмм при отключении опции «достоверность разности»: если готовая гистограмма не устраивает, можно взять копию таблицы значений и построить объемные 3D-гистограммы в EXCEL или других приложениях Windows:



6. Многопараметрическая доказательная статистика. РФА на примере аналитического ПО «ФАКТОР-ПРОГНОЗ»

Важные моменты:

РФА считается инструментом доказательства величин значимости влияния параметров на критерии.

РФА позволяет оценивать новые факторы (лечения) уже с результатов 2–3 случаев применения, при рассмотрении этих случаев на фоне уже имеющейся текущей рабочей базы данных (без создания специальных групп, как это делается традиционно в однопараметрической статистике на исследовании достоверной разности групп), это значительно экономит затраты на исследования.

6.1. Многопараметрическая доказательная статистика. РА и его развитие. РФА

РА всегда был инструментом доказательства на международном уровне для научного подтверждения различных гипотез – в техническом, медицинском ВАКе найденная весовая функция влияния параметров процесса на результат, полученный математическим методом РА, классифицируется как элемент открытия на защиту ученой степени доктора наук (технических, медицинских, ...). Но отсутствие специализированного программного обеспечения инструмента РА не могло преодолеть следующие проблемы:

- сложность неавтоматизированного решения систем уравнений требует подключать профессионального математика – нет возможности детальных исследований в виде быстрого перебора всех возможных вариантов для поиска решений – трудоёмкий и дорогостоящий процесс;
- при расчётах вручную с увеличением количества параметров сильно усложняется решение систем уравнений традиционными средствами;
- Формы представления результатов – визуализация решений должна быть не только в виде весовых коэффициентов, но и в виде гистограммы вкладов параметров в результат процесса в размерности результата для полной и наглядной картины представления приоритетов влияния параметров процесса на результат процесса, что очень важно для быстрых оценок эффективности;
- требуется быстрый инструмент выборки любых групп данных для перебора всех возможных вариантов групп для исследований, особенно в таких сложных биологических системных процессах, как онкология.

6.2. Принцип работы с РФА

Принцип работы с РФА в ПО «PLANETA» очень простой:

- указывая любые параметры результатов процесса (критерии) и набирая все остальные параметры в весовую функцию, можно сразу получать быстрые решения РФА в виде наглядной, полной и исчерпывающей картины влияния всех параметров (и качественных, и числовых) на результат процесса в единой математической схеме.

Результат получается в виде гистограммы, построенной по весовым вкладам параметров в результат и в виде системы высокоточного индивидуального прогноза значений критериев в размерности критериев.

Гистограмма решения позволяет наглядно видеть величины и приоритеты значимостей параметров на результат и решать задачи оптимизации и выбора оптимальных действий по максимизации (минимизации) критериев (результатов) процесса. А запись весовых функций в анкету позволяет получать автоматический прогноз значений всех критериев автоматически при заполнении требуемого для прогноза комплекта данных.

РА – международный инструмент работы с многопараметрическими системными процессами (как расчетный инструмент проведения научных работ на соискание учёной степени доктора наук в любой научной области)

6.3. Таблица сравнения доказательных методов однопараметрической статистики и многопараметрического РФА

<p>Однопараметрическая доказательная статистика типа достоверности разности групп (по Стьюденту, по Фишеру...)</p> <p>Ссылка на видео</p>	<p>Многопараметрическая доказательная статистика – РФА</p> <p>Ссылка на видео</p>
<p>Обязательные условия:</p> <p>соблюдение условия идентичности сравниваемых групп с ограничением воздействий для чистоты эксперимента</p>	<p>Обязательные условия:</p> <p>для решения системы уравнений количество заданных параметров в расчёте не должно быть больше количества опытов в расчёте весовой функции</p>
<p>Затраты</p> <p>Очень затратный, требует организации эксперимента – создания специальных групп</p>	<p>Затраты</p> <p>Не требует дополнительных затрат на организацию оценок – проведение оценок на текущем рабочем материале без ограничений на другие воздействия – очень важно в медицине, особенно в онкологии.</p> <p>Не нужно ограничивать проверенные методы лечения параллельно с новыми при условии их варьирования</p>

Недостатки	Преимущества
<ul style="list-style-type: none"> - Проблематичное соблюдение обязательной идентичности сравниваемых групп в области медицины, онкологии. - Не позволяет перейти к индивидуальному прогнозу. - Сложность соблюдения обязательного условия идентичности сравниваемых групп – в социальных процессах, финансах, биологии, медицине 	<ul style="list-style-type: none"> - Обладает свойством доказанности результата. - По картинам весовых значимостей факторов на результат позволяет легко дифференцировать общий поток данных процесса на однородные группы. - Позволяет проводить доказанную оценку эффективности воздействий на текущих данных в процессе работы. - Позволяет создавать системы высокоточного прогноза на всём объёме данных по весовым функциям РФА

6.4. Возможности РФА в медицине

1. РФА – международный инструмент доказательства свойств многопараметрических процессов.

2. РФА позволяет получить полную картину влияния параметров процесса на результаты процесса в единой математической схеме РА.

По полной картине влияния параметров на результат можно:

- разбивать исследуемые группы на подгруппы по картинам значимости параметров на критерий – 100-процентный прогноз по критериям свидетельствует о том, что в выборке только случаи одного процесса;
- оценивать эффективность воздействий внутри однородных групп на перечень критериев;
- для однородных групп получать системы высокоточного прогноза по всему комплексу результатов процесса (критериев).

3. Автоматизированное ПО РФА позволяет легко и быстро получать полные картины влияния параметров процесса на его результаты, что в десятки раз быстрее работы с однопараметрической статистикой.

4. Однопараметрическая статистика работает без учета влияний остальных параметров и может показывать противоположные результаты – РФА позволяет определить реальные знак и точную величину реального собственного вклада параметра в процесс в единой математической схеме.

5. РФА позволяет получать первые оценки эффективности новых препаратов и схем лечения на фоне уже имеющейся группы, начиная уже с 2–3 случаев; при явно отрицательном результате можно прекратить использование нового препарата с

минимальным ущербом для пациентов или при положительном результате продолжить уточнение эффективности с дальнейшей детализацией по условиям применения по мере добавления результирующих данных.

6. РФА позволяет делать оценки эффективности на рабочем материале без создания специальных исследовательских групп и без ограничений уже на существующих схемах лечения для чистоты эксперимента: для РФА не нужно ограничивать необходимые процедуры, а только чередовать их в опытах с новым воздействием.

7. РФА – индивидуальный подход – системы высокоточного прогноза по весовым функциям РФА – весь цикл индивидуальной работы с пациентом.

По мере набора пациентов в базе данных в РФА можно легко и быстро постоянно уточнять результаты РФА, и создавать более детальные однородные группы, и повышать точность результатов.

6.5. Импорт и подготовка данных СУБД «ФАКТОР-ПРОГНОЗ» («ПЛАНЕТА»)

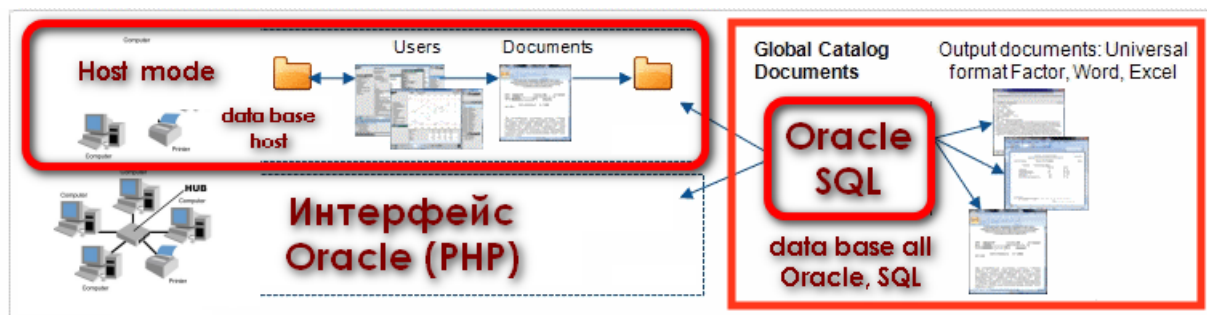
Все опции «ПЛАНЕТА», включая выборки данных для переноса в другие статистические пакеты на Windows, исключая расчёты аналитики, бесплатные.

Все группы по выборкам в «ПЛАНЕТА» можно переносить и дублировать расчёты в других статистических пакетах, в которых нет систем выборок данных типа «SPSS» или «Statistica».

В случае, когда у вас на предприятии уже есть рабочая СУБД для регистрации данных (например, на SQL, Oracle...):

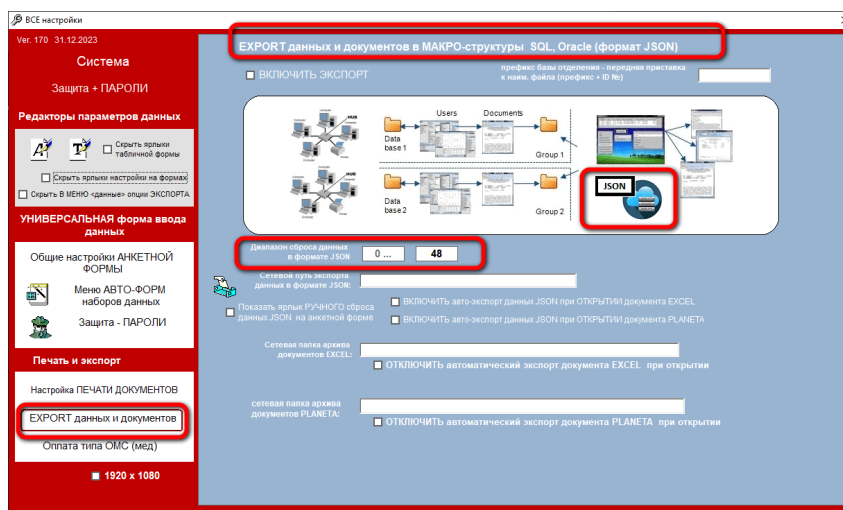
Используется 4 уровень СУБД «ФАКТОР-ПРОГНОЗ» – «Модуль» – сервер SQL, Oracle с интерфейсом ведения данных (PHP) + дополнительный интегрированный модуль «ФАКТОР-ПРОГНОЗ» для науки и создания систем прогноза – для того чтобы не вносить данные дважды.

Есть кнопка сброса регистрационных данных в любую СУБД (SQL...).

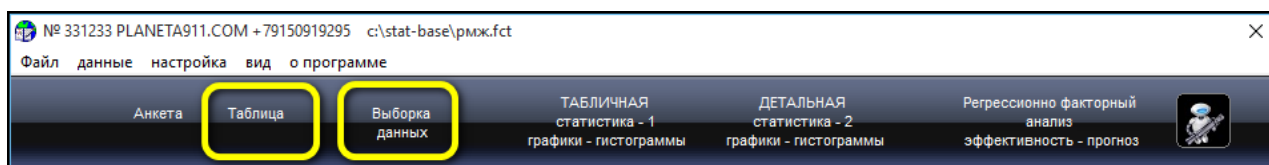


Для подключения СУБД «ФАКТОР-ПРОГНОЗ» как дополнительного модуля для научной работы и создания систем высокоточного прогноза необходимо:

- 1) сделать экспорт рабочей базы данных в таблицу Windows (Excel, Access) для автоматической установки в СУБД «ФАКТОР-ПРОГНОЗ»;
- 2) установить эту таблицу данных с той же последовательностью параметров, как и в рабочей базе данных для научной работы и создания систем прогноза;
- 3) настроить рабочий экспорт данных, набираемых в дополнительном модуле, в отдельную папку – для последующей автоматической загрузки этих данных в сеть.



Импорт – модификация – выборки



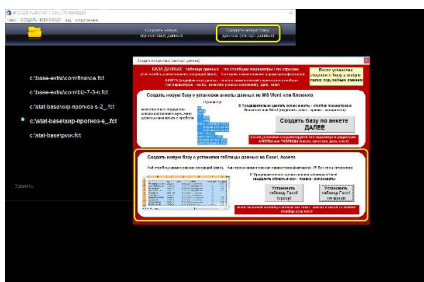
Импорт и подготовка данных в комплексе «PLANETA» включает в себя 3 этапа:

1. Импорт (см. подробно Раздел 6.1, см. видео – Импорт данных из таблиц Windows);
2. Модификация параметров данных для проведения расчётов и представления результатов анализа:
 - 2.1. Модификация базы (арифметика столбцов), см. видео;
 - 2.2. Для проведения расчётов в РФА необходимо качественные параметры, в которых > 2 состояний, разложить в конец анкеты данных в виде дополнительных столбцов двузначных параметров: нет этого качественного состояния и есть это качественное состояние, и такие двузначные параметры уже могут включаться в весовую функцию РФА (см. раздел 6.2 – разложить качественные параметры как факторы в конец анкеты);

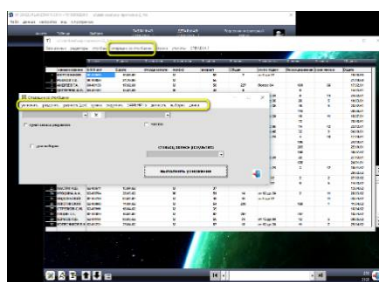
3. Выборки данных (см. видео – Любые выборки данных для анализа и отчётов) – этот вариант подготовки данных для анализа, в отличие от предыдущих, выполняется в процессе анализа с помощью специальной формы «Выборки данных».

Раздел «Детальные инструкции» по подготовке данных к анализу (видео):

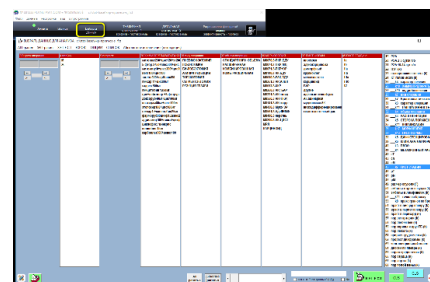
Импорт данных
из таблиц Windows



Модификация базы
(арифметика столбцов)



Любые выборки данных
для анализа и отчётов



Видео

Подробное
видео

Видео

Видео

Общие для любого анализа правила заполнения таблиц данных

Можно сделать экспорт любой демобазы в Excel и посмотреть, как она выглядит в Excel для установки в «PLANETA»

При заполнении текстов не используйте спецсимволы

Смотри номера в скобках [...] по картинке внизу

Форма "Таблица данных"
число, качество, дата, текст

Анкета данных
для качественных параметров -
список качественных состояний

Операции со столбцами

Редактор анкеты

номер	дата	срок	Прогресс	Рецидив	Реконст	МТС	возраст	возраст	пол	диагноз	Гистолог	гр	по	T	T	N	M	G	Стадия
56			нет	нет	нет	нет	56	50 - 59	ж	Липоса	Липоса	leg	1	1	0	0	0	G3	IIIa
89			нет	нет	нет	нет	18	ДО 20	ж	ЗЛО	ЗЛО с	leg	2	2	0	0	0	G3	IIIb
11			Прогресс	нет	нет	нет													
82			Прогресс	рецидив	нет	нет													
220			нет	нет	нет	нет													
102			Прогресс	рецидив	нет	нет													
228			нет	нет	нет	нет													
11			Прогресс	нет	нет	нет													
18			Прогресс	рецидив	нет	нет													
132			нет	нет	нет	нет	46	40 - 49	ж	ПС	ПС	тяж	1	1	0	0	0	G3	IIIa
72			Прогресс	нет	нет	нет	45	40 - 49	ж	ЗЛО	ЗЛО с	leg	2	2	0	0	0	Gx	x
38			Прогресс	нет	нет	нет	38	30 - 39	ж	Фиброс	Фиброс	leg	1	1	0	0	0	G3	IIIa
29			Прогресс	рецидив	нет	нет	62	60 - 69	ж	ПС	ПС	тяж	1	1	0	0	0	G3	IIIa
52			нет	нет	нет	нет	63	60 - 69	ж	ПС	ПС	тяж	2	2	0	0	0	G2	IIIb
32			Прогресс	нет	нет	нет													
40			Прогресс	рецидив	нет	нет													
58			нет	нет	нет	нет													
96			Прогресс	нет	нет	нет													
18			нет	нет	реконст	нет													
102			Прогресс	нет	нет	реконст													
64			нет	нет	нет	нет	64	60 - 69	ж	РИД	Ангисос	leg	1	1	0	0	0	G3	IIIb
54			МТС				54	50 - 59	ж	ПС	ПС	ж	x		0	0	0	G3	III
50			нет	нет	нет	нет	50	50 - 59	ж	РИД	Леомио	leg	2	2	0	0	0	Gx	x
42			МТС				42	40 - 49	ж	ПС	ПС	тяж	2	2	0	0	0	G3	IIIb
33			нет	нет	нет	нет	33	30 - 39	ж	Ангисос	Ангисос	leg	2	2	0	0	0	G2	IIIb
45			нет	нет	нет	нет	45	40 - 49	ж	ЗЛО	ЗЛО с	leg	1	1	0	0	0	G3	III
26			нет	нет	реконст	нет	26	20 - 29	ж	Ангисос	Ангисос	leg	2	2	0	0	0	G2	IIIb
63			Прогресс	рецидив	нет	нет	63	60 - 69	ж	ЗЛО	ЗЛО с	leg	2	2	0	0	0	G3	IIIb
16			нет	нет	реконст	нет	16	ДО 20	ж	Рабдом	Рабдом	leg	2	2	0	0	0	G1	IIIb

1. 1-й столбец таблицы это наименование строк (опытов, ФИО)
- 1-я строка таблицы – это наименование столбцов (наименования параметров).
2. В системах «PLANETA» в анкетной форме ведения данных:
 - 1-й столбец – это ФИО;
 - 2-й столбец используется под № (ID) записи (в медицине № истории болезни);
 - 3-й столбец – под дату регистрации (проставляется автоматически).
3. [7] В числовых параметрах (столбцах) должны быть только числа или пустые значения.
4. В качественных параметрах должны стоять наименования качественных состояний (или порядковые номера, начиная с цифры 1 (1, 2, 3... N) – если проставите порядковые номера в ячейки, то в системе вы должны будете заполнить наименования меню качественных состояний и поменять тип столбца с числового на качественный).
5. [5, 9, 11] Качественный параметр не должен состоять более чем одного качественного значения – T3aN0M0 – это четыре отдельных параметра: «Т3» (Т0, Т1, Т2, Т3, Т4), «а» («а», «в»), «N0» (N0, N1), «M0» (M0, M1).
6. Текстовые ячейки могут использоваться для формирования выборок данных по ключевым текстовым сочетаниям букв и слов.
7. [5, 9, 11] «Химиолучевая» – это два отдельных параметра одной качественной группы «вариант лечения» (или один для исследования действия совместного эффекта).
8. Даты – все даты должны быть в формате даты.
9. Нельзя ставить в один столбец параметры разных качественных групп – «лучевая» (лечение), а «стадия» – диагноз.

№	А	В	С	Д	Е	Г	Н	И	К	Л		
1	Пациент	№ истории болезни	Возраст	Пол	ДИАГНОЗ TNM	Адрес	Высота за Печенбис	У	Предоперационная	Дата операции	лечение	
2	Просветлов Александр	54691	63	1	м	T3N0M0(ст.2a)	Рязанская	1	ГБУЗ РО "О" не получал	10.09.2015	Лучевая, химио	
3	Сундуров Саид	8732-15	58	1	м	T3N0M0(ст.2a)	Республика	4.06x5.5	ГБУЗ РКОД не получал	10.09.2015	Лучевая, химио	
4	Александр С-156960					T1-2N0M0(ст.1)	Республика		ГБУЗ "РНХ" не получал	15.09.2015	Лучевая, химио	
5	Игнатовский Рюрик	37665	39	1	м	T1-2N0M0(ст.3b)	Воронеж		ГБУЗ ВО ВС лучевая терапия	19.09.2015	Лучевая, химио	
6	Банай Пл	37712	47	1	м	T1-2N0M0(ст.3b)	Воронеж		ГБУЗ ВО ВС химиолучевая	20.09.2015	Лучевая, химио	
7	Гусов Александр	16984	47	1	м	T3N0M0(ст.2a)	Рязанская		ГБУЗ РО ОК не получал	21.09.2015	Лучевая, химио	
8	Драч Александр	16984	47	1	м	T1-2N0M0(ст.3b)	Рязанская	3,5	убе. №87	21.09.2015	Лучевая, химио	
9	Хуми Рамис	60	60	1	м	T1-2N0M0(ст.3b)	Рязанская	4,8		25.09.2015	Укажите значение	
10	Чудов Алексей	21625	39	1	м	T1-2N0M0(ст.3b)	Рязанская	3,5	химиолучевая	30.09.2015	Стадия 3-4-мет	
11	Иванов Иван	7458	48	2	ж	T3N0M0(ст.2a)	Р-Ф, Никее	1,5	ГБУЗ "Т" химиолучевая	30.09.2015	Стадия 4-мет	
12	Булдин Дмитрий	9489-15	62	1	м	T1-2N0M0(ст.1)	РФ, Ишим	4,8	ГБУЗ РКОД не получал	01.10.2015	Укажите значение	
13	Пелешко Федор	9493-15	54	1	м	T1-2N0M0(ст.1)	Россия, Р	4	ГБУЗ РКОД не получал	01.10.2015	Укажите значение	
14	Шарипов Александр	46129	64	1	м	T1-2N1-2aM0(ст.3a)	Москва О	3,5	ММА	не получал	24.09.2015	Лучевая, химио
15	Шугуров Александр	45994/2015	64	1	м	T3-2N0M0(ст.3c)	г. Епифанов	3,5	ММА	не получал	29.09.2015	Укажите значение
16	Мавлюков Павел	44297/2015	66	1	м	T1-2N1-2aM0(ст.3a)	Магачевск	3,5	ММА	не получал	01.10.2015	Укажите значение
17	Фадеев Илья	178315-1	67	1	м	T1-2N0M0(ст.1)	г. Иваново	4,8	ОБУЗ Иван не получал	07.10.2015	Лучевая, химио	
18	Цибель Александр	C17575/a	60	2	ж	T1-2N0M0(ст.1)	Калмыкия	5,5	ФГБУЗ РНИИ не получал	07.10.2015	Лучевая, химио	
19	Вашинский Никс	7688	58	2	ж	T3N0M0(ст.2a)	Нижегород	1	ГБУЗ НО "Т" не получал	08.10.2015	Лучевая, химио	
20	Мартынов Александр	C-17523/d	61	1	м	T3N0M0(ст.2a)	Ростовск	5,5	ФГБУЗ РНИИ не получал	09.10.2015	Лучевая, химио	
21	Киселев Александр	9783-15	57	2	ж	T1-2N0M0(ст.1)	Россия, Б	4,8	ГБУЗ РКОД не получал	06.10.2015	Лучевая, химио	
22	Кавалеров Александр	37823	69	2	ж	T1-2N0M0(ст.3b)	Воронеж	4,5	ГБУЗ ВО ВС не получал	13.10.2015	Лучевая, химио	
23	Починков Владимир	C17850/d	41	1	м	T3-2N1-2bM0(ст.3c)	Ростовск	4,5	ФГБУЗ РНИИ не получал	13.10.2015	Лучевая, химио	
24	Цыганков Александр	37958	59	1	м	T1-2N1-2bM0(ст.3b)	Воронеж	4,5	ГБУЗ ВО ВС химиолучевая	15.10.2015	Стадия 4-мет	
25	Рябенко Александр	7840	56	1	м	T3N0M0(ст.2a)	Нижегород	4,5	ГБУЗ НО Н химиолучевая	15.10.2015	Стадия 4-мет	
26	Халипов Александр	C-19325/a	52	1	м	T1-2N0M0(ст.1)	респ. Даг	4.0x5.5	ФГБУЗ РНИИ не получал	02.11.2015	Лучевая, химио	
27	Лавров Александр	8137	80	2	ж	T3N0M0(ст.3a)	РФ, г. Наб	7,5	ГБУЗ НО Н лучевая терапия	02.11.2015	Стадия 4-мет	
28	Аскариев Мухамед	10197-15	59	1	м	T1-2N0M0(ст.1)	РФ, г. Уфа	3,5	ГБУЗ РКОД не получал	16.10.2015	Лучевая, химио	
29	Шенников Иван	37999	76	2	ж	T1-2N1-2aM0(ст.3a)	Воронеж	4,8	ГБУЗ ВО ВС не получал	20.10.2015	Лучевая, химио	

Автоматический импорт данных в СУБД «ФАКТОР-ПРОГНОЗ» («PLANETA») для анализа

При импорте текстовые строки автоматически формируются в качественные параметры с меню качественных состояний – если качественных состояний будет больше 20 – то программа предложит изменить тип параметра на текстовый и при отказе продолжит набирать меню состояний.

1. Выделяется таблица стандартной формы и делается её копия Ctrl + C

– 1-я строка – наименования параметров,

– 1-й столбец – наименование опытов (в медицине это ФИО пациентов).

2. Запускается программа «PLANETA», и по ярлыку «Создать новую базу» взятая копия таблицы устанавливается в систему «PLANETA» с автоматической установкой текстовых переменных как качественные параметры – подробно установку посмотрите на подробном видео.

Создать новую базу (импорт данных)

БАЗА ДАННЫХ - таблица данных - по столбцам параметры / по строкам
(1-й столбец наименование операций (фио), 1-я строка наименование параметров-факторов)
АНКЕТА (кодификатор) данных - список наименований параметров-столбцов
Тип параметров - число, качество (список состояний), дата, текст

После установки сохраните базу в любую папку под любым именем

Создать новую базу с установки анкеты данных из MS Word или блокнота

Пример

дата	муж	этанол	тест
пол	жен	отрицательный	положительный
группа	много (ед.)	протейн С (%)	гидроминос (%)
анализ	Антилазмин (%)		

!!! Предварительно сделать копию анкеты - столбик параметров в Блокноте или Word (выделить текст - правка - копировать)

Создать базу по анкете ДАЛЕЕ

после установки скорректируйте тип параметра в редакторе АНКЕТЫ или ТАБЛИЦЫ (число, качество, дата, текст)

Создать новую базу с установки таблицы данных из Excel, Access

(1-й столбец наименование операций (фио), 1-я строка наименование параметров-факторов) (!!! Без спец.символов)

	A	B	C	D	E	F	G	
1								
2	История	ТК	коп	05/2025	19.05.2009	Силами	34	1,97
3	ИСТОМЕРОВА	ИМ	05/2025	22.05.2009	Отрицательный		25	1,98
4	Тюковин	ИМ	05/1518	30.06.2009	Отрицательный		30	1,94
5	МАКЛЕРОВА	ТМ	05/1512	22.07.2009	Отрицательный		30	1,91
6	Плюкова	ОБ	05/1910	07.10.2009	Силами		23	1
7	ПОПОВА	ОБ	05/1502	12.10.2009	Силами		24	1,91
8	КОЗЛОВСКАЯ	ЖМ	05/1601	15.10.2009	Положительный		23	1,93
9	Ковач	ТТ	03/1505	22.10.2009	Силами		30	1,97
10	ДЕНИСОВА	ВА	05/1013	23.10.2009	Отрицательный		30	1,76
11	КАРАМАНОВА	ОБ	05/1509	23.10.2009	Неизвестно		40	1,92

!!! Предварительно сделать копию таблицы в Excel (выделить область ячеек - правка - копировать)

Установить таблицу Excel (сразу)

Установить таблицу Excel по шагам

если числовой столбец считался как текст - значит в какой-то ячейке столбца есть текст

20.01.2023

3. Подправьте в установленной базе порядок следования качественных состояний в качественных параметрах.

4. И если вы будете работать с РФА, разложите качественные параметры с более двух значений как двузначные факторы в конец анкеты, так как в весовые функции РФА можно добавлять только двузначные качественные состояния: «нет фактора» и «есть фактор» (состояние в качественном параметре «есть фактор» (1) должно стоять в позиции 2, а состояние «нет фактора» (0) должно стоять в позиции 1).

6.5.1. Разложение качественные параметры как факторы в конец анкеты (таблицы) данных для работы в РФА

В программном аналитическом комплексе «ФАКТОР-ПРОГНОЗ» («PLANETA») на рабочей или установленной из Excel таблице базы данных:

1. В двузначных качественных параметрах 1-я качественная позиция означает «нет фактора», 2-я качественная позиция – «есть фактор»; установите такую последовательность перестановкой пунктов по кнопкам редактора анкеты, если её нет.
2. Если в качественном параметре > 2 состояний, то необходимо разложить качественные состояния автоматически в конец анкеты – 1. позиция «нет фактора» и 2. позиция «есть фактор».

В РФА использовать качественные состояния можно только в двузначном виде:

0 (1 качественная позиция) – нет фактора;

и 1 (2 качественная позиция) – есть фактор.

В комплексе «PLANETA» есть программное автоматическое разложение качественных параметров > 2 состояний к виду двузначных факторов в аналитическом научном комплексе «PLANETA» («ФАКТОР») и запись преобразованных качественных параметров в двузначном виде в конец таблицы базы данных.

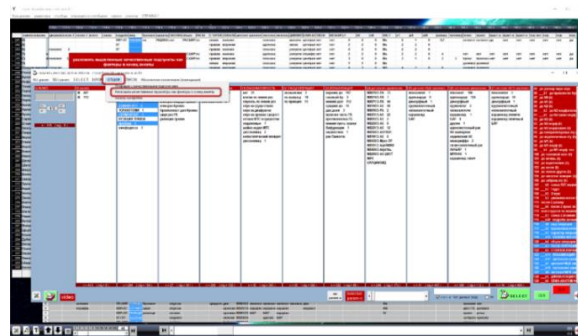
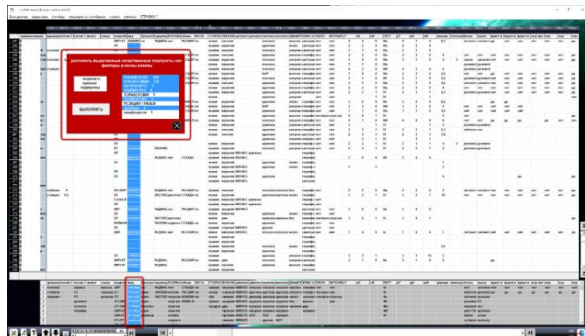
Автоматическое разложение качественных параметров с более 2-х качественных состояний в базы таблицы данных – добавляются двузначные параметры в виде дополнительных столбцов двузначных параметров в конец анкеты данных:

0 – «нет фактора»;

1 – «есть фактор»; можно отдельно для каждого параметра из двух форм:

из <Табличной формы> работы с данными
в разделе <Главное меню> <Столбцы>
<Разложить подпункты как факторы в
конец анкеты>

из формы <Выборки данных>
<Опции> <Разложить качественные
параметры в конец анкеты>





6.5.2. Окно быстрого формирования выборок данных

Это самый необходимый этап любого анализа информации с возможностью наложения ограничений по всем параметрам одновременно

Видео

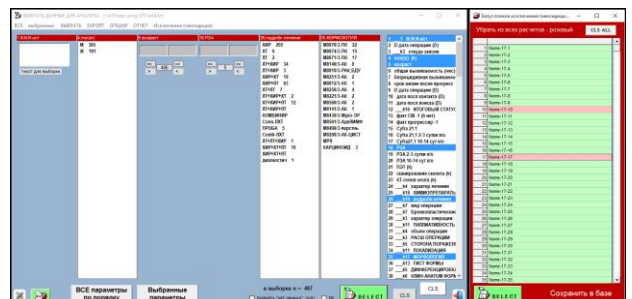
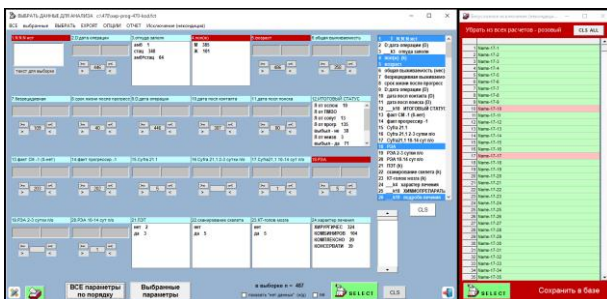
Важная часть любого анализа данных и формирования отчётов:

- наложение ограничений на числовые параметры в виде строгих и нестрогих неравенств;
- наложение ограничений на качественные параметры при их выделении. Если не выделено ни одно качественное состояние, то ограничений на этот параметр нет;
- строгие и не строгие ограничения по датам;
- ограничения на текстовые параметры заданием необходимого присутствия в тексте параметра указанного в текстовом окне обязательного текстового или буквенного фрагмента.

Наложением ограничений в РФА анализе данных можно реализовывать двухэтапный подход:

1 этап – разделять весь поток по несовместимым процессам: разделять весь поток данных по несовместимым однородным процессам по индивидуальным картинкам влияния факторов процесса на критерии результатов процесса с целью решения задач прогноза с целью идентификации принадлежности конкретного опыта к тому или иному процессу.

2 этап – разделять однородный процесс на подгруппы: разделять однородный процесс на группы внутри однородного процесса для подбора наиболее эффективных методов воздействия (например, подбор наилучших схем лечения по стадиям, морфологиям и гистологиям в области медицины, см. видео).



Можно выводить для наложения ограничений и просмотра результата как все параметры подряд, так и выводить только нужные параметры из разных мест базы.

Примечание по работе: Выбранные для работы из разных мест параметры для удобства можно визуализировать в одном месте, чтобы при работе с ограничениями и просмотром количеств не листать всю базу.

7. РФА в специализированном научно-практическом ПО «ФАКТОР-ПРОГНОЗ»

7.1. Проблемы использования РА при выполнении научно-исследовательской работы

РА всегда был доказательным инструментом на международном уровне для научного подтверждения различных гипотез – в ВАКе найденная весовая функция влияния параметров процесса на результат инструмента регрессионного анализа классифицируется как элемент открытия на защиту ученой степени доктора (технических, медицинских, биологических ...) наук.

Но отсутствие специализированного ПО инструмента РА не могло преодолеть следующие проблемы.

1. Требуется подключать профессионального математика для решения систем уравнений – нет возможности перебора всех вариантов для поиска решений.
2. С увеличением количества параметров усложняется решение систем уравнений традиционными средствами.
3. Визуализация решений должна быть не только в виде весовых коэффициентов, но и в виде гистограммы вкладов параметров в результат в размерности результата для точного представления приоритетов влияния параметров на результат – что важно для быстрых оценок эффективности.
4. Требуется очень быстрый инструмент выборки любых групп данных для перебора всех возможных вариантов групп для исследований, особенно в таких сложных биологических системных процессах, как онкология.
5. Возможность двухэтапной работы – когда на первом диагностическом этапе необходимо исследовать и разбить весь поток данных на однородные (по процессам и признакам исходных причин) группы с получением функций прогноза однородной группы; на втором оценочном этапе проводить оценочные работы уже внутри однородной группы для детальных исследований наилучших вариантов воздействий.
6. Как окончательный практический выход исследовательских работ:
 - прогноз (I этап – диагностика): создание высокоточных систем прогноза принадлежности конкретного случая к той или иной однородной группе;
 - прогноз (II этап – оценка) для подбора наилучших схем воздействий (лечения в области медицины, онкологии).

7. Кроме того, классический РА делался только для числовых параметров, а для качественных параметров использовался факторный анализ – два разных инструмента не позволяли получать единое решение в единой математической схеме, без которого оба результата были неполными.

Все эти проблемы для специалистов решил специализированный научно-практический комплекс ПО «ФАКТОР-ПРОГНОЗ», который работает без профессиональных математиков.

7.2. Получения весовых функций РФА

– дальнейшее развитие инструмента РА, в котором весь набор данных – числовые и качественные параметры – используется в единой математической схеме регрессионного анализа.

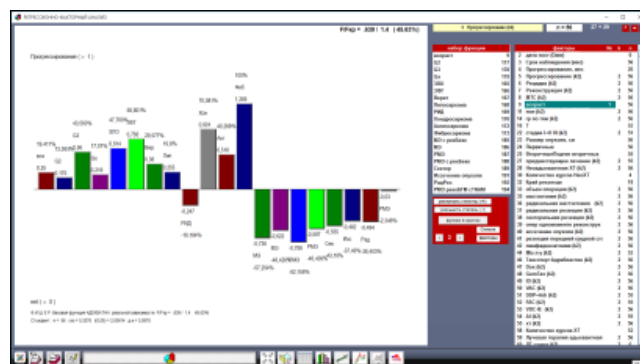
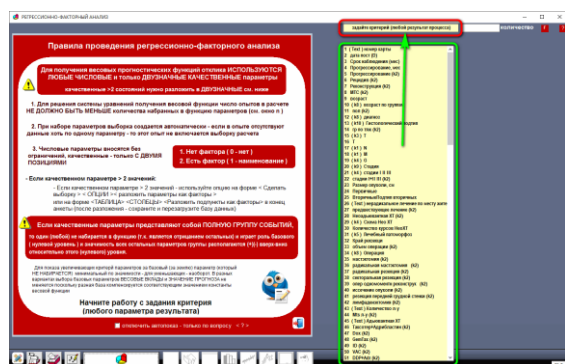
Технически в программе «ФАКТОР-ПРОГНОЗ» процесс получения весовых функций РФА очень простой: достаточно на форме РФА задать параметр критерия и далее набрать в весовую функцию РФА только те параметры процесса, которые увеличивают адекватность функции реальной зависимости по F критерию Фишера.

Как это делается в специализированном программном обеспечении «ФАКТОР-ПРОГНОЗ» в окне работы с РФА:

1 – задаём критерий;

2 – набираем параметры для РФА

и оставляем при улучшении адекватности.



1. В выкидном окне списка параметров задаёте параметр критерия (это любой параметр, который в вашем расчёте играет роль параметра результата).

2. В правом крайнем списке всех параметров набираете параметры в весовую функцию отклика РФА, соблюдая условие: параметр оставляется в весовой функции, если его участие увеличивает адекватность функции реальной зависимости по F-критерию Фишера (уменьшает ошибку между опытными и расчётными значениями весовой функции в прогнозе). Если же ошибка прогноза при добавлении параметра в весовую функцию увеличивается, то этот параметр исключается из весовой функции, то есть мы перебираем

все параметры, оставляя только те, которые увеличивают адекватность функции РФА реальной зависимости и уменьшают ошибку прогноза критерия.

7.3 Особенности работы РФА для медицины

В отличие от системных процессов в области экономики, финансов, техники, биологические процессы обладают свойством максимальной зависимости и сбалансированности всех параметров организма в такой степени, что каждый параметр организма имеет в организме и весовой функции РФА как свой функционал, так и может выступать одновременно как индикатор состояния всех остальных параметров организма с той или иной степенью точности.

Использование параметров как индикаторов очень широко используется в практике диагностики биологических систем в любой версии медицины – и в РФА позволяет получать функции высокоточного прогноза, где параметры работают и как косвенные индикаторы всех остальных параметров, не указанных в весовой функции.

7.4. Возможное смешение в одной исследуемой группе разных процессов и способы выделения однородных групп

В биологических системах отсутствие 100-процентного прогноза может означать смешение в исследуемой выборке разнородных процессов – эту коллизию можно устранить следующими способами:

1 способ – начать исследования с разбивки всего потока данных по фундаментальным исходным признакам. Например, в области онкологии начать разбивку на подгруппы по исходным признакам и условиям возникновения нарушений: группы случаев, вызванные действием радиации; случаи, вызванные наличием в питании определенных химических консервантов; случаи по особым локализациям в органах; случаи следствий генетических нарушений и т. д.

2 способ – подбирать более однородную выборку для расчётов дополнительными ограничениями; в этом случае нужно обязательно указывать эти ограничения в инструкциях к готовым системам прогноза.

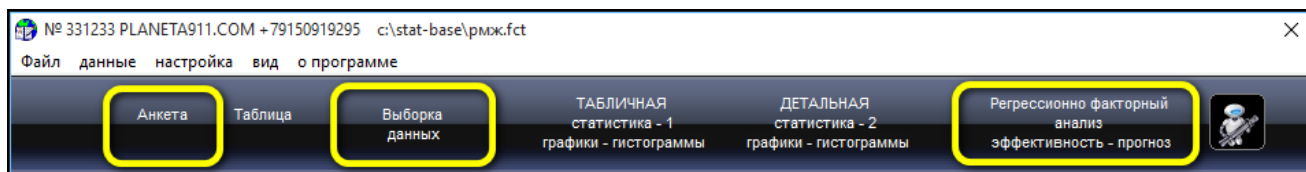
3 способ (основной в научных работах специалистов) – «отсеивать» удалять опытные данные выборки, начиная с тех, которые имеют максимальные ошибки в прогнозе (например, в онкологии допускается удаление до 5% случаев вследствие сильной выраженности индивидуальности процессов системных нарушений без возможности их идентифицировать другими средствами, а также в связи с недостаточным объёмом данных, что устраняется со временем практическим добором пациентов).

Нужно отметить, что очень простой механизм работы с созданием весовых прогностических функций РФА требует понимания природы и функциональных свойств исследуемых параметров: для работы с РФА требуются специалисты, которые на практике работают с исследуемыми группами, – это позволяет корректно указывать ограничения для функций прогноза РФА в практическом их использовании.

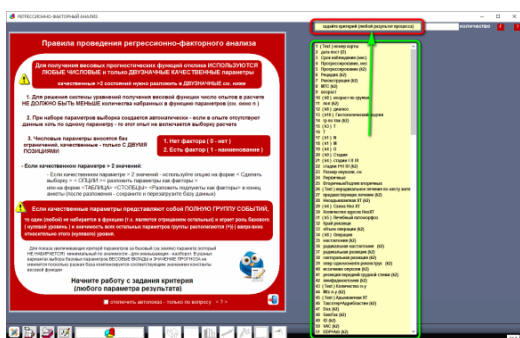
В любом случае разбивку исследуемой группы на однородные группы по однородным процессам должен делать сам специалист, имеющий практический опыт работы со своими данными.

7.5. Создание функций РФА с автоматическим прогнозом

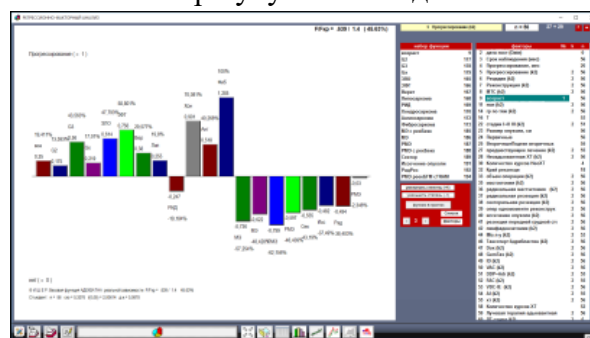
Работа с РФА в программном аналитическом комплексе «PLANETA»:



1 - задаём критерий



2 – набираем параметры для РФА и оставляем при улучшении адекватности




1. Делаем выборку данных.

2. Открываем окно РФА программы «ФАКТОР-ПРОГНОЗ», задаём параметр критерия и набираем параметры анкеты в список параметров весовой функции РФА при условии, что адекватность функции реальной зависимости увеличивается.

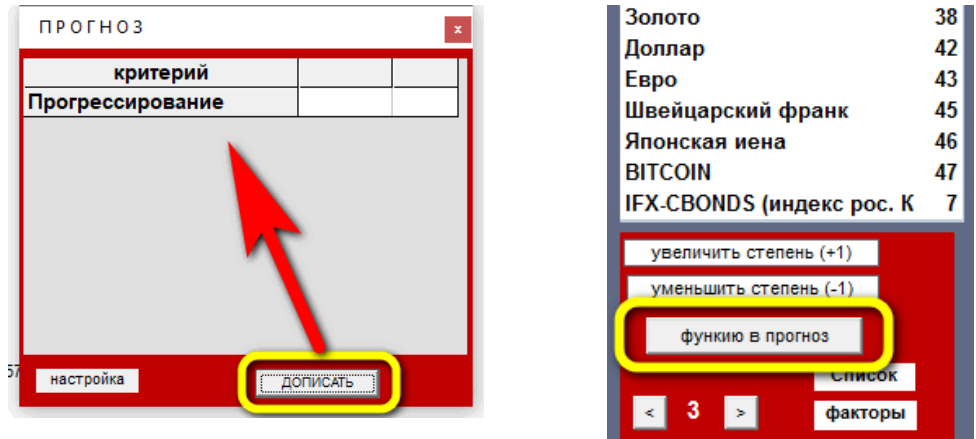
3. В окне РФА нужно указать мышкой критерий (параметр результата) в списке всех параметров.

4. Набрать мышкой в окно функции параметры для расчёта функции РФА из всего списка параметров; при соблюдении условия, что при добавлении параметра рассчитанное значение адекватности функции реальной зависимости улучшается, ошибка между расчётным и опытным значением меньше критического значения ($F < F_{\text{критич.}}$) и уменьшается.

5. Нажать кнопку расчёта , и получить исчерпывающую картину значимостей влияния параметров на критерий в условных единицах размерности критерия (параметра результата), и получить значение адекватности функции реальной зависимости по критерию Фишера.

6. Автоматический прогноз – отправляем полученный результат в анкетную форму регистрационной базы данных для прогноза критерия и записываем решение в список автоматического прогноза анкетной формы.

7. По мере заполнения данных регистрационной анкетной формы данными нового случая автоматически получаем значения прогнозируемых критериев.

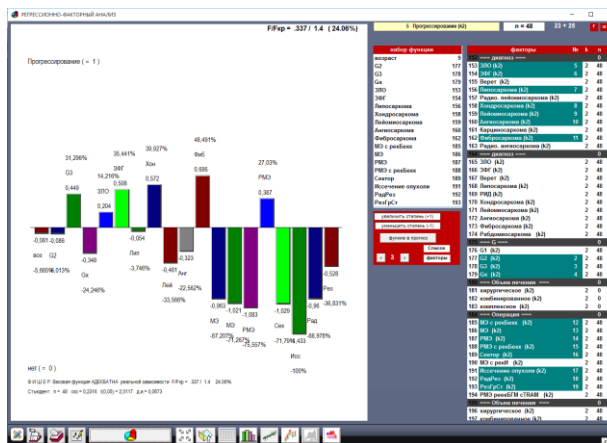


Медицина (критерий – качество)

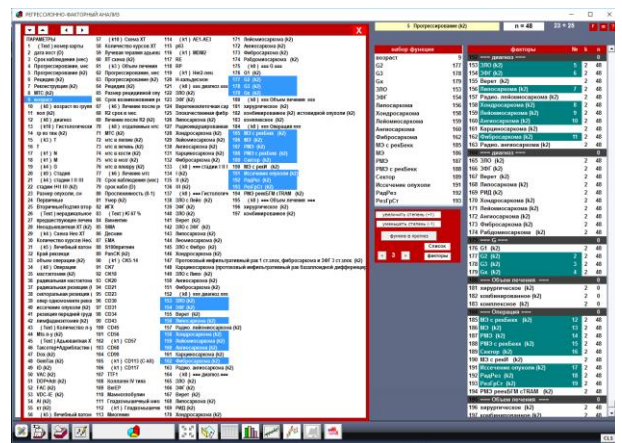
VIDEO MEDICINE

Набор параметров в РФА из правого списка
(не закрывается поле расчёта)

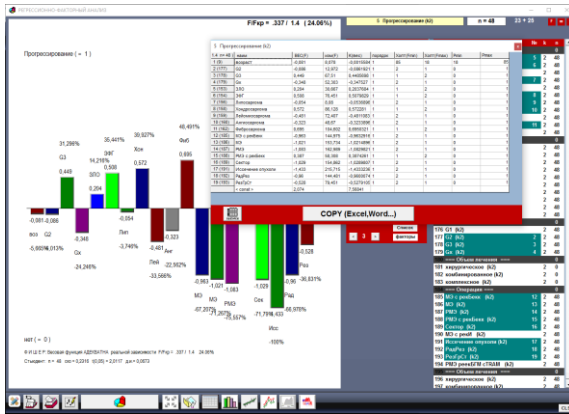
Набор параметров в РФА из панели всей анкеты (закрывается поле расчёта)



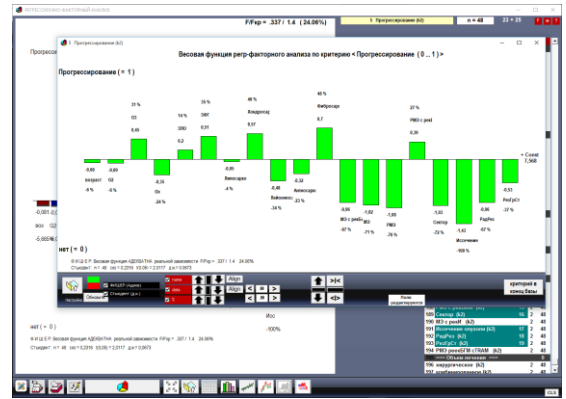
Гистограмма эффективности и таблица расчетов



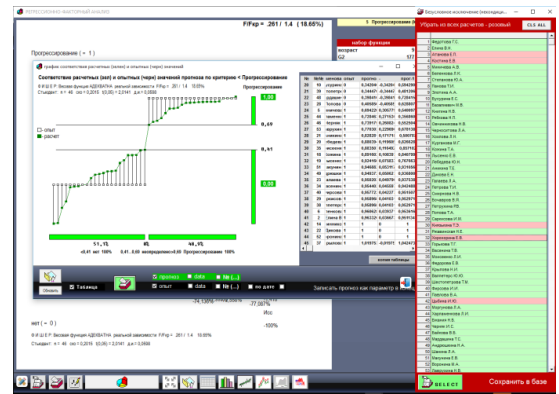
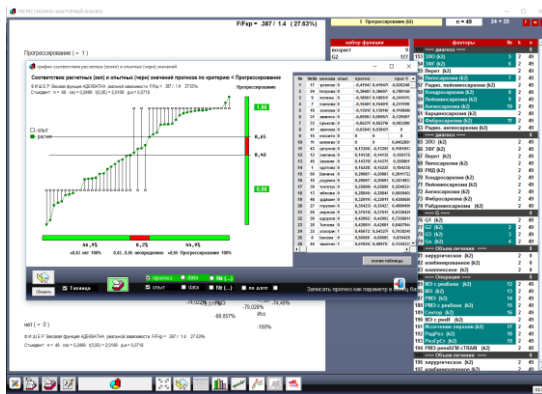
Гистограмма эффективности для экспорта в другие приложения



Соответствие опытных и расчётных значений



Удаление параметров из зоны неопределенности – 100-процентное разграничение



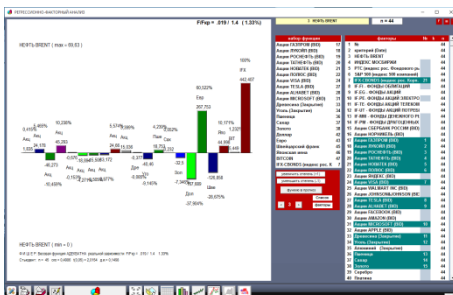
7.6. Различия РФА для числового и качественного критерия

Для числового критерия
(критерий – числовой параметр с любыми значениями)

1.7. Окончательный вариант оценки эффективности параметров

1.8. Графики соответствия расчётных и опытных значений по числовому критерию

1.9. Соответствие расчётных и опытных значений по схеме «Минус один»



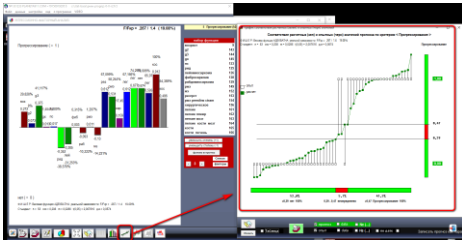
VIDEO

VIDEO

VIDEO

Для качественного критерия
(критерий – качественный параметр с двумя значениями)

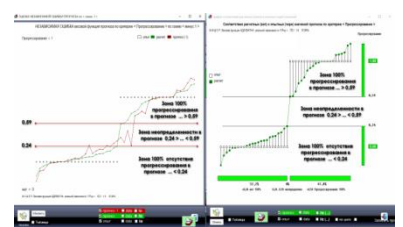
1.7. Окончательный вариант – просмотр на соответствие расчетных и опытных значений



1.8. Доводка РФА прогноза – удаление 5% неопределенных по данному критерию записей



1.9. Соответствие расчетных и опытных значений по схеме «Минус один»

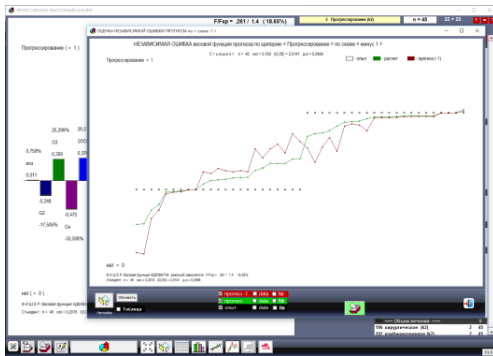


VIDEO

VIDEO

VIDEO

VIDEO

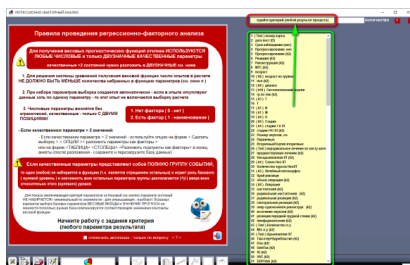


Соответствие опытных и расчётных значений по схеме «Минус один» – весовая функция строится без убираемого пациента – и для этого пациента считается прогноз как для вновь прибывшего.

Данная опция позволяет графически (визуально) показать, как будет выглядеть соответствие прогнозируемых и реальных значений критерия для вновь прибывающих больных, которые не участвовали в расчёте прогностической функции, и оценить ошибки в реальном прогнозе.

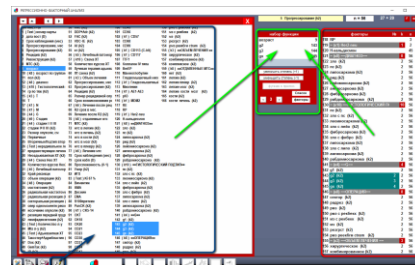
7.7. Создание линейных весовых функций РФА

1.1. Задание критерия в весовую функцию РФА



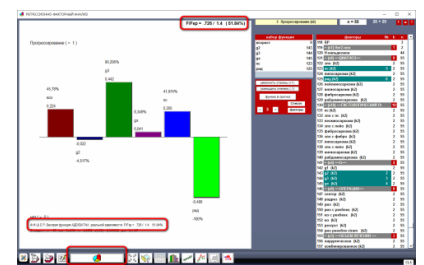
VIDEO

1.2. Задание – набор параметров в весовую функцию РФА



VIDEO

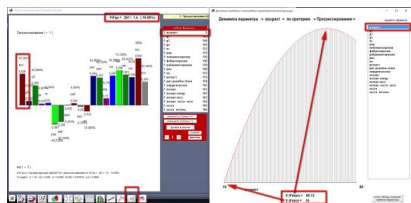
1.3. Расчет весовой функции РФА



VIDEO

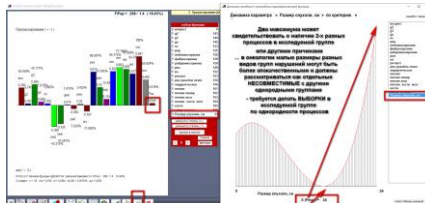
7.8. Создание нелинейной функции РФА n-го порядка

1.4. 2D функция РФА 2-го порядка по параметру «Возраст»



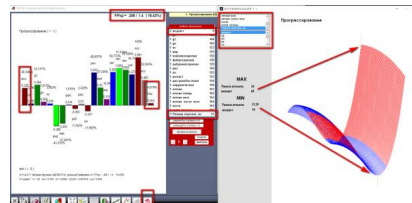
[VIDEO](#)

1.5. 2D функция РФА 3-го порядка по параметру «Размер объекта»



[VIDEO](#)

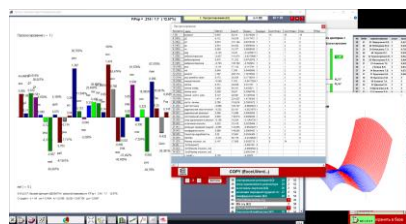
1.6. 3D (трёхмерная) функция РФА 3-го порядка в размерности 3D



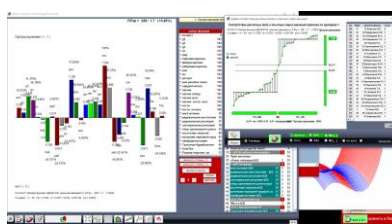
[VIDEO](#)

Нелинейная функция РФА 3-го порядка с большим количеством параметров уже не имеет зоны неопределенности

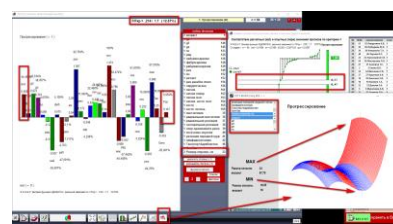
1.4. 2D функция РФА 3-й степени с большим количеством параметров



1.5. 2D функция РФА 3-й степени не имеет зон неопределенности – 100% разграничение по прогнозу



1.6. 3D (трёхмерная) функция РФА 3-го порядка в размерности 3D



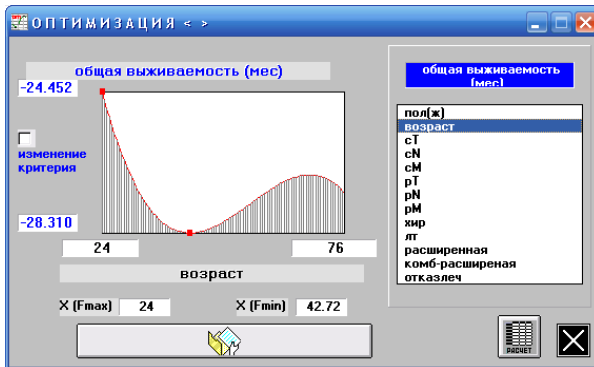
Нелинейные весовые функции отклика РФА любой степени в РФА «ФАКТОР-ПРОГНОЗ»

Получение нелинейных весовых функций осуществляется поднятием степени параметров в белом окне двойным щелчком кнопки мыши по нужному параметру (в примере – Доза (разовая)). Степень параметра обозначается цифрой перед наименованием – 2 Доза (разовая). Степень параметра поднимается только при условии $F(\text{большой степени}) < F(\text{меньшей степени})$, что свидетельствует о большем приближении функции к реальной зависимости. В противном случае степень понижается по кнопке **-1**. В примерах показана оптимизация по двум критериям – «Гранулемы в легких» и «Гранулемы в селезенке» для определения оптимальной дозы препарата ОФЛ по достижению минимума критерия.

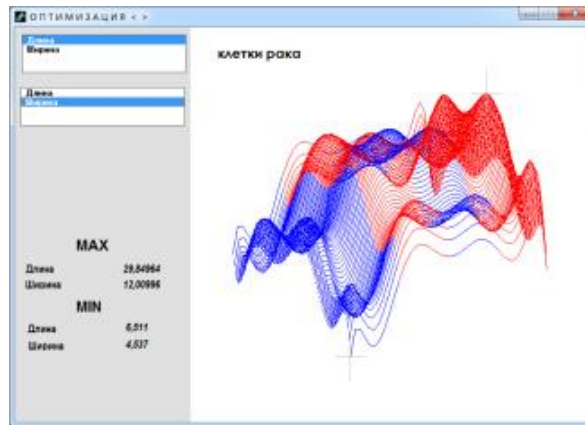
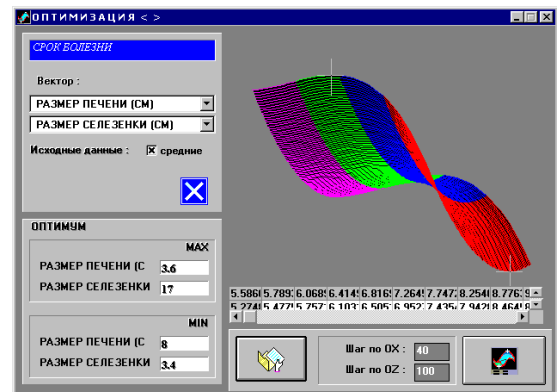


7.9. Одновременная оптимизация по двум и более нелинейным параметрам – функции РФА n-го порядка

Графическая интерпретация оптимизации по одному нелинейному параметру «возраст» (см. ниже)



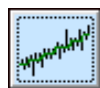
Графическая интерпретация оптимизации по двум нелинейным параметрам: «размер печени» и «размер селезёнки» (см. ниже)



7.10. Мониторинг результатов РФА

Наглядное сравнение расчётных и опытных значений

После получения весовой функции РФА всегда имеется возможность визуально и математически посмотреть соответствие опытных и рассчитанных по функции РФА значений.

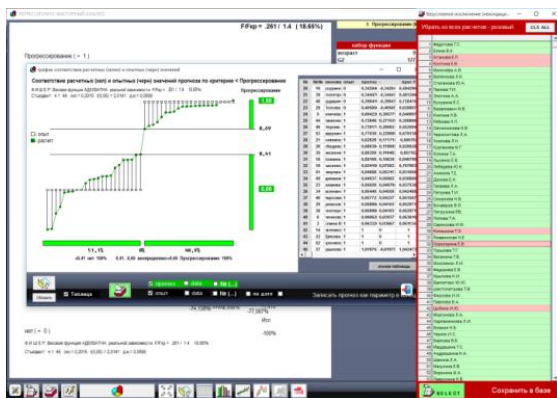


Для числовых критериев

Визуальную картину соответствия опытных и расчетных значений можно просмотреть на форме РФА по кнопке <График>.

Графики соответствия расчётных и опытных значений по числовому критерию. Реальные значения по опытам обозначены дискретно (точки).

Для качественных параметров-критериев (параметров результатов) в системе «PLANETA» («ФАКТОР») также имеется форма полной визуализации графической картины областей 100% прогноза и зон неопределённости (см. рис. ниже):



Для качественных критериев

Зона $< 0,27$ и $> 0,71$

Зона – 100% прогноза состояния «рак или не рак» (79,2% случаев)

В прогнозе:

– более 0,69 – «прогрессирование» (48,9% случаев);

– менее 0,41 – «нет прогрессирования» (51,1% случаев);

– зона 0,41–0,69 – зона неопределенного состояния: присутствуют оба состояния (в 0% случаев).

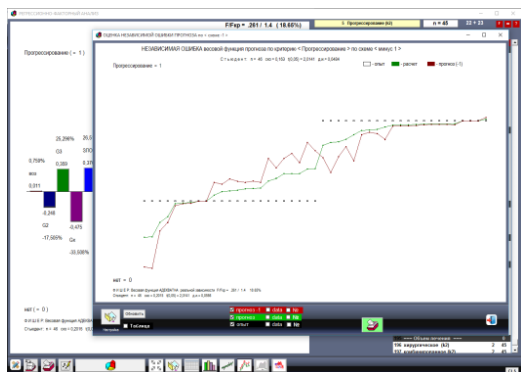
По данному графику наглядно представлено, что в 79,2% случаев прогностическая функция РФА будет однозначно определять состояние «рак» (более 0,71) – «не рак» (менее 0,27) и в 20,8% потребуются дополнительные исследования для уточнения состояния.

Непрерывная функция прогнозирования имеет более точную интерпретацию состояния «рак – не рак», чем качественная дискретная двузначная маркировка состояния «рак – не рак».

Проверка практической ошибки прогноза на практике

Проверка соответствия расчётных и опытных значений по схеме «Минус один»

Такая проверка показывает, как будет проходить прогноз на новых случаях, не участвовавших в получении функции прогноза РФА, и позволяет определить ошибку прогноза на практике.

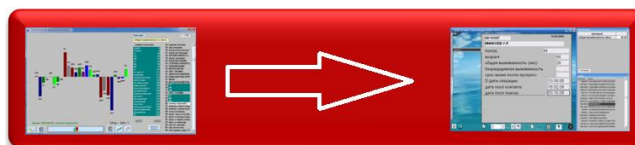


Соответствие опытных и расчётных значений по схеме «Минус один»: РФА строится без пациента – и для этого пациента считается прогноз как для вновь прибывшего.

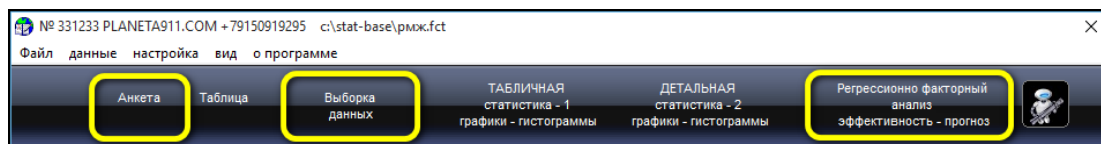
Данная опция позволяет графически показать, как будет выглядеть ошибка прогнозируемых и реальных значений критерия для вновь прибывающих больных, которые не участвовали в расчёте прогностической функции, и оценить ошибки в реальном прогнозе.

8. Прогноз. Создание систем высокоточного прогноза на весовых функциях РФА

8.1. Создание систем высокоточного прогноза на весовых функциях РФА



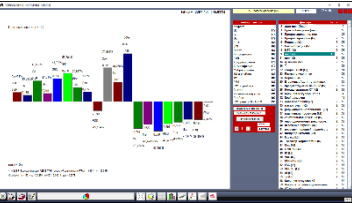
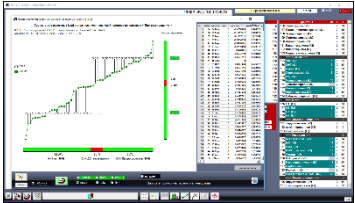
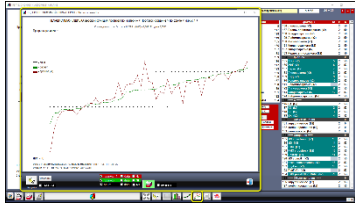
1. Получаем весовую функцию влияния параметров на результат.
2. Добавляем ее в список автоматического прогноза анкеты ведения данных.
3. Получаем автоматический высокоточный прогноз результата при заполнении данных текущих или значений прогнозируемых ситуаций.



Два этапа прогноза для любых процессов:

- первый этап: прогноз принадлежности к однородной группе;
- второй этап: оптимизация воздействий и индивидуальный прогноз уже внутри однородной группы.

8.3. Прогноз по критерию «Качество» (0–1). Медицина «Прогрессирование» **VIDEO MEDICINE**

Функция анализа (РФА)	Соответствие расчетных и опытных значений	Соответствие прогнозируемых значений по «схеме 1»
		
VIDEO	VIDEO	VIDEO

ВНИМАНИЕ! В таблице расчётов за зеленым столбцом коэффициентов идут столбцы диапазонов допустимых минимальных и максимальных значений параметров – эти значения нужно обязательно указывать в инструкции к прогнозу. При выходе параметров за пределы диапазонов прогноз нужно считать условным.

8.4. Создание систем прогноза по рабочим базам данных

Гистограмма решения позволяет определять приоритет значимостей параметров на результат и решать задачи оптимизации и выбора оптимальных действий по максимизации (минимизации) критериев (результатов) процесса, а запись весовых функций в анкету позволяет получать автоматический прогноз значений всех критериев автоматически при заполнении требуемого для прогноза комплекта данных.

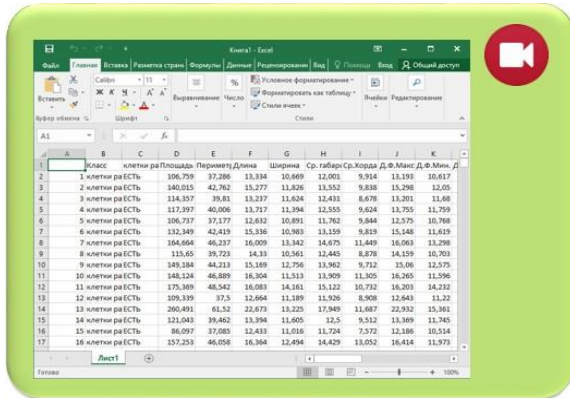
При создании системы прогноза к каждой позиции (весовой функции) прогноза необходимо прилагать подробную инструкцию применения этой системы прогноза, в которой должны быть чётко и однозначно перечислены все ограничения, которые были использованы при создании этой конкретной весовой функции прогноза. Обязательно должна прилагаться таблица со списком параметров с указанными допустимыми диапазонами их изменений, которые можно посмотреть в расчётной таблице РФА.

В случаях, когда параметры случая выходят за диапазоны допустимых параметров, прогноз можно указывать как условный и прилагать значения параметров, вышедших за рамки допустимого диапазона прогноза.

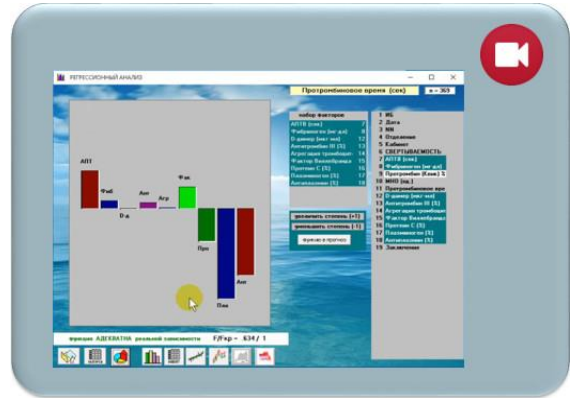
[Ссылка на видео](#)

8.5. Демонстрационные видео по созданию систем прогноза по линейным и нелинейным функциям РФА

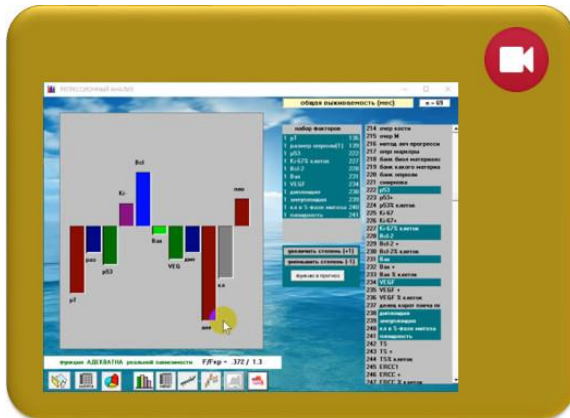
1. Создание систем прогноза по таблицам данных признаков картинок клеток с микроскопа (лаборатория Боткинской больницы)



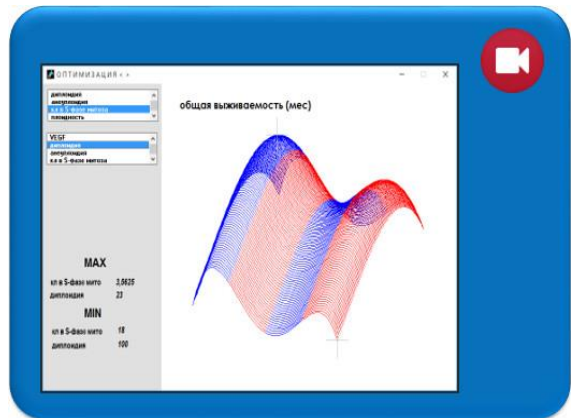
2. Создание систем высокоточного прогноза по данным клинико-диагностических лабораторий группы гемостаза (РОНЦ)



3. Создание систем высокоточного прогноза по медицинским научным исследовательским базам НМИЦ им. Н.Н. Блохина



4. Создание систем высокоточного прогноза на нелинейных весовых функциях любого порядка НМИЦ им. Н.Н. Блохина



Комментарии к видео:

1. Создание систем высокоточного прогноза по таблицам данных в Excel (лаборатория Боткинской больницы).

Это видео показывает, как легко и просто создать систему высокоточного прогноза по любой таблице данных Excel, полученной по программе распознавания и определения геометрических элементов и фигур по графическим и яркостным признакам картинки – позволяет автоматизировать тяжелый процесс визуальной диагностики по картинкам с микроскопа и рентгенограммам.

2. Создание систем высокоточного прогноза по данным лабораторий на примере группы гемостаза (РОНЦ):

– показывает, как автоматизировать и сделать прогноз дорогостоящих анализов по комплексу стандартных анализов – хорошая вспомогательная информация по перечню анализов, которые нельзя получить на имеющемся оборудовании;

– позволяет дублировать результаты анализов и возможность избежать грубых ошибок лабораторных систем.

3. Создание систем высокоточного прогноза по медицинским научным исследовательским базам (НМИЦ им. Н.Н. Блохина).

Простое демонстрационное видео с очевидными результатами для демонстрации работы РФА в научной части без детализаций, что вы обычно получаете в общем случае с первых расчётов без шлифовки результатов по выборкам.

4. Создание систем высокоточного прогноза на нелинейных весовых функциях любого порядка (НМИЦ им. Н.Н. Блохина).

Демонстрационное видео по работе с весовыми функциями РФА n -го порядка с очевидными результатами. Для демонстрации процесса работы нелинейными функциями РФА любой степени в научной части.

8.6. Основополагающие условия оптимизации и прогнозирования

Прогнозирование – возможность определить значение результата при заданных вариантах течения процесса, в том числе и при различных воздействиях.

В системе «ФАКТОР-ПРОГНОЗ» для прогнозирования используются весовые функции РФА, проходящие по условию адекватности реальной зависимости по F -критерию Фишера. Чем меньше значение F -критерия по отношению к критическому $F_{кр}$, тем точнее прогнозируемая величина и меньше зона неопределенности. Прогностические функции высокой точности могут быть получены для выборок групп данных при обязательном соблюдении условий ограничений выборки для прогнозируемого случая.

Принципы работы системы «ФАКТОР-ПРОГНОЗ» как системы высокоточного прогнозирования:

- прогноз по весовым функциям, полученным по всему качественно-числовому объему информации в единой схеме без ограничений на объем и состав информации;
- прогнозирование по весовым функциям, полученным по собственному текущему рабочему материалу, с учетом постоянно меняющихся собственных внешних и внутренних условий работы.

Одномерная оптимизация

В случае, когда необходима оптимизация параметров воздействия без учета эффектов совместного действия, то достаточно однопараметрической оптимизации.

Многомерная оптимизация

При необходимости учета эффекта совместного действия (при одновременном воздействии различных комбинаций доз нескольких препаратов) в базе данных формируются дополнительные столбцы данных из произведений значений этих параметров, весомой вклад которых в критерий и будет определять дополнительно эффект совместного действия. В этом случае оптимальную комбинацию значений двух воздействий (доз) необходимо определять на трёхмерной поверхности, где по X и Y – значения параметров воздействия (доз), по Z – значение критерия (результата).

Оптимум (глобальный экстремум), значения X и Y , при которых $Z = Z_{\min}$ (\max), полученный с учетом совместного действия, не будет совпадать с оптимумами отдельно по параметрам.

8.7. Необходимый объем информации для решения поставленных задач

Можно подходить к процессу обработки информации по принципу «как есть». Завести в базу данных параметры «как есть» и решать задачи по оценке эффективности воздействий. В этом случае будут решаться задачи по конкретным ситуациям. Это приемлемо, когда процесс ограничивается стандартными повторяющимися ситуациями, но этого недостаточно при работе с многовариантными процессами. Многовариантные процессы необходимо рассматривать с более широких позиций, добавляя в базу данных правила поведения процесса во внешней среде.

С точки зрения математического объема данных РФА можно проводить в минимуме с матрицы 10 операций на 5–10 параметров. В норме (разбивка базы данных на выборки) – 30 операций на 5–20 параметров. При увеличении объема данных в однородной выборке точность расчета возрастает. В системе «ПЛАНЕТА» («ФАКТОР-ПРОГНОЗ») нет ограничений на максимальный объем операций и факторов в схеме РФА. Все остальные параметры, не указанные в расчете, косвенно войдут и распределятся между указанными, поэтому желательно включать в расчёт все ведущие воздействия, определяющие значение выбранного критерия.

Решение задач оптимизации: при решении задачи поиска оптимальных воздействий по максимизации результата в минимуме достаточно базы данных по этим воздействиям и результату. Точность и адекватность решения будет увеличиваться при расширении объема используемой информации.

Решение задач прогнозирования: при решении задачи прогнозирования результата, в базе данных необходимо иметь главные ключевые параметры, определяющие значение результата.

9. Примеры научных работ в РФА в аналитическом программном комплексе «ФАКТОР-ПРОГНОЗ»

1 этап

Получение весовой функции РФА, оценка эффективности влияния параметров на результат.

2 этап

Оценка соответствия расчётных и опытных значений.

3 этап

Оценка ошибки прогноза по вновь прибывшим по схеме «Минус 1» – опыт убирается из расчёта РФА и после расчёта подставляется в функцию РФА, которая была определена без этого опыта, и так в цикле по всем опытам для расчёта ошибки прогноза по вновь прибывшим.

4 этап

Этап записи весовой функции РФА в регистрационную анкету ведения данных для автоматического прогноза после набора всего комплекта данных для прогноза по данному критерию.

5 этап

Автоматический комплексный прогноз по всем критериям в панели прогноза.

9.1. пример оформления научной статьи с РФА по оценке значимости параметров и созданию системы прогноза

Прогнозирование синдрома задержки роста плода у беременных высокого риска

Член-корр. РАН, проф., д. м. н. И.В. Игнатко, к. м. н. М.М. Мирющенко

Впервые термин «Синдром задержки роста плода» (СЗРП) появился в 1973 году на Первом международном симпозиуме, посвященном этой проблеме. Последние достижения медицины прояснили этиологию и патогенез данного осложнения перинатального периода. Но до настоящего времени синдром задержки роста плода остается одной из ведущих причин перинатальной заболеваемости и смертности. В 2013 году задержка роста плода названа Американской коллегией акушеров и гинекологов «наиболее распространенной и сложной проблемой современного акушерства» (Creasy Robert K., Resnik R., Iams J.D., 2014).

Новорожденные, родившиеся с СЗРП, страдают от дыхательных расстройств, гипогликемии, внутрижелудочковых кровоизлияний, нарушений функции ЦНС и паренхиматозных органов, что в перспективе ведет в детском возрасте к церебральному параличу, задержке развития и поведенческой дисфункции, а у взрослых – к повышенному риску развития ожирения, сахарного диабета II типа, артериальной гипертензии, дислипидемии, ИБС и метаболическому синдрому (Barker D.J.P., 1998; Unterscheider J., O'Donoghue Keelin, Malone F.D., 2015; Громова О.А., 2015). Крайней степенью тяжести СЗРП является антенатальная гибель плода (Стрижаков А.Н., Игнатко И.В., Тимохина Е.В., Белоцерковцева Л.Д., 2014; Попова Ю.Ю. 2013; Карданова М.А., 2015).

Цель. Разработка комплексной персонифицированной системы обследования беременных с целью предикции, ранней диагностики и превенции синдрома задержки роста плода на основании оценки клинико-анамнестических и лабораторно-инструментальных факторов риска для снижения перинатальной заболеваемости и смертности.

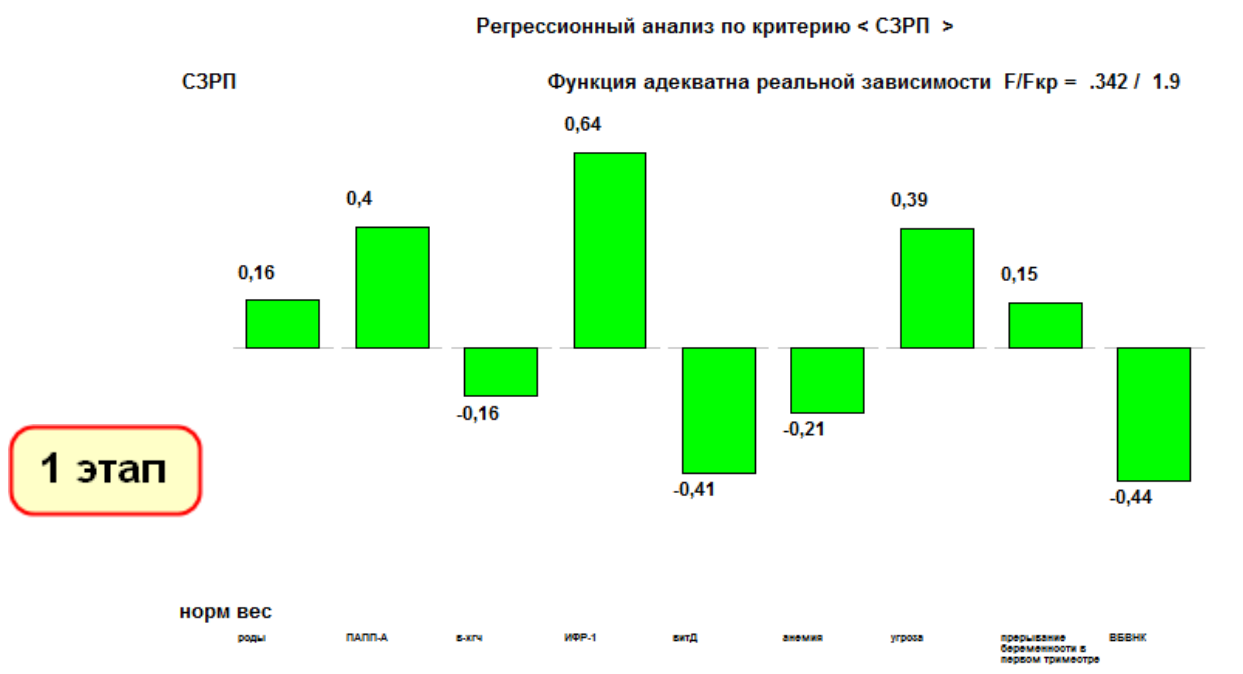
Пациенты и методы. В соответствии с поставленной целью нами проведено комплексное проспективное обследование 184 беременных с первого триместра гестации, 184 новорожденных, а также проведен анализ перинатальных и материнских исходов гестации. Основную группу составили 146 женщин, беременность которых осложнилась плацентарной недостаточностью и закончилась рождением живых детей с синдромом задержки роста плода (СЗРП); группу сравнения – 38 женщин, беременность которых протекала без признаков плацентарной недостаточности и закончилась рождением живых доношенных детей. Оценивались социальное и семейное положение, место жительства; наличие вредных привычек (курение, употребление алкоголя, прием наркотических веществ); соматический и акушерско-гинекологический анамнез; течение настоящей беременности; особенности родоразрешения; состояние новорожденного. В обеих группах проводилась оценка результатов первого перинатального скрининга и уровня сывороточных маркеров СЗРП (РАРР-А, β -ХГЧ, ИФР-1 и витамин D) в первом триместре беременности. Беременным также проводилось комплексное динамическое ультразвуковое фетометрическое, плацентографическое и доплерометрическое исследования кровотока в маточных артериях, артерии пуповины, средней мозговой артерии плода и аорте плода.

Результаты. Для оценки значимости каждого из изученных факторов риска в формировании СЗРП проводился РФА. Проанализировав все факторы риска, мы смогли выявить совокупность наиболее значимых факторов. Самым важным фактором оказался уровень ИФР-1, а минимальным – количество β -ХГЧ, число родов в анамнезе и наличие прерывания беременности в первом триместре в предыдущих беременностях. В группу отрицательных факторов вошли: уровни β -ХГЧ, витамина D ($1,25(\text{OH})_2$ витамина D), анемия, заболевания почек. В группу положительных факторов вошли: РАРР-А, количество родов в анамнезе, угроза прерывания беременности, уровень ИФР-1 и наличие прерывания беременности в первом триместре в предыдущих беременностях.

Проведение РФА позволило получить функцию отклика, адекватную реальной зависимости. Адекватность реальной зависимости $F/F_{кр} = 0,342 / 1,9$; $n = 29$.

Самым важным фактором оказался ИФР-1, а минимальным – β -ХГЧ, количество родов в анамнезе и наличие прерывания беременности в первом триместре в предыдущих

беременностях. В группу отрицательных факторов вошли: β -ХГЧ, витамин D, анемия и варикозная болезнь вен нижних конечностей. В группу положительных факторов вошли: ПАПП-А, количество родов в анамнезе, угроза прерывания беременности, ИФР-1 и наличие прерывания беременности в первом триместре в предыдущих беременностях. Для наглядности на рисунке 1 представлена гистограмма, отражающая соответствие расчетных прогностических факторов, полученных по весовой функции РФА, и опытных значений.



1 этап

Рисунок 1. Гистограмма весовой функции РФА по критерию «СЗРП»

В результате проведенных исследований выявлены предикторы высокого риска развития СЗРП. С использованием метода РФА разработана математическая модель раннего прогнозирования СЗРП, заключающаяся в определении прогностического индекса (PI) по формуле:

$$PI = -0,43739X_1 - 0,016018X_2 - 0,212004X_3 - 0,108084X_4 + 0,1497046X_5 + 0,079163X_6 + 0,3928565X_7 + 0,3081384X_8 + 0,00327X_9 - 0,706$$

X_1 – отсутствие/наличие варикозной болезни вен нижних конечностей (0/1);

X_2 – уровень витамина D, нг/мл;

X_3 – отсутствие/наличие анемии (0/1);

X_4 – уровень β -ХГЧ в первом перинатальном скрининге, МоМ;

X_5 – отсутствие/наличие прерывание беременности в первом триместре в анамнезе (0/1);

X_6 – количество родов в анамнезе;

X_7 – отсутствие/наличие угрозы прерывания беременности (0/1);

X_8 – уровень ПАПП-А в первом перинатальном скрининге, МоМ;

X_9 – уровень ИФР-1, нг/мл;

0,706 – constanta.

График соответствия расчетных прогностических значений по весовой функции РФА и опытных значений по критерию «СЗРП»

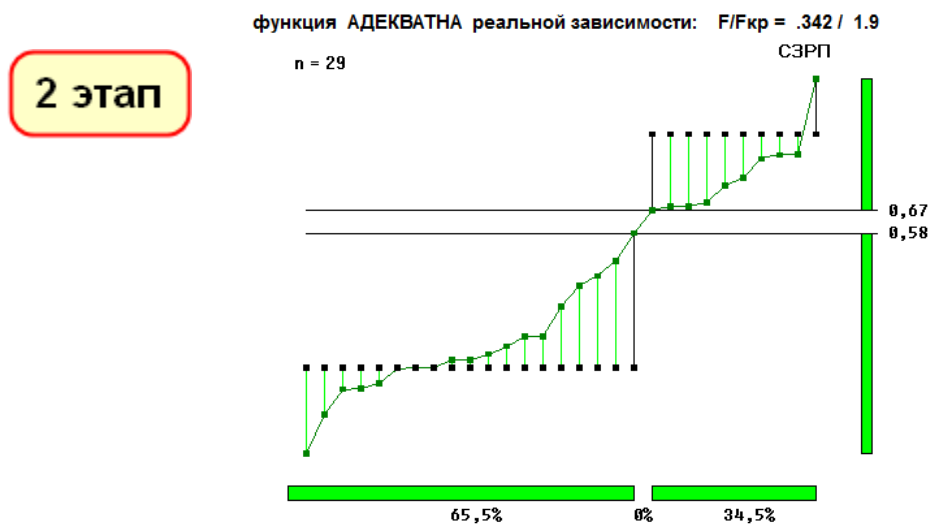


Рисунок 2. При $PI \leq 0,58$ прогнозируется низкий риск развития СЗРП, а при $PI \geq 0,67$ можно сделать заключение о высоком риске развития СЗРП

На основании полученных результатов исследования разработан алгоритм обследования беременных женщин в ранние сроки беременности с целью выявления групп риска по развитию СЗРП (см. рисунок 3). В первом триместре беременности рекомендуется проводить определение прогностических критериев, позволяющих дифференцировать физиологическое и патологическое ее течение в дальнейшем. При попадании пациенток в зону неопределенности по риску развития СЗРП необходимо проводить дополнительное обследование беременных женщин с использованием наиболее информативных показателей и с подсчетом прогностических индексов для определения групп риска формирования синдрома задержки роста плода.

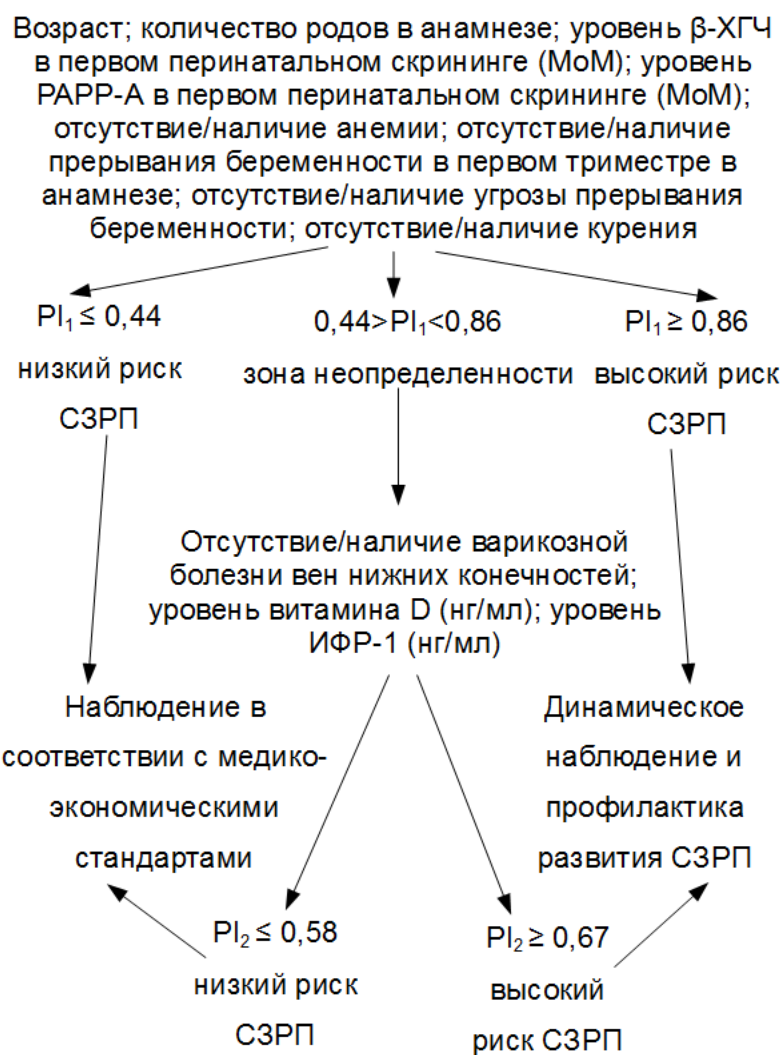


Рисунок 3. Алгоритм обследования женщин в ранние сроки беременности с целью выявления групп риска развития СЗРП

Заключение. Разработанный алгоритм прогнозирования формирования СЗРП у беременных группы риска развития СЗРП позволил значительно повысить раннюю диагностику данного синдрома (риск развития СЗРП в группе высокого риска – 79,35%). Доказана эффективность применения ранней профилактики и медикаментозной коррекции эндотелиальной дисфункции и железодефицитных состояний, что позволило снизить риск развития СЗРП с 79,35% до 27,87%.

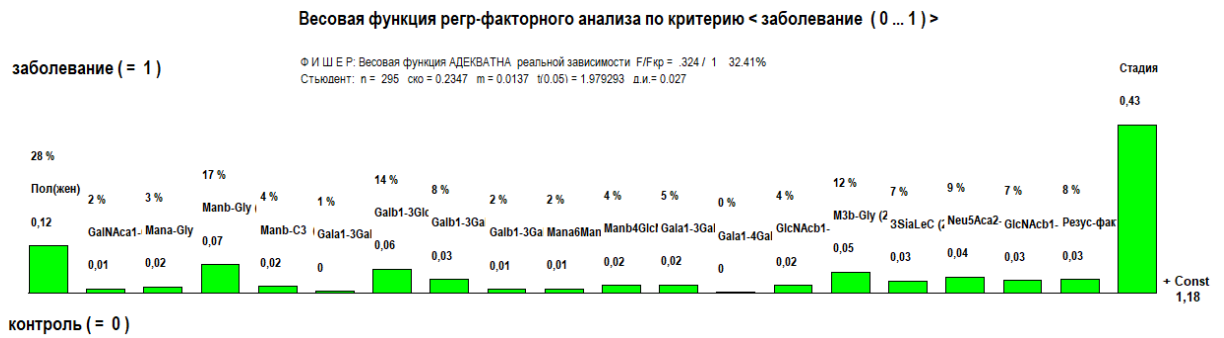
Таким образом, комплексное клиническое, инструментальное и лабораторное обследование беременных высокого риска развития СЗРП дает возможность не только своевременно диагностировать данные осложнения, но и со значительной вероятностью прогнозировать их формирование еще на предгравидарном этапе, что позволит провести комплекс преемственных коррекционных мероприятий.

9.2. Пример работы РФА с переходом от менее точных дискретных оценок состояния больного к более точным аналоговым оценкам в диагностике раком желудка

(по материалам исследования уровней антигликановых антител у больных раком желудка и практически здоровых доноров.

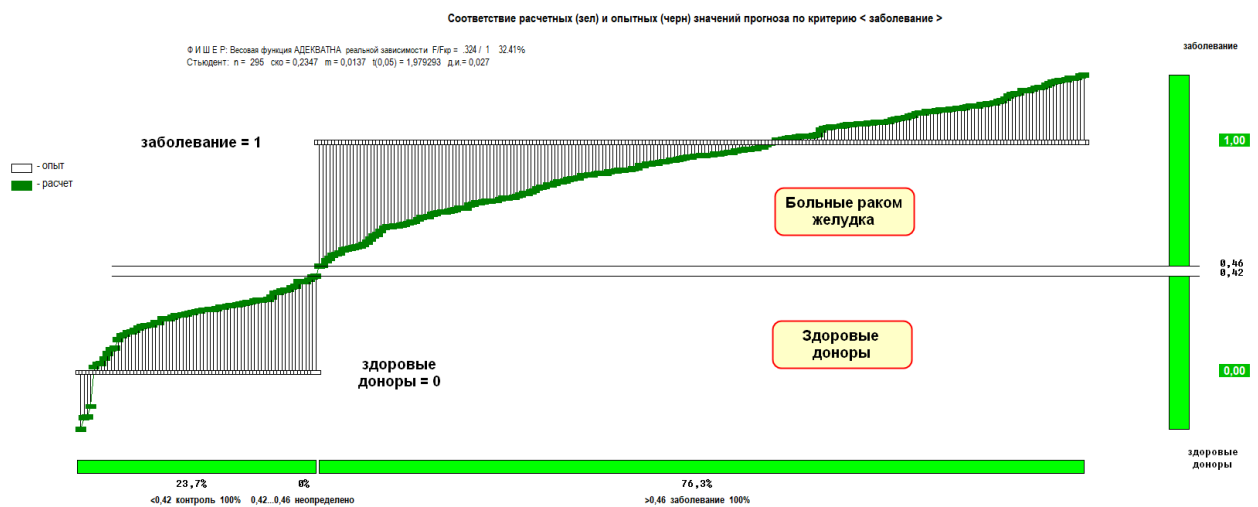
Колосов Е.А., Никулин М.П., отделение абдоминальной онкологии № 1)

1 этап



2 этап

Прогностическое значение уровня антигликановых антител для больных раком желудка

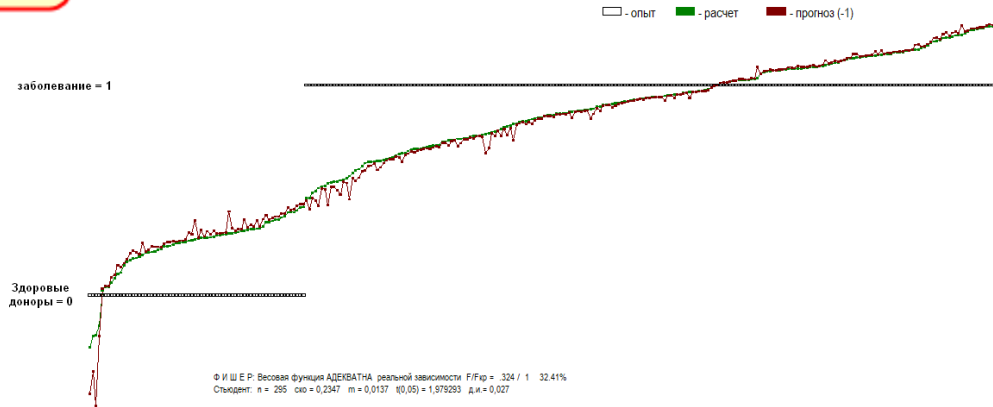


Большое количество параметров оценки в единой математической схеме РФА обеспечивает высокую точность и минимальные ошибки прогноза.

3 этап

НЕЗАВИСИМАЯ ОШИБКА весовой функция прогноза по критерию < заболевание > по схеме < минус 1 >

Стьюдент: n = 295 ско = 0,0342 m = 0,002 ц(0,05) = 1,979293 д.и. = 0,0039



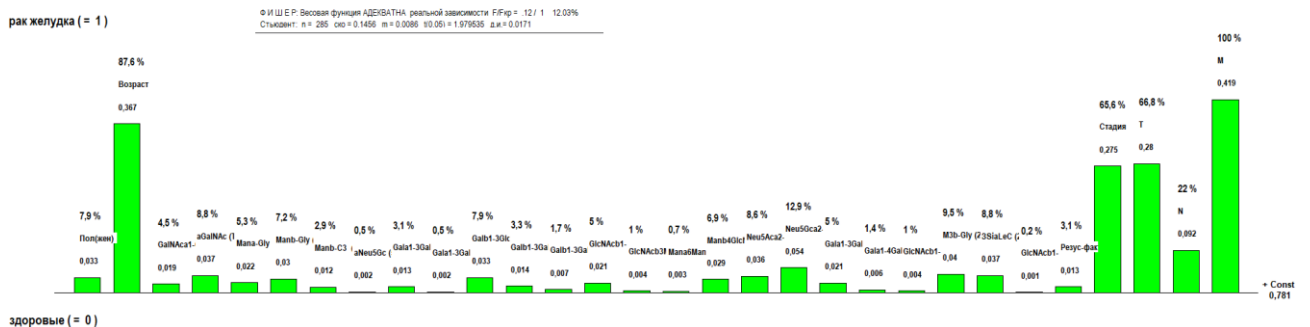
Насколько увеличится адекватность функции реальной зависимости в диагностике больных раком желудка при добавлении к уже имеющимся 27 диагностическим параметрам ($F/F_{кр} = 0.324 / 1$ 32,41%) ещё трёх параметров, составляющих значение стадии T, N, M ($F/F_{кр} = 0.12 / 1$ 12,03%).

Адекватность функции реальной зависимости улучшилась в 2,7 раза: с $F / F_{кр}$ от 0,324 / 1 до 0,12 / 1 (с 32,4% до 12%), с зоной разделения 0,26–0,59.

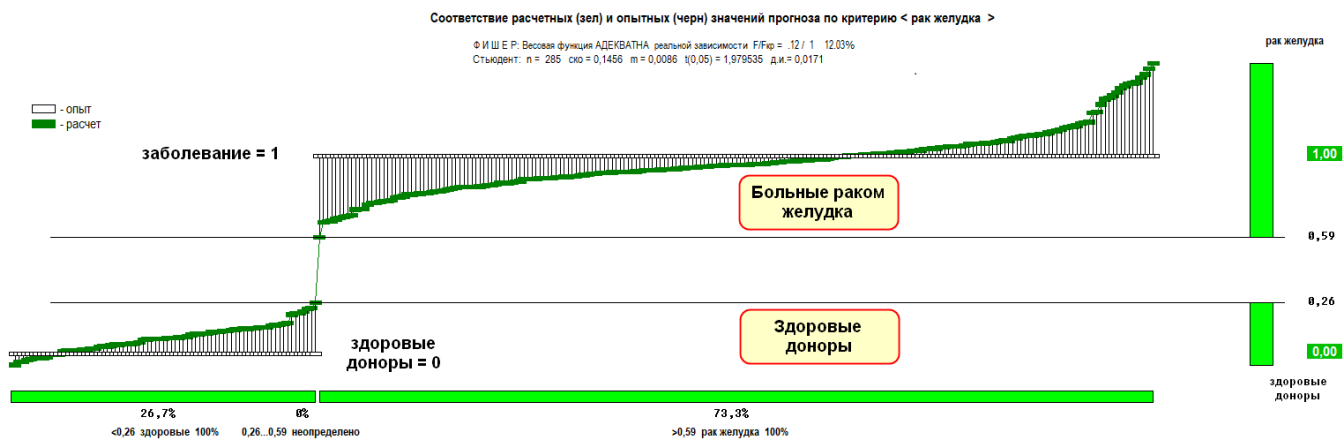
1 этап

Весовая функция РФА по критерию «Рак желудка» (0–1)

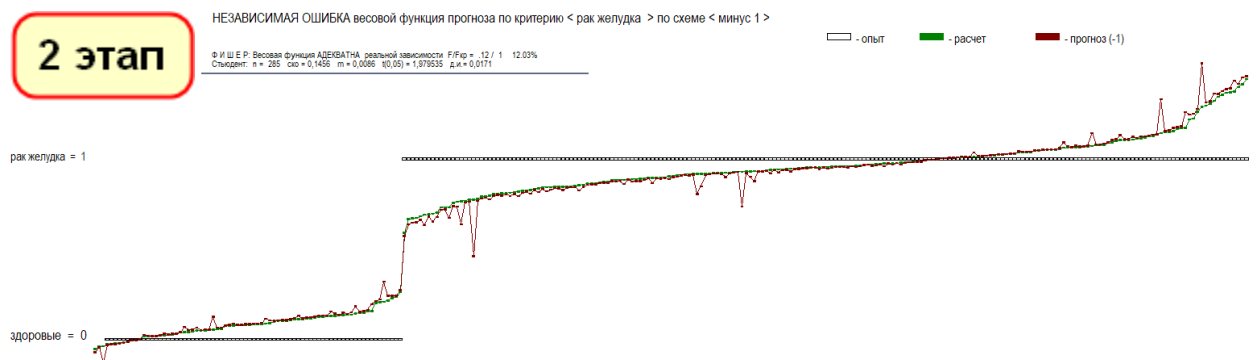
Весовая функция регр-факторного анализа по критерию < рак желудка (0 ... 1) >



Прогностическое значение уровня антигликановых антител для больных раком желудка



Большое количество пациентов и параметров оценки в единой математической схеме РФА обеспечивает высокую точность и минимальные ошибки прогноза



9.3. РФА в области лучевой диагностики в онкологии «МРТ в диагностике метастатического поражения печени и оценка эффективности лечения больных нейроэндокринными опухолями»

Акад. РАН, д. м. н., проф. Б.И. Долгушин, д. м. н., проф. В.А. Горбунова, к. м. н.
Лаптева М.Г. Дисс., Москва, 2021. – 121 с.
НМИЦ онкологии им. Н.Н. Блохина

РФА оценки значимости параметров по качественному критерию «Локализация первичной НЭО» по всему потоку без выделения однородных групп не позволяет получить весовую функцию высокой адекватности реальной зависимости ($F/F_{кр} = 0,439/1,3$), при $n = 65$ ошибка 33,8% от критического значения ошибки по F-критерию Фишера.

1 этап

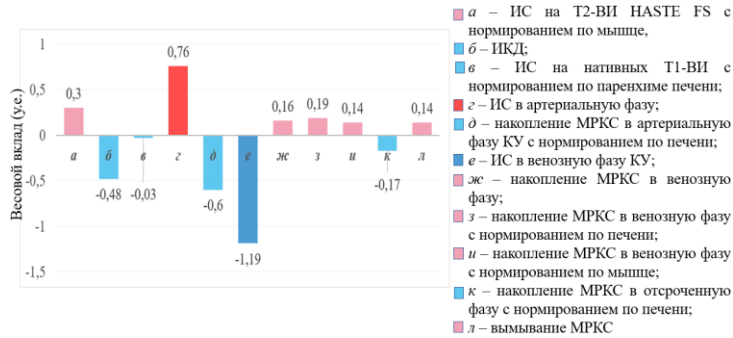


Рисунок 1. Весовые вклады РФА факторов по критерию «Локализация первичной НЭО»

2 этап

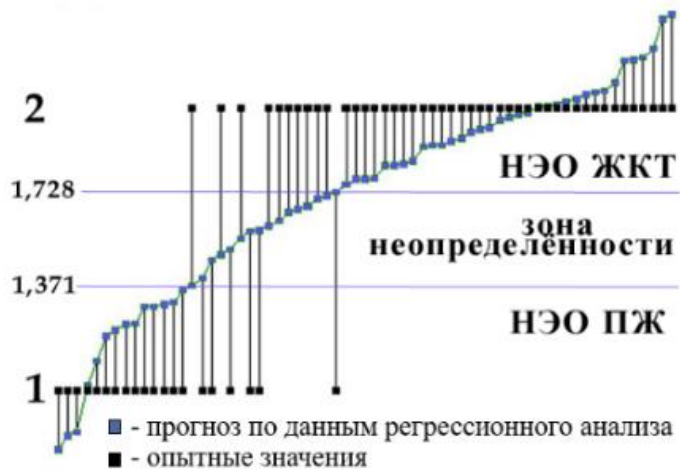


Рисунок 2. График соответствия расчётных (РФА) и опытных значений по локализации первичной НЭО по данным РФА. Анализ РФА по всему потоку без выделения однородных групп не позволяет 100-процентно разделить состояния критерия и получить 100-процентный прогноз (есть зона неопределённости 1,371–1,728)

5 этап



Рисунок 3. Интерфейс прогноза степени злокачественности метастазов НЭО

Автоматический прогноз диагноза

Автоматический прогноз локализации первичной опухоли и степени злокачественности МТС после введения необходимых МРТ-характеристик метастазов в печени (рисунок 3) после заполнения данных анкеты.

9.3. Пример выделения однородных групп по разным критериям

Возможности РФА позволяют решать следующие задачи:

- 1) задачи выделения однородных групп по критериям системных нарушений в биологических процессах (например, медицинские нарушения, онкология...);
- 2) задачи прогноза принадлежности системного нарушения к своей однородной группе;
- 3) задачи оценки влияния параметров процесса на критерии процесса внутри однородной группы.

1 и 2 пункты могут решаться сразу при анализе исследовательских групп.

3 пункт требует, помимо определения общей картины влияния параметров на критерии, проводить дополнительные исследования по определению функционала параметра в конкретной весовой функции РФА – в этой конкретной функции РФА параметр дает свой вклад в критерий или он в этой весовой функции РФА является индикатором других параметров, которые не указаны в этой конкретной весовой функции.

Выводы: Прежде чем проводить оценочные расчёты влияния параметров на комплекс критериев оценки эффективности любых процессов в областях техники, финансов, экономики, медицины, онкологии и т. д., необходимо иметь в виду соблюдение принципа однородности группы по исследуемым критериям.

В случаях, когда в решении не удастся получить весовые функции со 100-процентным разделением состояний булевых критериев и 100-процентным прогнозом критерия, это свидетельствует о присутствии в выборке нескольких разных процессов, что требует предварительно провести исследования на разделение выборки на однородные группы, внутри которых вы уже получите 100-процентное разделение по состояниям критерия и 100-процентный прогноз.

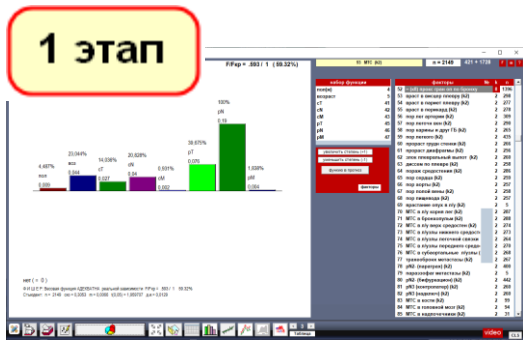
Правильная разбивка общего потока на однородные группы может осуществляться, только начиная с общего потока, но отсутствие финансовых возможностей полной системной диагностики всех систем организма для всего потока пациентов делает такой подход практически не реализуемым на практике.

На практике, при проведении исследовательских научных работ, например в области онкологии, используются небольшие текущие рабочие базы данных подразделений без выделения однородных групп по общему потоку. В этом случае можно улучшать условие однородности процесса, исключая до 5% случаев, явно не вписывающихся в закономерности большинства группы. Это происходит вследствие отсутствия точной диагностики на основе однопараметрических методов оценки. Также можно улучшать ситуацию добавлением новых ключевых параметров в базу данных, для улучшения идентификации состояний критерия.

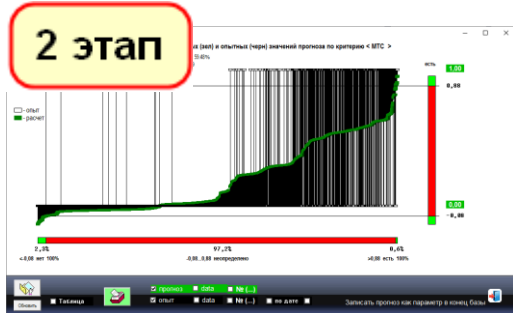
9.4.1. Пример подобия и различия решений РФА по критерию наличия метастазов для общего потока пациентов

n = 2149 и однородной группы по критерию наличия метастазов при 100-процентном прогнозе (пример подготовлен по материалам научных баз данных НИИ КО)

Общий поток пациентов n = 2149. Оценка в РФА всего потока данных по критерию «Наличие метастазов» не позволяет получить высокую степень адекватности реальной зависимости по критерию Фишера $F/F_{кр} = 0,593 / 1$ (59,3%), но позволяет определить общие тенденции по критерию «Наличие метастазов»

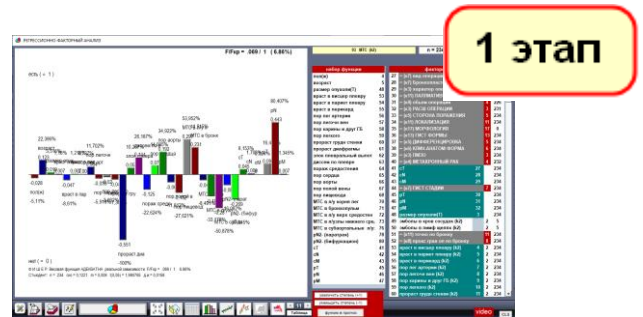


Даже по всему потоку пациентов прослеживается однозначная зависимость по критерию «Наличие метастазов» с выделением на графике различных однородных групп по критерию «Наличие метастазов»

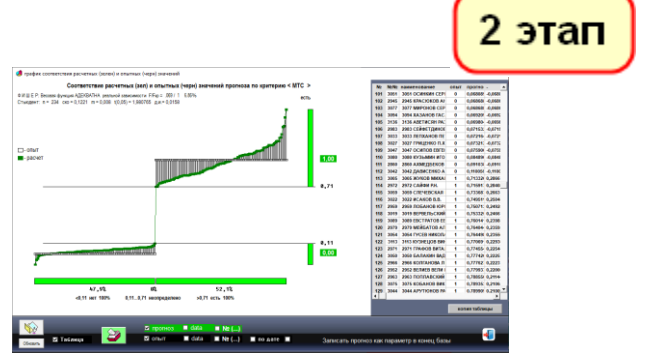


Прогноз по вновь прибывшим по схеме «Минус 1» показывает ошибку прогноза по вновь прибывшим, принадлежность к своей однородной группе и степень

Однородная группа пациентов n = 234 по критерию «Наличие метастазов» и получение 100-процентного прогноза – 100-процентного разделения состояний критерия при добавлении параметров из группы критерия при высокой степени адекватности реальной зависимости по критерию Фишера $F/F_{кр} = 0,069 / 1$ (6,86%)



100-процентное разделение состояний и прогноз по весовой функции РФА дополнительно определяет и степень тяжести по шкале критерия «Наличие метастазов»

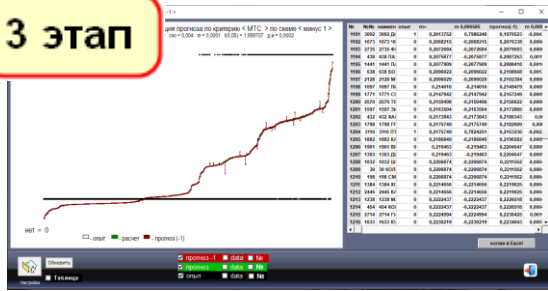


Прогноз по вновь прибывшим по схеме «Минус 1» определяет не только факт «Наличия метастазов», но и степень тяжести системного нарушения по шкале критерия «Наличие метастазов»

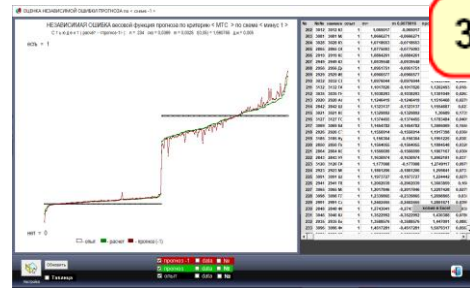
2 этап

«Наличие МТС»
по критерию «Наличие МТС»

3 этап



3 этап



Выводы: РФА по общему потоку данных не позволяет добиться 100-процентного разделения по состояниям критерия, но позволяет обозначить однородные группы для детальных исследований процессов уже внутри однородных групп.

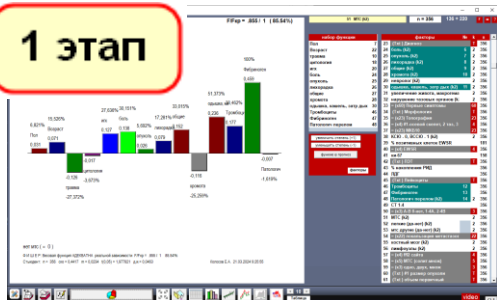
9.4.2. Пример разделения общего потока пациентов по критерию

В примере представлено, как можно разделить общий поток пациентов по критерию МТС с $n = 356$ на однородные группы, добавив в РФА параметры локализации МТС при 100-процентном прогнозе события «Наличие МТС»

(пример подготовлен по материалам научных баз данных НИИ ДОиГ им. Л.А. Дурнова)

Функция РФА без параметров группы локализаций по критерию «Наличие МТС»
 $n = 356$

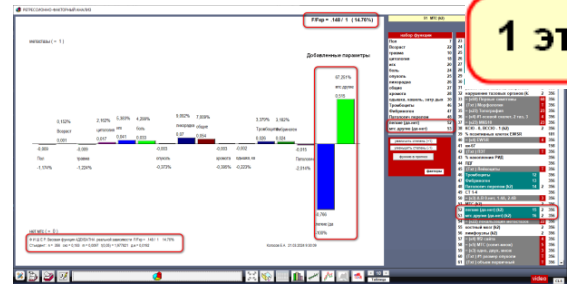
1 этап



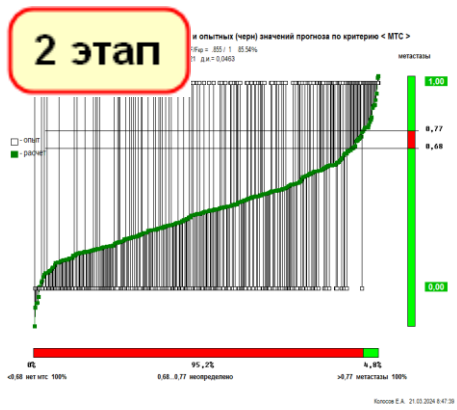
По критерию «Наличие МТС» не наблюдается 100-процентного разграничения групп без МТС и групп с МТС

Разбивка общего потока пациентов на однородные группы и получение 100-процентного прогноза по критерию «Наличие МТС» при добавлении двух параметров из группы критерия ($n = 356$)

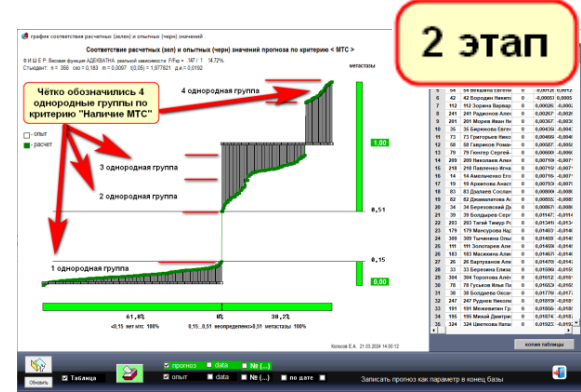
1 этап



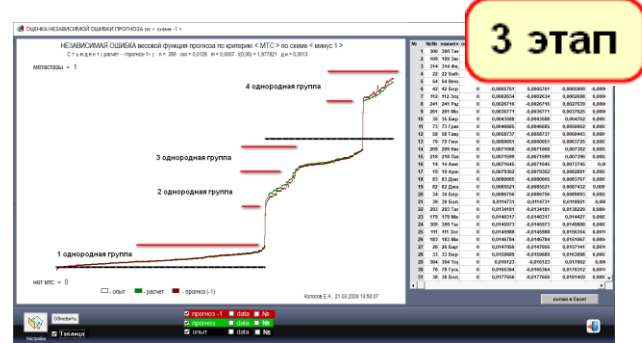
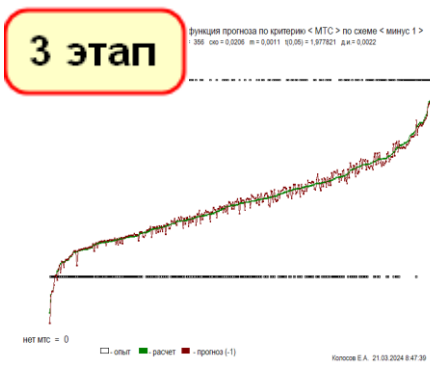
Прогноз по данной весовой функции определяет не только прогноз-оценку по аналоговой шкале МТС, но и принадлежность к своей однородной группе. 100-процентное разделение групп по критерию «Наличие МТС»



Прогноз по вновь прибывшим по схеме «Минус 1». Нет детализации и разделения групп по критерию «Наличие МТС»

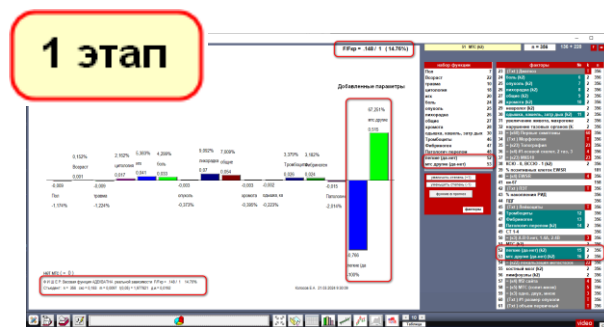


Прогноз по вновь прибывшим по схеме «Минус 1». Есть чёткое разделение и детализация групп по критерию «Наличие МТС» при добавлении в РФА параметров локализации МТС – получаем 100-процентный прогноз наличия МТС и принадлежности к своей однородной группе системных нарушений

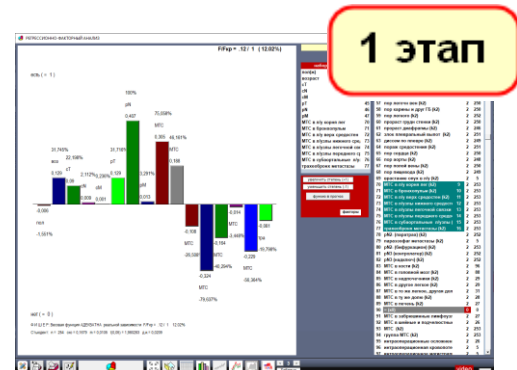


9.4.3. Сравнение результатов исследований по детской БД (НИИ ДОиГ) и взрослой БД (НИИ КО) пациентов по критерию «Наличие метастазов»

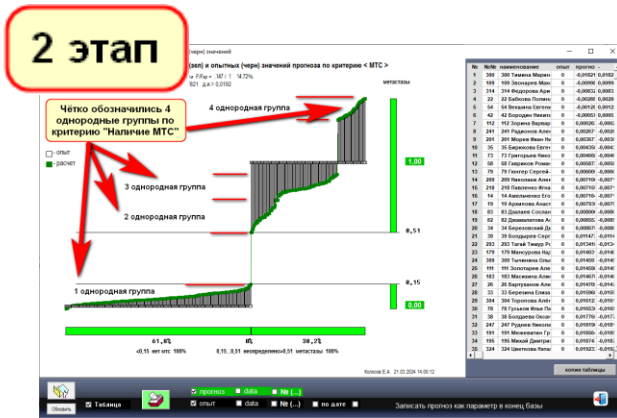
НИИ ДОиГ (детская БД)



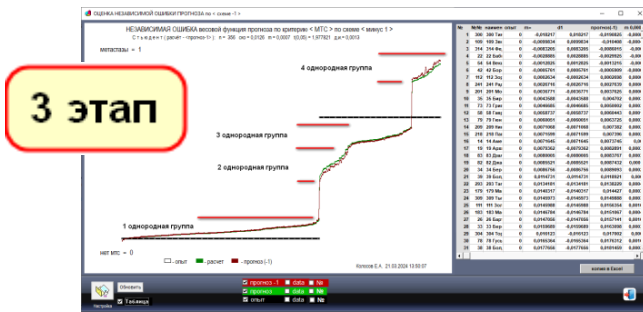
НИИ КО (взрослая БД)



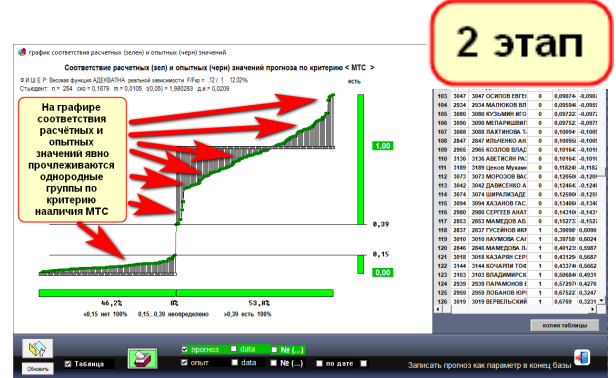
Прогноз по данной весовой функции определяет не только прогноз-оценку по аналоговой шкале МТС, но и принадлежность к своей однородной группе



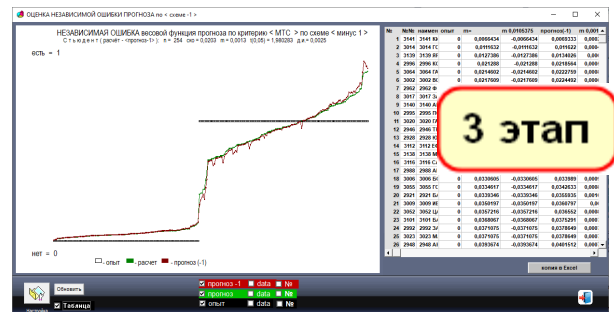
Прогноз по вновь прибывшим по схеме «Минус 1»



Прогноз по данной весовой функции определяет не только прогноз-оценку по аналоговой шкале МТС, но и принадлежность к своей однородной группе



Прогноз по вновь прибывшим по схеме «Минус 1»



Выводы: РФА позволяет по булевым критериям 100-процентно разделять состояния по критерию «Наличие МТС», а также выявлять однородные группы процессов и определять принадлежность пациента к своей однородной группе.

Также мы имеем на порядок более точную оценку тяжести системных нарушений по аналоговой шкале при оценке состояний по критерию «Наличие МТС» – это уже не критерий «Наличие МТС», а более детальный аналоговый критерий «Степень тяжести МТС».

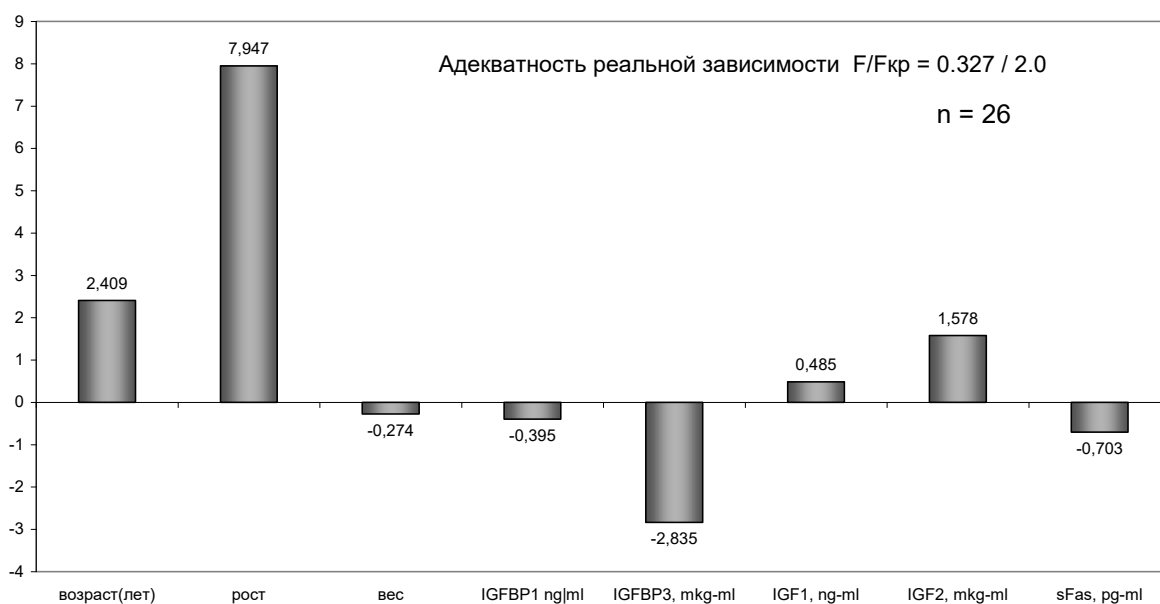
9.5. Дифференцированный подход к диагностике и тактике ведения больных с гиперпластическими процессами эндометрия в постменопаузе

(по материалам совместной научной работы Первого МГМУ им. Сеченова (Москва, РФ) и РОНЦ им. Н.Н. Блохина РАМН (Москва, РФ) в научной статье: академик РАН Стрижаков А.Н., д. м. н., проф. Шахламова М.Н., к. м. н.

Пример показывает, как РФА даже в малых выборках ($n = 26$) позволяет 100-процентно разделять в прогнозе однородные группы: «...Прогноз РФА совпал с гистологическим заключением в 100% наблюдений атрофии и полипов эндометрия. Это позволяет использовать данную весовую функцию РФА для прогноза высокой степени точности...» (с. 11).

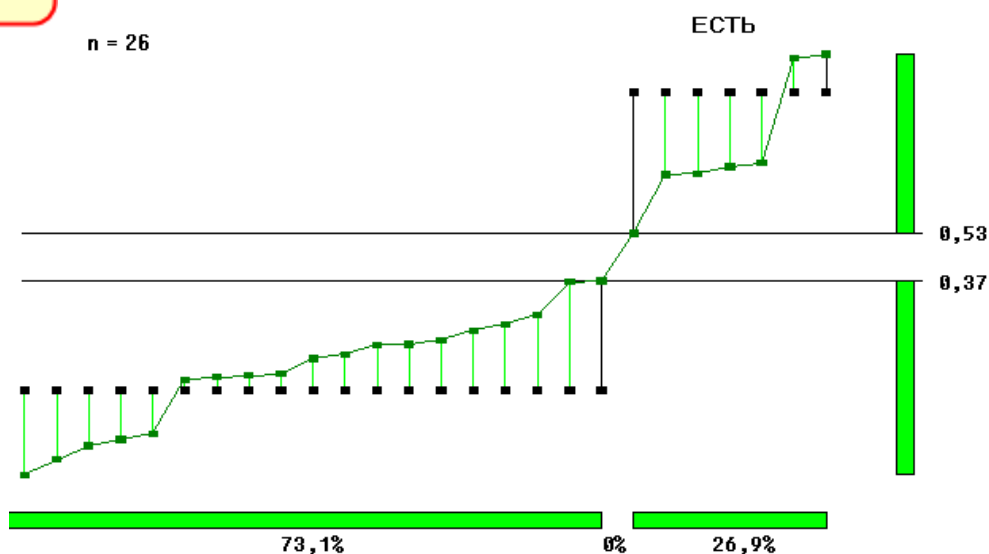
Весовая функция 100% прогноза состояния «Атрофия» – «Полипы эндометрия» в группе «Атрофия» – «Полипы эндометрия» по критерию «Атрофия» при высокой степени адекватности реальной зависимости по F-критерию Фишера $F/F_{кр.} = 0,327 / 2,0$ (16,3% от $F_{кр.}$)

Весовая регрессионно-факторная функция по критерию "атрофия"
в группе атрофия + полипы



Прогноз принадлежности к группе «Атрофия» при $p > 0,53$ или группе «Полипы эндометрия» при $p < 0,37$ (73,1%)

2 этап



В зависимости от того, какие процессы рассматриваются, какой набор параметров используется и какой набор опытных данных имеется, можно получать как 100-процентный прогноз состояния, так и многомерные координаты точки искомого состояния в многомерной области возможных состояний, и с увеличением количества опытного материала результат можно бесконечно уточнять и детализировать для подбора наилучших воздействий.

Таблица весовых вкладов (в условных единицах критерия),
весовых коэффициентов и диапазонов параметров

Наименование параметров	ВЕС (F) в у. е.	Изм (F)	К (вес)	Порядок	Хопт (Fmin)	Хопт (Fmax)	Pmin*	Pmax*
Возраст (лет)	2,409	243,04	0,0367	1	51	81	51	81
Рост	7,947	788,78	0,0486	1	152	172	152	172
Вес	-0,274	36,088	-	1	150	60	60	150
IGFBP1 ng-ml	-0,395	50,32	-	1	32,1	5,6	5,6	32,1
IGFBP3, mkg-	-2,835	289,13	-	1	4,89	2,13	2,13	4,89
IGF1, ng-ml	0,485	56,998	0,0037	1	61,5	244,2	61,5	244,2
IGF2, mkg-ml	1,578	160,69	0,8059	1	1,1	2,87	1,1	2,87
sFas, pg-ml	-0,703	81,402	-	1	988,5	131	131	988,5
< const >	-6,942		-6,942					

* Для передачи окончательных решений для диагностического прогноза необходимо также указывать диапазоны параметров, участвующих в прогнозе и, в случае выхода параметров прогнозируемого случая за границы этих диапазонов** указывать результат прогноза как «условный».

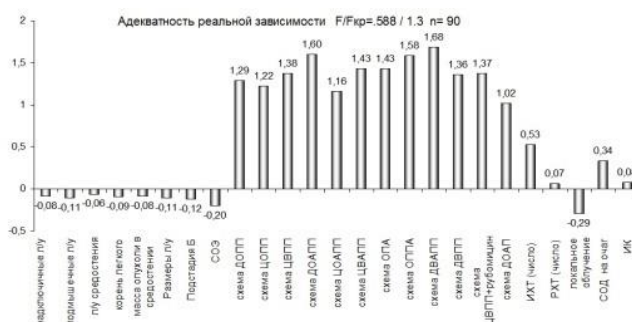
** Допущения «условности» прогноза при выходе значений параметра за границы исследуемой группы, для которой был получен РФА, могут быть свои для каждого параметра и определяться практическим опытом.

9.6. Пример оценки эффективности схем лечения

(из материалов докторской работы «Стратегия лечения лимфогранулематоза у детей», академик РАМН, д. м. н., проф. Л.А. Дурнов, д. м. н. Н.А. Сусулева) НИИ ДОиГ РОНЦ РАМН

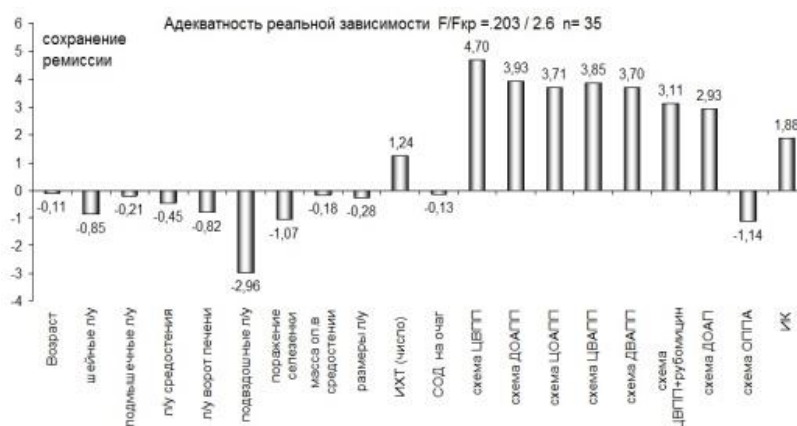
Весовая функция РФА по критерию «Сохранение ПР» для II стадии СК

1 этап



Четко прослеживается высокая эффективность схем лечения по критерию «Положительная ремиссия» на II стадии. Значение значимости факторов схем лечения превышают значения диагностических факторов в 10–15 раз. Также мы видим приоритет значимостей эффективности схем лечения для II стадии СК.

Весовая функция РФА по критерию «Сохранение ПР» при III стадии СК

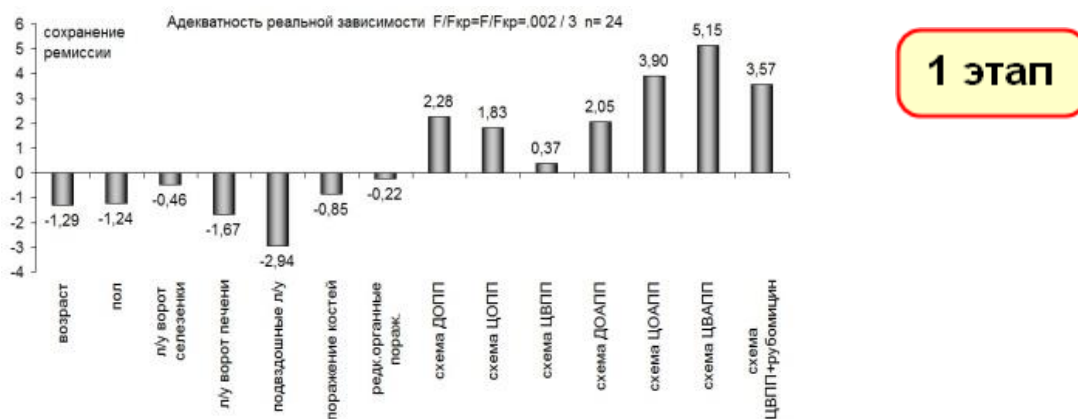


1 этап

Эффективность схем лечения по критерию «Положительная ремиссия» на III стадии уже меньше: значения значимости факторов схем лечения превышают значения

диагностических факторов в 3–7 раз. Также мы видим приоритет значений эффективности схем лечения для III стадии СК.

Весовая функция РФА по критерию «Сохранение ПР» при IV стадии СК



Эффективность схем лечения по критерию «Положительная ремиссия» на IV стадии еще меньше: значения значимости факторов схем лечения превышают значения диагностических факторов в 2–3 раза. Также мы видим приоритет значений эффективности схем лечения для IV стадии СК.

9.7. Пример определения оптимума

Основатель кафедры клинической фармакологии Российского государственного медицинского университета профессор ГОУ ВПО РНИМУ им. И.Н. Пирогова, зав. кафедрой клинической фармакологии лечебного факультета, член-корреспондент РАН Ю.Б. Белоусов

Получение нелинейных весовых функций РФА позволяет определять оптимумы – наиболее эффективную дозу препарата, полученную с учетом всех параметров процесса.



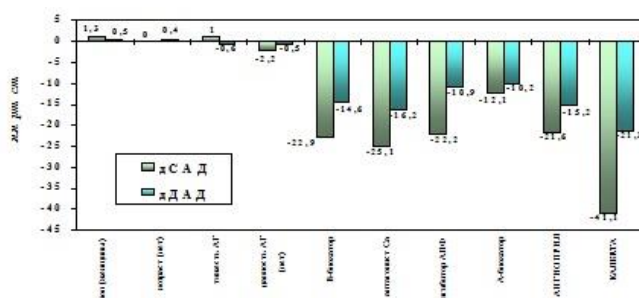
Нелинейные весовые функции n-го порядка позволяют изучить дозозависимости действия гипотензивных препаратов. С помощью построения математической функции зависимости степени снижения АД от дозы препарата можно установить оптимальные эффективные дозы и проследить результат передозировки препарата (см. рисунок 2). Например, для препарата пропранолол наибольшее снижение АД достигается в дозе около 80 мг, дальнейшее увеличение дозы препарата может уменьшать эффект и поэтому нецелесообразно. Максимальная степень снижения АД у данной группы больных мягкой и умеренной тяжести заболевания составляет 15 мм рт. ст.

Общая весовая функция оценки эффективности действия групп препаратов у больных с АГ через 4 недели лечения.

Кафедра клинической фармакологии Российского государственного медицинского университета

(зав. кафедрой – член-корреспондент РАЕН, профессор Ю.Б. Белоусов)

Сравнение гипотензивного действия препаратов (по степени снижения систолического и диастолического АД) у больных с артериальной гипертонией, исключив влияние таких независимых факторов, как пол, возраст больных, тяжесть и давность заболевания, а также изучив значение этих факторов в развитии гипотензивного эффекта препаратов.



9.8. РФА в области научной стоматологии

Центральный Научно-исследовательский институт стоматологии. Хаустова Е.А. Оценка качества реставрации современными композитными материалами (научный руководитель – член-корреспондент РАМН, профессор В.К. Леонтьев). Дисс. канд. мед. наук. – Москва, 1999. – 222 с.

Цель исследования: оценка качества реставраций зубов, выполненных из отечественных и зарубежных пломбировочных материалов светового и химического отверждения, с использованием различных адгезивных систем. Изучение краевой проницаемости (электрометрический метод) на границе «искусственный материал – ткани зуба» с учетом влияния различных факторов.

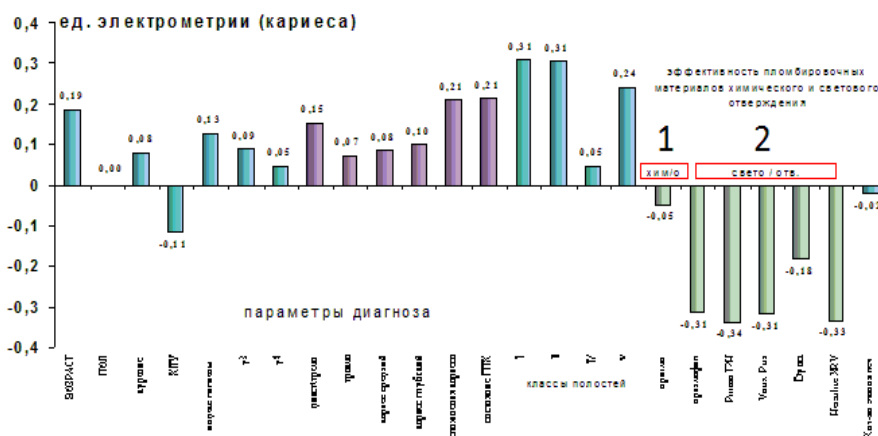
Было обследовано 110 человек в возрасте от 16 до 64 лет. Исследованы характеристики реставрации 374 зубов, в 219 случаях – фронтальных зубов, в 155

случаях – зубов жевательной группы, с интервалами наблюдения: после лечения – 1 нед. – 3 мес. – 6 мес. – 12 мес. – 18 мес. (диаграмма).

Весовая функция влияния клинико-диагностических факторов и пломбировочных материалов на значение электрометрии через 18 месяцев после лечения.

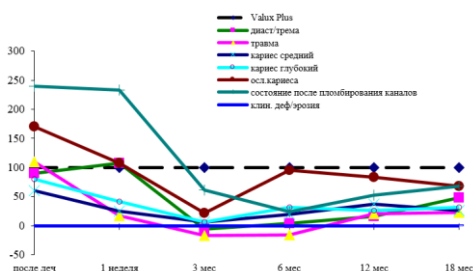
Гистограмма показывает однозначное преимущество светоотверждаемых материалов (2) над материалами химического отверждения (1) в 4–6 раз.

1 этап



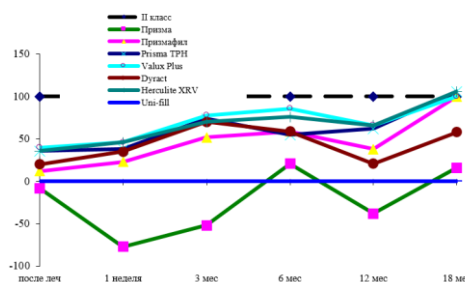
Исследование с использованием инструмента РФА позволило не только получить комплексные оценки значимости факторов внутри этапа наблюдения, но и проследить динамику изменения весовых значимостей факторов по этапам наблюдения:

Динамика значимости пломбировочных материалов по критерию электрометрии относительно пломбировочного материала (Valux)



Исходно преобладающая весовая значимость факторов диагноза (по отношению к значимости фактора пломбировочного материала Valux Plus, принятого за 100%) не устраняется проведенным лечением, а снижается постепенно, достигая минимума к 3-му месяцу обследования, и, далее,

Динамика значимости пломбировочных материалов по критерию электрометрии относительно II класса кариозных полостей



Малая весовая значимость пломбировочных материалов на начальном этапе лечения значительно возрастает к 18-му месяцу наблюдения и становится ведущим по значимости фактором. Эффективность светоотверждаемых материалов превышает

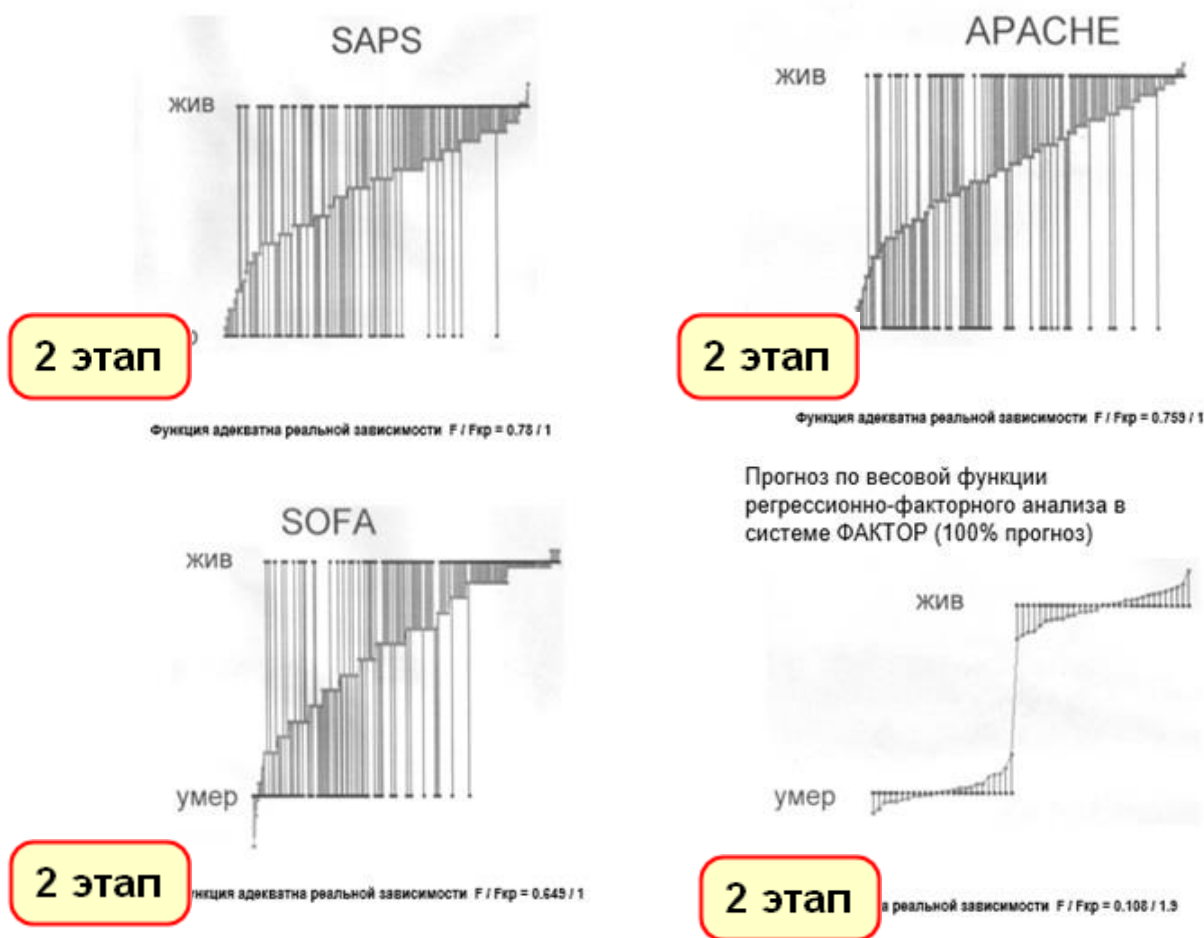
значимость факторов диагноза медленно возрастает вследствие особенностей патологического процесса, но уже не превышая значимости пломбировочных материалов.

эффективность материалов химического отверждения в 3–4 раза на начальном этапе наблюдения и в 8–10 раз через 18 месяцев.

9.7. Сравнение расчётных и опытных значений

РОНЦ РАМН 2007 по материалам базы данных отделения реанимации № 2 на защиту учёной степени д. м. н.

Сравнение расчетов прогноза выживаемости одной и той же группы реанимационных больных по общепринятым международным критериям и по весовой функции регрессионно-факторного анализа в системе ФАКТОР



Регрессионно-факторный анализ с использованием всего стандартного набора качественных и числовых факторов диагноза реанимационных больных позволяет получить 100% весовую функцию прогноза выживаемости с высокой адекватностью реальной зависимости ($0.108 / 1.9$ - 5% от критического значения – против (APACHE) $0.649 / 1$ – 64,9% от критического значения по критерию Фишера) с увеличением точности прогноза в 13 раз.

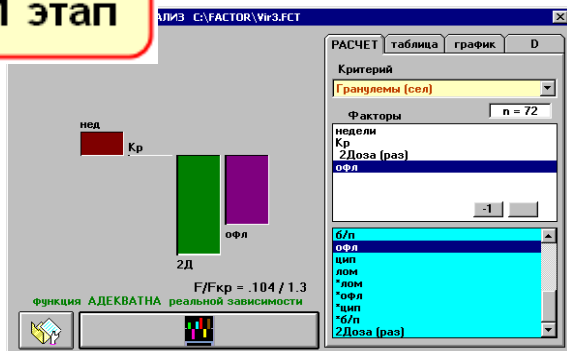
9.10. Работа с нелинейными функциями РФА порядка > 1

Примеры определения оптимальных доз на нелинейных функциях РФА (по результатам версии программы «ФАКТОР-ПРОГНОЗ», 2000 г.).

По материалам исследовательских работ по определению оптимальных доз (ОФЛ, офлоксацин – опыты на мышах) на защиту учёной степени д. м. н. ФГБНУ «ЦНИИТ» – «Центральный научно-исследовательский институт туберкулеза».

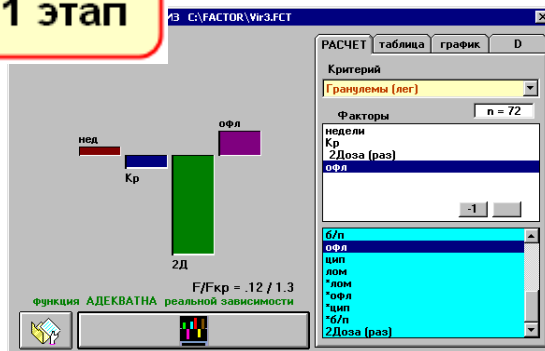
Функция 2-й степени по критерию «Гранулемы в селезенке»

1 этап



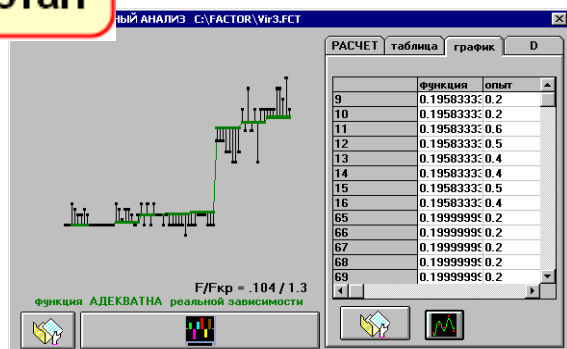
Функция 2-й степени по критерию «Гранулемы в легких»

1 этап



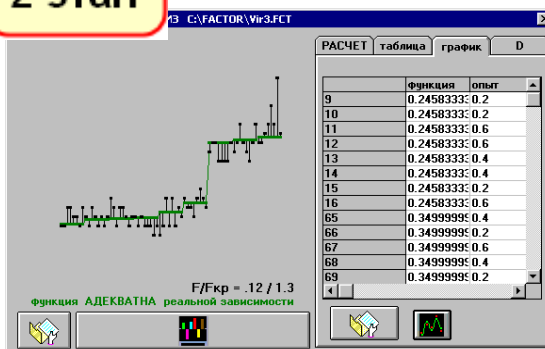
Соответствие расчетного и опытного значений функции 2-й степени по критерию «Гранулемы в селезенке»

2 этап

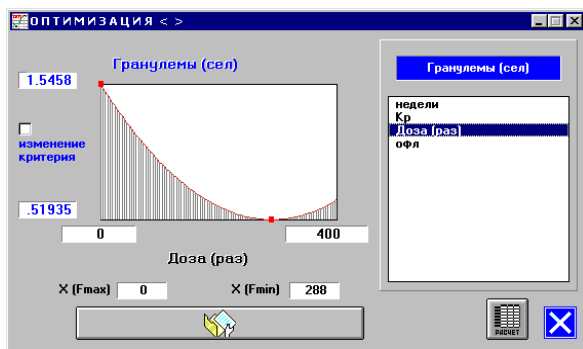


Соответствие расчетного и опытного значений функции 2-й степени по критерию «Гранулемы в легких»

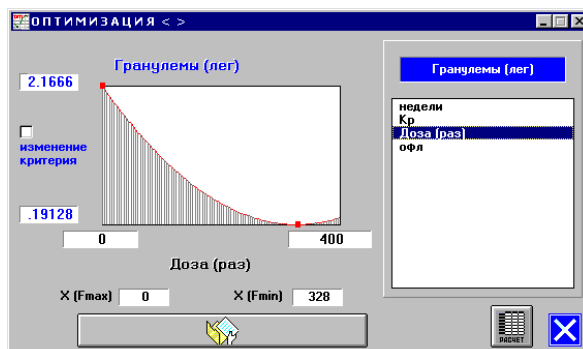
2 этап



Выбрав кнопкой мыши нелинейный параметр – 2 Доза (раз), определяем оптимальную дозу – 288 ед. по минимизации критерия «Гранулемы в селезенке»



Выбрав кнопкой мыши нелинейный параметр – 2 Доза (раз), определяем оптимальную дозу – 320 ед. по минимизации критерия «Гранулемы в легких»



На последних двух картинках – значения оптимальных доз препаратов, при которых достигается минимум гранулём в легких и селезенке при задании параметров доз препаратов в виде нелинейностей 2 степени.

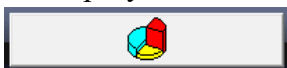
10. Математический аппарат РФА. на примере аналитического программного комплекса «ФАКТОР-ПРОГНОЗ»:

10.1. Программная реализация РФА

Получение линейных весовых функций
(все параметры в 1-й степени)

Операция получения весовой функции проста:

- 1 В верхнем (желтом) поле выбора задается критерий $X_{кр}$ (любой параметр результата).
- 2 В списке параметров анкеты кнопкой мыши набираются параметры (любые – качественные и числовые) X_1, X_2, \dots, X_n в состав весовой функции, значимость которых на результат нас интересует. Расчет осуществляется нажатием кнопки



Решается система уравнений вид $X_{кр} = K_1 X_{1i} + K_2 X_{2i} + \dots + K_n X_{ni}$

и определяются неизвестные коэффициенты системы уравнений – K_1, K_2, \dots, K_n и C (const)

Количество уравнений в системе – количество опытов (строк базы данных) в выборке, значения $X_{1i}, X_{2i}, \dots, X_{ni}$ в каждом уравнении – значения параметров в i -м опыте.

Результат решения визуализируется в виде гистограммы ВЕС (F) со значениями и таблицы

под ярлычком <Расчет>



Определяется адекватность функции реальной зависимости по F критерию Фишера:
 $F / F_{кр} = 0.726 / 1.4 \quad n = 55 \quad (51,84\%)$

51,84% – процент ошибки F РФА от критического значения $F_{кр}$.

Если при добавлении параметра % ошибки уменьшается – параметр оставляется; при увеличении ошибки параметр убирается из функции РФА.

Полученная функция РФА считается адекватной реальной зависимости в случае $F / F_{кр} < 1$

И чем меньше это значение, тем больше приближение к реальной зависимости.

Весовая функция РФА любой степени адекватна реальной зависимости по F-критерию Фишера при обязательном условии $F < F_{\text{критич.}}$ (табл. Фишера)

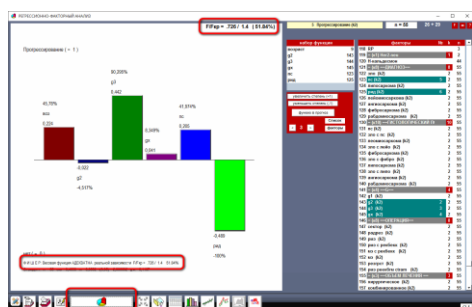
$$F / F_{\text{кр}} = 0.726 / 1.4 \quad n = 55$$

Таблица вкладов параметров, весовых коэффициентов и диапазонов изменения параметров

Весовые вклады – среднее значение эффективности.

Весовые коэффициенты – коэффициенты для прогноза.

Диапазоны изменения параметров – это ограничения при прогнозе



ВНИМАНИЕ! Расчет функции не будет производиться в следующих случаях:

- количество опытов меньше количества набранных в функцию параметров;
- в числовом или качественном параметре присутствует только одно значение;
- набраны все качественные параметры, составляющие полную группу событий (из полной группы событий один (любой) параметр должен исключаться, иначе система не будет иметь решения – этот параметр будет инверсией остальных событий группы);
- в функцию набраны равнозначные (идентичные) по содержанию параметры, что соответствует повторению столбцов в матрице параметров. Оставляется любой из равнозначных параметров с более широкой трактовкой его физического смысла.

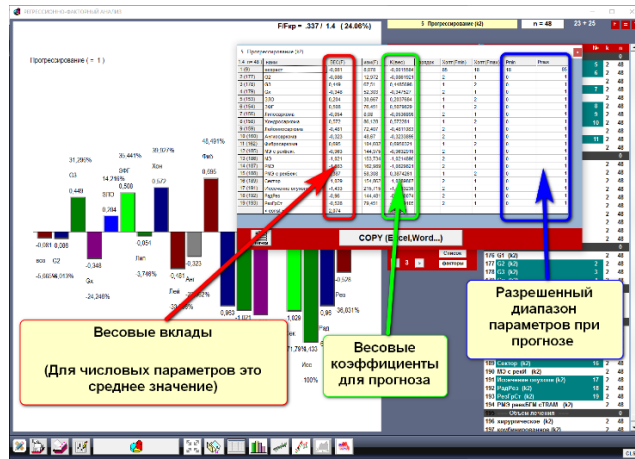
K (вес) – найденные весовые коэффициенты системы уравнений.

ВЕС (F) – гистограмма – характеристика весового вклада параметра в значение критерия. Если необходимо прямое прогнозирование результата (значения критерия) в конкретном случае, то используются весовые коэффициенты РФА – **K (вес)**. Сумма произведений конкретных значений параметров на соответствующие весовые коэффициенты и свободного члена функции РФА определяет ожидаемое значение результата (критерия).

Если необходимо оценить степень влияния параметров на результат (критерий), то используется характеристика **ВЕС (F)**:

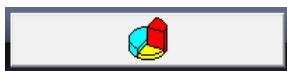
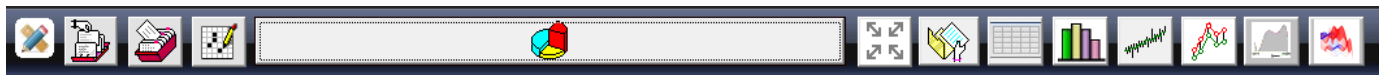
ВЕС (F) – характеристика вклада параметра, осредненная по всем опытам в выборке.

Для качественного параметра ВЕС (F) = K (вес), для числового параметра $ВЕС(F) = \frac{\sum_{i=1}^N K_j \cdot X_i}{N}$



ВНИМАНИЕ! В таблице расчётов синей рамкой выделены столбцы диапазонов допустимых минимальных и максимальных значений параметров – эти значения нужно обязательно указывать в инструкции к прогнозу. При выходе параметров за пределы диапазонов прогноз нужно считать условным.

Меню РФА



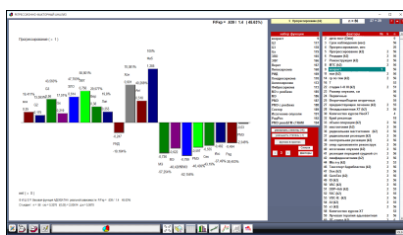
1. Функция РФА



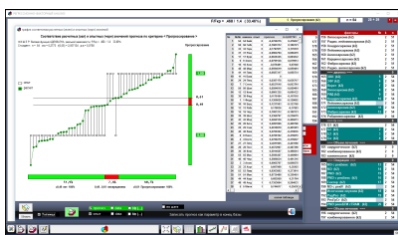
2. Соответствие расчетных и опытных значений



3. Соответствие прогнозируемых значений по «Схеме 1»



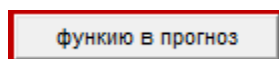
[VIDEO](#)



[VIDEO](#)



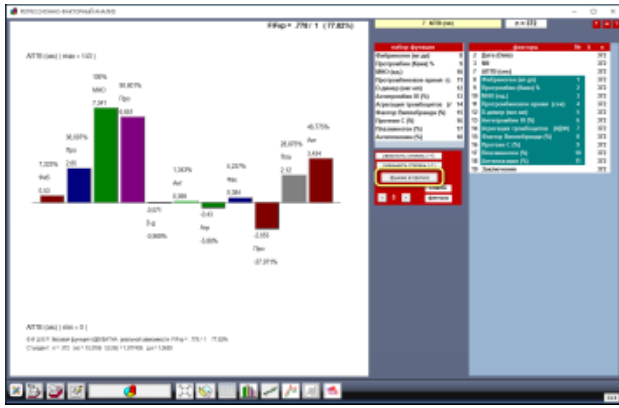
[VIDEO](#)



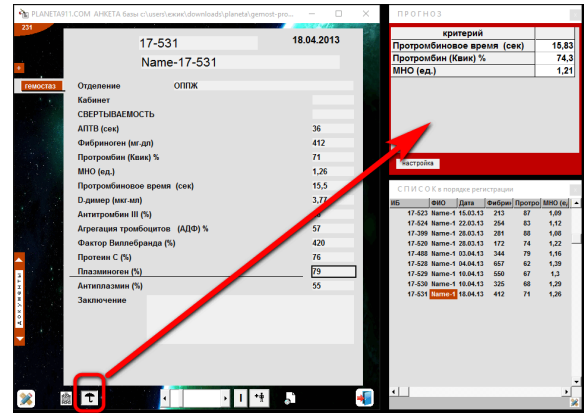
4. Функцию РФА параметров гемостаза в прогноз анкеты



5. Автопрогноз критериев гемостаза в анкете регистрации новых данных



VIDEO



VIDEO

Создание нелинейных функций РФА n-го порядка

увеличить степень (+)

уменьшить степень (-)



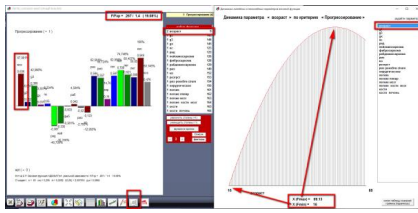
2D функция РФА 2-го порядка по параметру «Возраст»



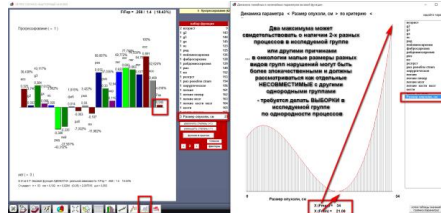
2D Функция РФА 3-го порядка по параметру «Размер объекта»



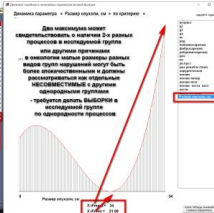
3D (3-х мерная) функция РФА 3-го порядка в 3D



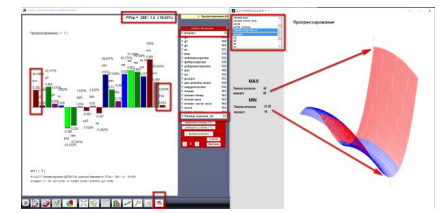
VIDEO



VIDEO



VIDEO



VIDEO

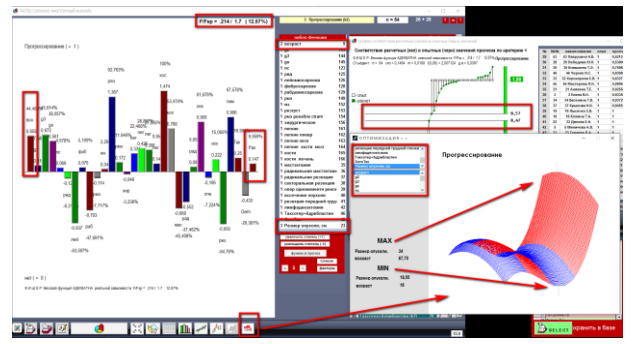
Нелинейная функция РФА 3-го порядка с большим количеством параметров уже не имеет зоны неопределенности.



Функция РФА 3-степени не имеет зон неопределенности – 100-процентное разграничение по прогнозу



3D (3-х мерная) функция РФА 3-го порядка в размерности 3D



10.2. Правила работы в РФА

Единая по всей информации весовая функция может быть получена по всему качественно-числовому набору параметров без ограничений, но при соблюдении следующих правил.

При получении весовых функций РФА числовые и качественные параметры оставляются в составе набранной расчётной группы только при условии, если их добавление улучшает адекватность функции реальной зависимости – ошибка адекватности (% ошибки по отношению к критическому значению) уменьшается. В противном случае параметр убирается из расчётной группы повторным щелчком мыши по наименованию параметра в правом списке анкеты базы данных.

⌋ – количество набранных параметров не должно превышать количество опытов (n) в выборке.

Количество опытов в выборке можно проконтролировать в поле **n = 45**;

Опыты, в которых отсутствуют данные хоть по одному набранному в функцию параметру, автоматически исключаются из расчёта текущей выборки – это можно проследить по уменьшению количества параметров в расчётной группе в поле **n = 45** Чем больше опытов в выборке, тем точнее весовая функция.

⌋ – качественные параметры включаются в функцию в двузначном виде: 0 – нет (фактора, события); 1 – есть (фактор, событие).

Если в вашей базе данных качественные параметры имеют более двух состояний, то для их использования в расчётах РФА воспользуйтесь опциями разложения таких параметров к виду факторов (качественных двузначных состояний типа 1 – НЕТ фактора (состояния); 2 – ЕСТЬ фактор (состояние)) в дополнительные столбцы, добавленные в конце базы данных (см. далее раздел 10.3).

Есть важный момент в наборе в расчёт качественных параметров-факторов: из полной группы событий – один (любой) фактор из полной группы событий не должен набираться в расчёт, так как его добавление сделает матрицу системы вырожденной и она не будет иметь решения – этот не включённый в расчёт параметр будет нулевой точкой отсчёта (на

$Y = 0$), относительно которого будут располагаться те факторы полной группы, которые включены в расчёт.

В исследованиях используйте расчёты РФА по выборкам данных, если позволяет количество опытов исследуемой группы.

Инструмент РФА, особенно в биологических (сильно взаимосвязанных по параметрам) системах, позволяет получать весовые функции высокой точности: если полученная весовая функция не дает высокой точности, то это свидетельствует о присутствии и смещении в исследуемой группе несовместимых по используемому критерию процессов – необходимо при исследованиях разбивать такие совокупности опытов на более однородные (по исследуемым критериям) группы.

10.3. Использование качественных параметров в расчётах РФА

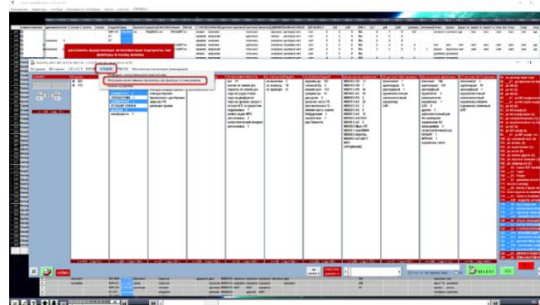
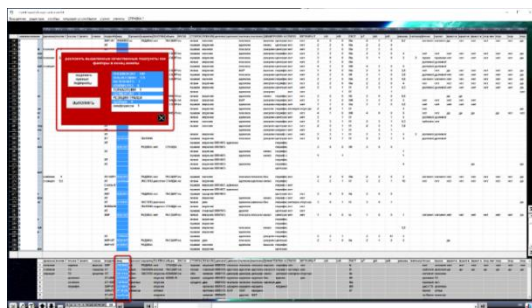
В РФА качественные состояния можно использовать только в двузначном виде: 0 – нет фактора; 1 – есть фактор.

Есть автоматическое преобразование качественных параметров > 2 состояний к виду двузначных факторов в аналитическом научном комплексе «PLANETA» («ФАКТОР») и запись преобразованных качественных параметров в двузначном виде в конец таблицы базы данных.

Разложить качественные пункты с состояниями > 2 как двузначные факторы (0 – «нет фактора, события»; 1 – «есть фактор, событие») можно отдельно для каждого параметра из панели «Таблица данных»: <Опции> – <Разложить как факторы в конец таблицы данных>.

Можно разложить качественные параметры из табличной формы работы с данными в разделе <Главное меню> <Столбцы> <Разложить подпункты как факторы>.

Или можно разложить все качественные параметры > 2 состояний из формы выборки данных.



Разложенный на двузначные факторы качественный параметр "Стадии"

Разложенный на двузначные факторы качественный параметр "Гистологический подтип"

Для служебного пользования МЗ РФ
(матричную часть расчётов в Интернете не размещать)

10.4. Математический аппарат РФА

По материалам методики МАИ, 1986 год, «Применение методов планирования экспериментов в процессе формирования облика объекта», д. т. н., профессор В.Е. Руднев, к. т. н., доцент В.Б. Петров, раздел 2. Методы планирования эксперимента 2.1. Основы РА Москва, 1986 год, Московский ордена Ленина и ордена Октябрьской революции авиационный институт имени Серго Орджоникидзе

Существующие направления в математике для изучения системных процессов

РА (для числовых параметров)

– работа с числовой составляющей любой базы данных. В случае, когда мы не сами выбираем экспериментальные точки $K(1)$, $K(2)$, ... , $K(N)$ и они представляют собой результаты пассивного наблюдения, применяется метод минимизации невязок (классический метод наименьших квадратов).

Факторный анализ (для качественных параметров – факторов)

– работа с качественной составляющей любой базы данных. В факторном анализе принято решать задачу только для трех или четырех ведущих факторов, что также ограничивает возможности и точность решения.

Если можно выбирать экспериментальные точки, то это делается так, чтобы матрица параметров $K \times K$ (K – матрица параметров) была диагональной. Весовые коэффициенты в этом случае будут вычисляться независимо друг от друга, такое планирование называется ортогональным и является основой факторного анализа.

Одним из основных путей получения ортогонального плана является построение полного факторного эксперимента. Главный принцип построения полного факторного эксперимента состоит в том, что каждый фактор варьируется на двух уровнях, то есть может принимать всего два значения, $K(i)_{max}$ и $K(i)_{min}$, при этом количество экспериментальных точек выбирается таким образом, чтобы были рассмотрены все возможные сочетания

уровней факторов. В такой схеме числовые параметры кодируются как факторы через преобразование вида:

$$\tilde{K}_i = \frac{2K_i - K_{i \max} - K_{i \min}}{K_{i \max} - K_{i \min}}$$

В силу известной ограниченности двузначных характеристик, малой чувствительности к взаимосвязям между параметрами, не говоря об отсутствии возможности в реальной практической работе ставить полные факторные эксперименты, имеет смысл ограничиться этой информацией по классической постановке факторного эксперимента. Кроме того, практика показывает, что предполагаемое постоянство весовых коэффициентов, достигнутое по двузначности характеристик, и их независимость друг от друга не определяют корректность решения практических задач.

Основной и фатальный недостаток этих двух методов – невозможность работать со всем качественно числовым набором информации в единой математической схеме. Вследствие этой причины эти методы не нашли широкого применения в науке и практике, поскольку затраты на математику не оправдывают получение однобоких результатов.

Система «ФАКТОР-ПРОГНОЗ» – РФА-инструмент для работы с любыми процессами

В комплексе ПО программы «ФАКТОР-ПРОГНОЗ» разработан новый программный математический механизм РФА (РФА), который позволяет анализировать всю имеющуюся информацию в единой математической схеме без ограничений на состав данных. Этот механизм позволяет работать с любыми процессами в любых вариантах исследований.

Сутью решения РФА (как в регрессионном и факторном анализе) является весовая функция влияния параметров на параметры результата процесса, полученная в результате решения системы уравнений по выборке из таблицы опытных данных. Эта функция имеет аналогичный физический смысл, но является всеобъемлющей и окончательной, так как охватывает весь объем информации (числовой и качественный) в единой математической схеме.

С приемами и методами решения систем уравнений можно ознакомиться в математической литературе в разделах, посвященных РА. В системе «ФАКТОР» такое решение организовано с использованием алгебры матриц и со специальным механизмом согласования качественной и числовой составляющих информации.

Решение, полученное по всей информации (числовой и качественной) в единой схеме, является исчерпывающим для данного набора данных системного процесса и может быть использовано для оптимизации и индивидуального прогнозирования.

10.5. Краткая схема математического аппарата РФА

(приводится для общего представления)

Система «ФАКТОР-ПРОГНОЗ» предназначена для анализа системных процессов с использованием всего имеющегося объема информации (количественного и качественного состава) в единой математической схеме.

Система «ФАКТОР-ПРОГНОЗ» обеспечивает решение практических задач оценки эффективности, точного прогнозирования и оптимального управления по весовым факторным функциям отклика любого порядка.

Использование усовершенствованного механизма РФА позволяет избежать недостатков классического факторного анализа при изучении динамики взаимовлияний параметров (факторов) со сложными взаимосвязями по всему объему информации в единой математической схеме.

РФА – анализ всего объема информации в единой схеме числовых и качественных показателей (факторов) в виде весовых функций n-го порядка.

В качестве основы для математического аппарата получения таких весовых функций представлен классический метод получения весовых функций – метод наименьших квадратов (минимизация невязок) с применением специального масштабирования по качественным и числовым показателям.

Такой подход позволил объединить всю необходимую для анализа информацию (и качественную, и числовую) в единую схему.

Суть метода рассматривается для случая, когда функция отклика ищется в виде полинома первой степени (система «ФАКТОР» позволяет получать и работать с весовыми функциями отклика в виде полиномов любой степени).

$$X = B_0 + B_1K_1 + B_2K_2 + \dots + B_\alpha K_\alpha$$

Функцию отклика называют уравнением регрессии. Результаты опытов используются для нахождения неизвестных коэффициентов:

$B_0, B_1, B_2, \dots, B_\alpha$ – весовые коэффициенты регрессии

Число опытов берется больше числа неизвестных коэффициентов $N \geq \alpha + 1$ или равным ему.

Коэффициенты регрессии должны выбираться таким образом, чтобы величины

$$\xi_i = X^{(i)} - (B_0 + B_1K_1^{(i)} + B_2K_2^{(i)} + \dots + B_\alpha K_\alpha^{(i)}) \quad i = 1, 2, \dots, n$$

были как можно меньше.

□ ξ_i НЕВЯЗКА – разница между экспериментальным и опытным значениями X (критерия).

Требование минимизации невязок записывается следующим образом:

$$\min_{V_0, V_1, \dots, V_\alpha} \sum_{i=1}^N \xi_i^2 = \min_{V_0, V_1, \dots, V_\alpha} \sum_{i=1}^N (X^{(i)} - V_0 + V_1 K_1^{(i)} + V_2 K_2^{(i)} + \dots + V_\alpha K_\alpha^{(i)})^2$$

Минимум функции (если он существует) достигается при равенстве нулю частных производных по всем неизвестным.

Обозначив $U = \sum_{i=1}^N \xi_i^2$, необходимые условия минимума будут выглядеть следующим образом:

После дифференцирования получаем систему нормальных уравнений относительно неизвестных

$$V_0, V_1, V_2, \dots, V_\alpha:$$

$$B_0 N + B_1 \sum_{i=1}^N K_1^{(i)} + B_2 \sum_{i=1}^N K_2^{(i)} + \dots + B_\alpha \sum_{i=1}^N K_\alpha^{(i)} = \sum_{i=1}^N X^{(i)};$$

$$B_0 \sum_{i=1}^N K_1^{(i)} + B_1 \sum_{i=1}^N (K_1^{(i)})^2 + B_2 \sum_{i=1}^N K_1^{(i)} K_2^{(i)} + \dots + B_\alpha \sum_{i=1}^N K_1^{(i)} K_\alpha^{(i)} = \sum_{i=1}^N K_1^{(i)} X^{(i)};$$

.....

$$B_0 \sum_{i=1}^N K_\alpha^{(i)} + B_1 \sum_{i=1}^N K_\alpha^{(i)} K_1^{(i)} + B_2 \sum_{i=1}^N K_\alpha^{(i)} K_2^{(i)} + \dots + B_\alpha \sum_{i=1}^N (K_\alpha^{(i)})^2 = \sum_{i=1}^N K_\alpha^{(i)} X^{(i)}.$$

Определение коэффициентов $B_0, B_1, B_2, \dots, B_\alpha$:

$$B_j = C_{j0} \sum_{i=1}^N X^{(i)} + C_{j1} \sum_{i=1}^N K_1^{(i)} X^{(i)} + C_{j2} \sum_{i=1}^N K_2^{(i)} X^{(i)} + \dots + C_{j\alpha} \sum_{i=1}^N K_\alpha^{(i)} X^{(i)}$$

$$j = 1, 2, \dots, \alpha$$

где

$$\left\| \begin{array}{cccc} C_{00} & C_{01} & \dots & C_{0\alpha} \\ C_{10} & C_{11} & \dots & C_{1\alpha} \\ \dots & \dots & \dots & \dots \\ C_{\alpha 0} & C_{\alpha 1} & \dots & C_{\alpha\alpha} \end{array} \right\| = (K^* K)^{-1}$$

K – матрица значений факторов
(матрица планирования)

K^* – транспонированная
матрица K

$$K = \left\| \begin{array}{cccc} 1 & K_1^{(1)} & K_2^{(1)} & \dots & K_\alpha^{(1)} \\ 1 & K_1^{(2)} & K_2^{(2)} & \dots & K_\alpha^{(2)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & K_1^{(N)} & K_2^{(N)} & \dots & K_\alpha^{(N)} \end{array} \right\|$$

Операция обращения матрицы $(K^*K) - A$ осуществляется следующим образом:

- определяется детерминант (определитель) матрицы $|A|$;
- каждый элемент матрицы A заменяется его алгебраическим дополнением A_{jk} ;
- полученная матрица транспонируется.

В результате получается присоединенная матрица: $A \square$

После деления каждого элемента присоединенной матрицы на детерминант получаем матрицу, обратную матрице A :

$$A^{-1} = \frac{A \nabla}{|A|} = \left[\frac{A_{ki}}{|A|} \right]$$

Проверка функции на адекватность реальной зависимости по F-критерию Фишера

По F-критерию Фишера проверяется адекватность весовой функции РФА реальной зависимости. В найденную функцию РФА подставляются опытные значения параметров, по которым она была рассчитана, и сравниваются дисперсии отклонения рассчитанных по функции и опытных значений результирующего параметра, принятого в качестве критерия, в виде соотношения $F / F_{кр}$

В качестве оценки величины отклонения опытных данных от прогнозируемых значений, полученных в функции РФА, используется расчёт остаточной дисперсии (дисперсия адекватности):

$$D_{ад} = \frac{\sum_{i=1}^N (\tilde{X}_i - X_i)^2}{f}; \quad D_B = \frac{\sum_{i=1}^N (\bar{X} - X_i)^2}{N - 1}, \quad \text{где } \bar{X} = \frac{\sum_{i=1}^N X_i}{N}; \quad F = \frac{D_{ад}}{D_B}$$

$D_{ад}$ – дисперсия адекватности, для $D_{ад}$ $f = N - (\square + 1)$,

где f – число степеней свободы;

\square – количество факторов;

D_B – дисперсия воспроизводимости (характеризует разброс значений по опытам);

X_i – значение критерия, вычисленное по функции РФА для значений факторов в i -м опыте;

\tilde{X}_i – значение критерия непосредственно в i -м опыте;

N – число опытов.

После нахождения значения критерия F его сравнивают с критическим значением, найденным по таблице с учетом числа степеней свободы. При соблюдении условия $F < F_{\text{критическое}}$ (то есть ошибка меньше критического значения) функция считается адекватной реальной зависимости.

11. РА и планирование экспериментов с использованием ПО «ФАКТОР-ПРОГНОЗ»

Классические математические инструменты регрессионного (РА) и факторного (ФА) анализа в области медицины и особенно онкологии лишены главного – результата, полученного по всему объёму информации (и качественного, и числового) в единой математической схеме; это приводит к однобоким, некорректным оценкам эффективности реального влияния параметров на результаты процесса и не позволяет получать исчерпывающие точные функции прогноза критериев, но позволяет делать сравнительные оценки параметров процесса.

Математический инструмент РФА (в отличие от РА и ФА) работает над всей информацией (и числовой, и качественной) в единой математической схеме РА и в ПО «ФАКТОР-ПРОГНОЗ» имеет на порядок большие возможности по оценке эффективности, оптимизации на нелинейных функциях РФА порядка > 1 и позволяет получать высокоточные системы прогноза на всём имеющемся объёме данных (числовых и качественных) в единой математической схеме.

Совместный материал представлен сотрудниками НМИЦ им. Н.Н. Блохина: Колосов Е.А. и сотрудники кафедры 703 МАИ, национальный исследовательский университет:
к. т. н. Петров В.Б., Киреев А.А.

11.1. РА

(Е.А. Колосов, комментарий: инструменты анализа для любых процессов математически работают одинаково, но процессы в разных областях отличаются по природе, функционалу и сложности. Например, в биологических процессах все параметры жестко завязаны между собой в состоянии собственного биологического равновесия, что приводит к отсутствию идентичности значений параметров между опытами (пациентами), но позволяет получать весовые функции высокой точности. В технике нет жесткого состояния взаимного равновесия всех параметров, но можно наблюдать состояние идентичности опытов и групп при исследованиях технических процессов. Такие различия в природе процессов приводят к различиям в идеологии использования доказательного однопараметрического и многопараметрического анализа по оценке эффективности параметров в процессе.)

Процесс формирования оптимального облика сложной технической системы (СТС) с использованием непосредственно в процессе оптимизации имитационных или полунатурных моделей, которые воспроизводят функционирование СТС, может оказаться нереализуемым из-за того, что просчет на таких моделях одного варианта занимает по отношению к общему сроку разработки слишком много времени. Поэтому бывает трудно просмотреть необходимое число вариантов для принятия оптимального решения.

Для устранения этого недостатка используют РА. РА – процедура построения уравнения регрессии (регрессионной модели) методом наименьших квадратов (МНК) и анализ полученного уравнения с помощью аппарата математической статистики.

Берется некоторое ограниченное число вариантов облика СТС (несколько точек $k^{(1)}, k^{(2)}, \dots, k^{(N)}$ в многомерном пространстве конструктивных параметров), и для каждого варианта (каждой точки) на сложных имитационных моделях определяется её эффективность (определяются соответствующие точки $w^{(1)}, w^{(2)}, \dots, w^{(N)}$ в пространстве параметров эффективности). Указанную процедуру можно трактовать как проведение на имитационных моделях ряда опытов (экспериментов). Результаты экспериментов затем обрабатываются для получения функций отклика (уравнений регрессии):

$$\begin{aligned} \bar{w}_1 &= \varphi_1(k_1, k_2, \dots, k_\alpha); \\ \bar{w}_2 &= \varphi_2(k_1, k_2, \dots, k_\alpha); \\ &\dots\dots\dots \\ \bar{w}_n &= \varphi_n(k_1, k_2, \dots, k_\alpha). \end{aligned}$$

Функции отклика представляют собой достаточно простые аналитические выражения; как правило, это полиномы какой-либо степени. Простота функций отклика позволяет использовать их в дальнейшем при выделении доминирующих оптимальных решений.

РА включает в себя операции оценивания неизвестных параметров функции отклика, проверки их статистической значимости и проверки адекватности полученной функции отклика.

Функции отклика должны быть определены в таком виде, при котором не нарушалась бы их адекватность реальным зависимостям $w=f(k)$. В противном случае могут быть отброшены лучшие решения. Для того чтобы этого не произошло, применяются методы, позволяющие правильно спланировать эксперименты и проверить адекватность полученной по их результатам функции отклика.

Планирование эксперимента – это процедура выбора числа и условий проведения опытов, необходимых и достаточных для получения функций отклика с требуемой точностью. Под числом опытов N понимается число вариантов (точек), прогоняемых через сложные имитационные модели, а под условиями проведения – конкретные значения конструктивных параметров в этих точках.

В общем случае величина w является случайной, так как она определяется на основании обработки результатов некоторой выборки. Поэтому следует иметь в виду, что аналитическая зависимость $\bar{w}=\varphi(k)$ отражает связь математического ожидания случайной величины w с неслучайными параметрами $k_1, k_2, \dots, k_\alpha$. Случайность величины w связана с тем, что имитация поведения сложной динамической системы происходит под воздействием большего числа случайных внешних процессов, в том числе под влиянием человеческого фактора. Определение зависимости $\bar{w}=\varphi(k)$ – это задача регрессионного анализа (\bar{w} обозначает математическое ожидание величины w).

При линейной функции отклика (регрессионная модель ищется в виде полинома первой степени):

$\bar{w} = b_0 + b_1 k_1 + b_2 k_2 + \dots + b_\alpha k_\alpha$, где $b_0, b_1, \dots, b_\alpha$ - коэффициенты регрессии.

Система уравнений для нахождения неизвестных коэффициентов регрессии в матричной форме имеет следующий вид:

$$B = [K^* K]^{-1} [K^* \bar{W}],$$

где $K = \begin{bmatrix} 1 & k_1^{(1)} & k_2^{(1)} & \dots & k_\alpha^{(1)} \\ \vdots & & \ddots & & \vdots \\ 1 & k_1^{(N)} & k_2^{(N)} & \dots & k_\alpha^{(N)} \end{bmatrix}$ - матрица значений факторов (матрица планирования);

$\bar{W} = \begin{bmatrix} \bar{w}^{(1)} \\ \vdots \\ \bar{w}^{(N)} \end{bmatrix}$ - матрица-столбец наблюдений;

$B = \begin{bmatrix} b_0 \\ \vdots \\ b_\alpha \end{bmatrix}$ - матрица-столбец неизвестных коэффициентов регрессии;

K^* - транспонированная матрица K (матрица, в которой по отношению к исходной строки и столбцы меняются местами);

$K^* K$ - матрица коэффициентов нормальных уравнений для нахождения неизвестных коэффициентов регрессии;

$$[K^* K]^{-1} = \begin{bmatrix} c_{00} & c_{01} & \dots & c_{0\alpha} \\ \vdots & \ddots & & \vdots \\ c_{\alpha 0} & c_{\alpha 1} & \dots & c_{\alpha\alpha} \end{bmatrix} - \text{матрица, обратная матрице}$$

$K^* K$ ($[K^* K]^{-1} [K^* K] = I$ - единичная матрица).

В развернутой форме соотношения для нахождения неизвестных коэффициентов регрессии имеют следующий вид:

$$b_j = \sum_{v=0}^{\alpha} (c_{jv} \sum_{i=1}^N k_v^{(i)} \bar{w}^{(i)}), \quad j=0,1,2,\dots,\alpha, \quad (*)$$

при этом $k_0^{(j)} = 1, \forall j$.

Матрица коэффициентов нормальных уравнений $K^* K$ не должна быть вырожденной (определитель матрицы не должен быть равен нулю), а это требование налагает дополнительное условие на выбор экспериментальных точек, в матрице планирования K элементы одного столбца не должны быть взвешенными суммами соответствующих элементов других столбцов.

Анализ соотношения (*) показывает, что если мы захотим изменить вид функции отклика, например отбросить часть членов в уравнении регрессии, то придется все коэффициенты регрессии перед оставшимися членами пересчитывать заново, поскольку изменятся элементы в матрице $[K^* K]^{-1}$, и, следовательно, изменится число слагаемых и в выражении (*). Другими словами, непосредственное использование МНК не позволяет определять коэффициенты регрессии независимо друг от друга. Поэтому этот метод в

основном используется в том случае, когда мы не сами выбираем экспериментальные точки $k^{(1)}, k^{(2)}, \dots, k^{(N)}$, а они представляют собой результаты пассивного наблюдения за объектом (пассивного эксперимента) – в частности, в медицине.

11.2. Планирование экспериментов

Если же можно самим выбирать экспериментальные точки, то делать это надо так, чтобы матрица K^*K была диагональной (все члены матрицы, за исключением лежащих на главной диагонали, должны быть равны нулю). В этом случае соотношения для определения коэффициентов регрессии будут иметь следующий вид:

$$b_j = \frac{\sum_{i=1}^N k_j^{(i)} \bar{w}^{(i)}}{\sum_{i=1}^N (k_j^{(i)})^2}, \quad j=0,1,2,\dots, \alpha,$$

а это значит, что эти коэффициенты будут вычисляться независимо друг от друга. Увеличение или уменьшение числа членов в функции отклика (вычеркивание или добавление строк и столбцов в матрице K^*K) не будет изменять значения уже вычисленных коэффициентов регрессии, и их не потребуется пересчитывать. Это очень важно, поскольку вид функции отклика приходится изменять, когда она не удовлетворяет условиям адекватности. Такое планирование называется ортогональным.

Одним из основных путей получения ортогонального плана является построение полного фактора эксперимента (ПФЭ), состоящего в том, что каждый фактор (каждый конструктивный параметр) k_j варьируется на двух или более уровнях. При этом число экспериментальных точек $k^{(1)}, k^{(2)}, \dots, k^{(N)}$ выбирается таким образом, чтобы были рассмотрены все возможные сочетания уровней факторов. Например, необходимое число точек (число опытов) для реализации двухуровневого ПФЭ $N=2^\alpha$, где α – число факторов.

При выборе уровней варьирования для каждого фактора необходимо учитывать следующие два основных противоречивых момента:

– с одной стороны, стремление учесть возможный диапазон изменения параметра k_j приводит к увеличению интервала варьирования: $\Delta k_j = k_{j \max} - k_{j \min}$

($k_{j \max}$ – верхний уровень, $k_{j \min}$ – нижний уровень);

– с другой стороны, чем меньше интервал варьирования факторов, тем меньше степень полинома, выражающего с требуемой точностью функцию отклика, и тем проще процедура её определения.

Поэтому назначать уровни варьирования факторов необходимо на основе тщательного анализа априорной информации о поведении функции отклика.

Для упрощения записи условий эксперимента значения факторов кодируются таким образом, чтобы верхний уровень соответствовал +1, а нижний -1. Кодирование производится по следующей формуле:

$$\tilde{k}_j = \frac{2k_j - k_{j\max} - k_{j\min}}{k_{j\max} - k_{j\min}}, \text{ где } \tilde{k}_j - \text{ кодированное значение фактора}$$

(в дальнейшем знак « \sim » будем опускать).

Рассмотрим правило построения матрицы планирования для двухуровневого ПФЭ в случае линейной функции отклика.

Число строк в матрице равно числу опытов, а число столбцов равно числу факторов плюс один (первый) столбец, соответствующий свободному члену b_0 в функции отклика. Все элементы первого столбца всегда равны +1.

Каждому столбцу в матрице присваивается буквенное значение соответствующего ему фактора. Элемент какого-либо столбца (за исключением первого), находящийся на пересечении с i -строкой, обозначает уровень, на котором находится соответствующий данному столбцу фактор в i -й экспериментальной точке. Обычно для простоты записи единицы в матрице планирования опускают.

При $\alpha = 3$ число опытов $N = 2^3 = 8$, и матрица планирования имеет вид:

$$K = \begin{array}{c|cccc} & & k_1 & k_2 & k_3 \\ \hline & + & - & - & - \\ + & + & - & - & \\ + & + & - & - & \\ + & - & + & - & \\ + & + & + & - & \\ + & - & - & + & \\ + & + & - & + & \\ + & - & + & + & \\ + & + & + & + & \end{array}$$

Нетрудно увидеть, что матрицу планирования для трех факторов получают из матрицы планирования для двух факторов, повторив её дважды: один раз – при значении k_3 , находящемся на нижнем уровне, второй раз – на верхнем уровне. Аналогичным образом могут быть построены матрицы планирования для любого числа факторов.

Матрицы планирования в ПФЭ обладают следующими свойствами:

– свойство симметричности

$$\sum_{i=1}^N k_j^{(i)} = 0, \quad j=1, 2, \dots, \alpha;$$

– свойство нормировки

$$\sum_{i=1}^N \left(k_j^{(i)}\right)^2 = N, \quad j=0, 1, 2, \dots, \alpha;$$

– свойство ортогональности

$$\sum_{i=1}^N k_j^{(i)} k_u^{(i)} = 0, \quad j \neq u; \quad u, j = 0, 1, 2, \dots, \alpha.$$

Из свойства ортогональности следует, что матрица коэффициентов нормальных уравнений K^*K будет диагональной.

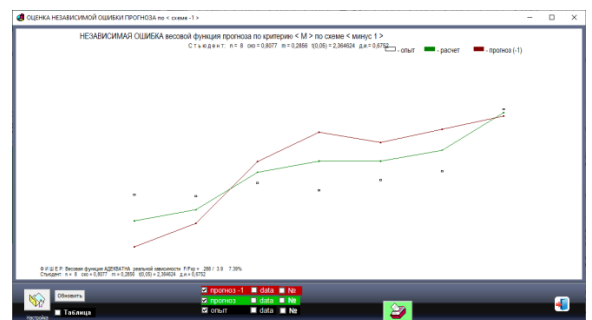
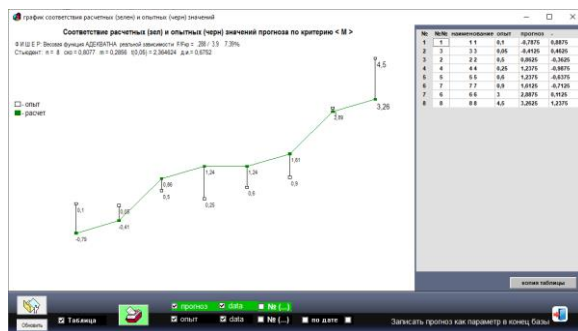
Из свойства нормировки следует, что все диагональные элементы этой матрицы будут равны N , а коэффициенты регрессии – вычисляться по следующим простым формулам:

$$b_j = \frac{\sum_{i=1}^N k_j^{(i)} \bar{w}^{(i)}}{N}, \quad j = 0, 1, 2, \dots, \alpha.$$

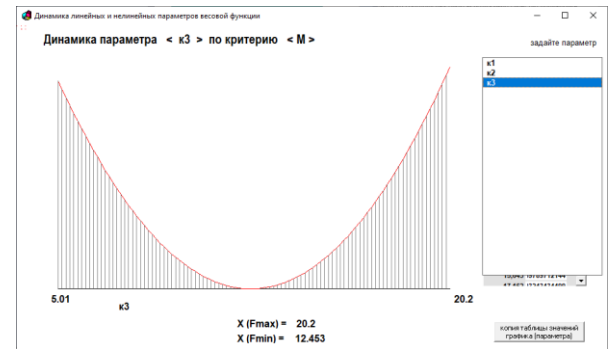
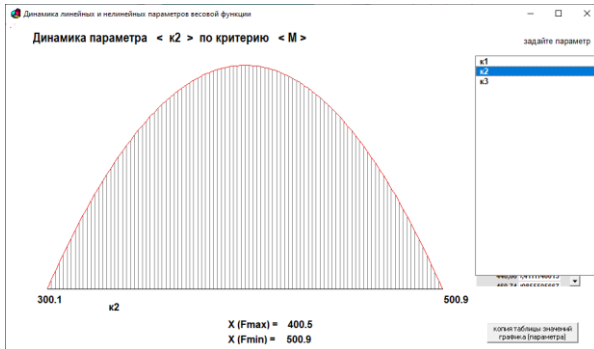
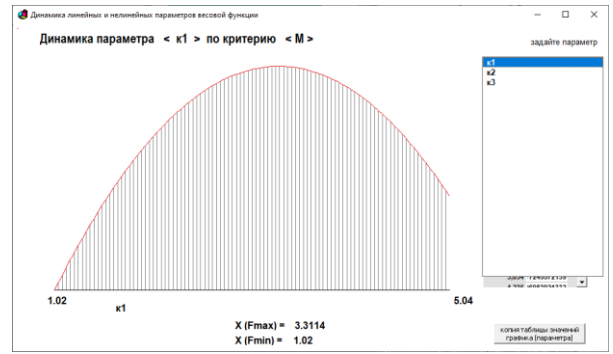
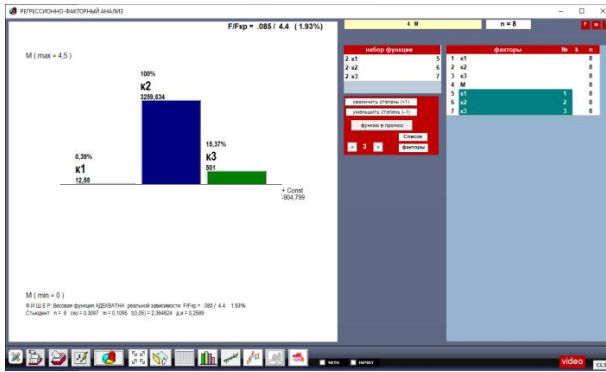
Получение функций РА и РФА в аналитическом программном комплексе «ФАКТОР-ПРОГНОЗ»



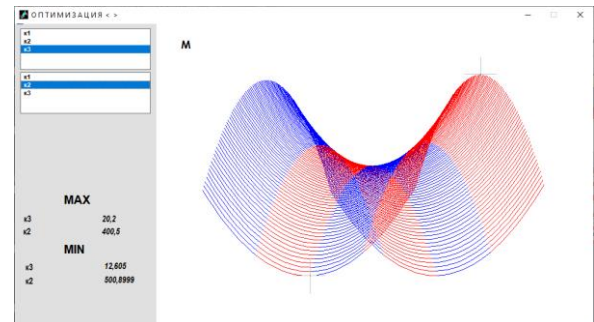
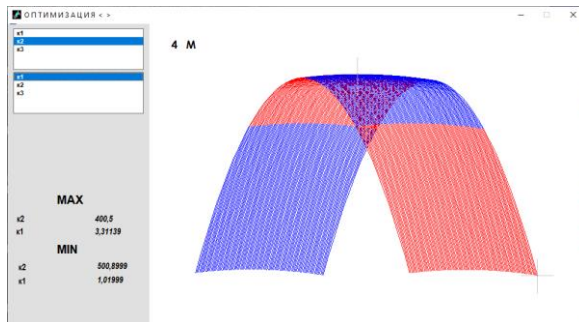
Графическая интерпретация соответствия расчётных и опытных значений – проверка весовых функций РА и РФА на адекватность реальной зависимости



2D оптимизация на функциях РФА в ПО «ФАКТОР-ПРОГНОЗ» порядка > 1



3D оптимизация на функциях РФА в ПО «ФАКТОР-ПРОГНОЗ» порядка > 1



Найденные коэффициенты b_1, \dots, b_α характеризуют силу влияния факторов на показатель эффективности w , а вклад фактора k_j в величину показателя w при переходе фактора с нижнего уровня на верхний называется эффектом фактора и равен удвоенному коэффициенту регрессии $2b_j$.

Планируя эксперимент, на первом этапе всегда стремятся получить линейную функцию отклика. Однако нет гарантии, что в выбранных интервалах варьирования факторов линейная функция с требуемой точностью сможет заменить реальную зависимость $w=f(k)$.

Если линейная функция непригодна (неадекватна реальной зависимости) и надо переходить к нелинейной функции, то возникает вопрос, как количественно оценить нелинейность, пользуясь ПФЭ. Наиболее часто встречающийся вид нелинейности обусловлен тем, что эффект одного фактора зависит от уровня, на котором находится другой фактор. Это так называемый эффект взаимодействия двух факторов.

Учёт эффекта взаимодействия μ -го и ν -го факторов заключается в том, что функцию отклика ищут в виде:

$$\bar{w} = b_0 + b_1 k_1 + b_2 k_2 + \dots + b_\alpha k_\alpha + b_{\mu\nu} k_\mu k_\nu.$$

Искомый эффект взаимодействия определяется удвоенным коэффициентом $b_{\mu\nu}$, для нахождения которого в матрице планирования добавляется ещё один столбец, равный произведению столбцов для факторов k_μ и k_ν . Полученная матрица сохранит полезные свойства ПФЭ (симметричность, свойство нормировки, ортогональность), поэтому коэффициент $b_{\mu\nu}$ можно вычислять при помощи известных соотношений, используя при этом новый столбец, то есть:

$$b_{\mu\nu} = \frac{\sum_{i=1}^N k_\mu^{(i)} k_\nu^{(i)} \bar{w}^{(i)}}{N}.$$

Как уже отмечалось, ортогональность матрицы планирования позволяет получить независимые друг от друга оценки коэффициентов регрессии. Однако это будет справедливо лишь в том случае, если функцию отклика искать в виде, содержащем линейные члены и члены, учитывающие эффекты взаимодействия факторов. Если функцию отклика искать в виде, содержащем квадратичные члены $b_{jj} k_j^2$, то при помощи ПФЭ нельзя найти коэффициенты b_{ii} , так как при добавлении в матрицу планирования столбцов, соответствующих квадратичным членам, нарушается свойство ортогональности из-за совпадения этих столбцов между собой и с первым столбцом.

После отыскания коэффициентов регрессии производится проверка их статической значимости. Статически незначимые коэффициенты не несут полезной информации об объекте и исключаются из уравнения регрессии. Некоторый коэффициент b_j является оценкой соответствующего коэффициента β_j теоретической модели, причем $M[b_j] = \beta_j$. Кроме того, каждый коэффициент представляет собой некоторую взвешенную сумму случайных нормально распределенных значений отклика $\bar{w}^{(i)}$ с весовыми коэффициентами $1/N$. Поэтому для установления значимости коэффициента используется t -критерий Стьюдента. При этом дисперсия коэффициента регрессии может быть найдена по следующей формуле:

$\sigma_{\beta_j}^2 = \frac{\sum_{i=1}^N (k_j^{(i)})^2 D_B}{N^2} = \frac{D_B}{N}$ (при выводе данного соотношения было использовано свойство нормировки матрицы планирования),

где $D_B = \frac{\sum_{i=1}^N \sum_{u=1}^m (\bar{w}^{(i)} - w_u^{(i)})^2}{N(m-1)}$ – дисперсия воспроизводимости, которая находится постановкой так называемых параллельных опытов, то есть опытов, повторенных в одинаковых условиях; m – число параллельных опытов (полагается, что в каждой экспериментальной точке проводится одинаковое число параллельных опытов):

$$\bar{w}^{(i)} = \frac{\sum_{u=1}^m w_u^{(i)}}{m} \text{ – среднее значение выходного параметра в } i\text{-й экспериментальной точке.}$$

Указанные соотношения применяются, если в каждой точке проведено $m > 1$ параллельных опытов.

Если $m = 1$, то есть при каждом наборе факторов проводится один опыт, дисперсию воспроизводимости можно определить по ряду параллельных опытов в другой точке

области эксперимента, например в центре плана (в точке с координатами $k_1 = k_2 = k_3, \dots, k_\alpha = 0$).

После проведения m_0 опытов дисперсию воспроизводимости находят по следующей формуле:

$$D_B = \frac{\sum_{u=1}^{m_0} (\bar{w}^{(0)} - w_u^{(0)})^2}{m_0 - 1},$$

где $\bar{w}^{(0)} = \frac{\sum_{u=1}^{m_0} w_u^{(0)}}{m_0}$ – среднее значение выходного параметра в центре плана.

Находя значимость коэффициента b_j , проверяют основную статистическую гипотезу:

$$M[b_j] = \beta_j \neq 0 \text{ против альтернативы } M[b_j] = \beta_j = 0.$$

Значение t -критерия вычисляется по формуле: $t_j = \frac{|\beta_j - b_j|}{\sigma_{\beta_j}} = \frac{|b_j| \sqrt{N}}{\sqrt{D_B}}$.

Число степени свободы статистики t_j равно числу степеней свободы дисперсии воспроизводимости (число степеней свободы – это разность между числом слагаемых в сумме квадратов и числом связей, наложенных на эти слагаемые) и зависит от способа определения этой дисперсии.

При определении дисперсии при $m > 1$ число степеней свободы $f_B = f_t = N(m-1)$, а при $m = 1$ по формуле $f_t = m_0 - 1$.

Из таблиц распределения Стьюдента по уровню значимости $P_{\text{ош}}$ и числу степеней свободы f_t определяется критическое значение величины $t_{j\text{кр}}$, которую статистика t_j не должна превысить.

Если это обеспечивается, то есть $t_j < t_{j\text{кр}}$, то с доверительной вероятностью $1 - P_{\text{ош}}$ удовлетворяется гипотеза $M[b_j] = \beta_j = 0$.

На практике для сокращения вычисления рассчитываются не все значения t_j , а одно значение $b_{\text{кр}} = t_{\text{кр}} \sqrt{\frac{D_B}{N}}$. Все коэффициенты уравнения регрессии, удовлетворяющие условию $b_j \leq b_{\text{кр}}$, являются статистически незначимыми, и их следует положить равными нулю.

После нахождения коэффициентов регрессии функции отклика и проверки их статистической значимости необходимо проверить, насколько хорошо полученная функция отклика $\bar{w} = \varphi(k)$ заменяет реальную существующую зависимость $w = f(k)$, то есть проверить адекватность функции отклика.

В основе этой проверки лежит сравнение двух видов рассеивания – рассеивания случайной величины w относительно своего среднего значения и рассеивания (разброса) этих средних значений относительно предсказанной по уровню регрессии поверхности отклика.

Так как рассеивание характеризуется дисперсией, то сравнение рассеиваний заменяется сравнением соответствующих дисперсий (точнее – их оценок) и базируется на аппарате дисперсионного анализа, сущность которого состоит в том, что общая сумма квадратов отклонений измеренных значений от их предсказанных значений разбивается на две независимые суммы (два слагаемых).

Первое слагаемое зависит от рассеивания значений $w_u^{(i)}$ относительно своих средних $\bar{w}^{(i)}$ и характеризует оценку дисперсии воспроизводимости D_B .

Второе слагаемое зависит от разброса средних $\bar{w}^{(i)}$ относительно модели $\varphi(k^{(i)})$ и характеризует оценку так называемой дисперсии адекватности

$$D_{ад} = \frac{\sum_{i=1}^N [\bar{w}^{(i)} - \varphi(k^{(i)})]^2}{f_{ад}},$$

где $f_{ад} = N - l$ – число степеней свободы оценки дисперсии адекватности равно разности между числом слагаемых в сумме квадратов и числом наложенных связей (число связей равно числу коэффициентов регрессии в найденной функции отклика).

Найденная модель $\bar{w} = \varphi(k)$ будет адекватной в том случае, если дисперсия адекватности не будет превышать дисперсию воспроизводимости. Для того чтобы эти дисперсии были равны между собой, их оценки D_B и $D_{ад}$ должны быть однородны. В статистике однородность оценок дисперсий проводится по F-критерию (критерий Фишера).

Расчетное значение критерия (статистики) вычисляется по формуле

$$F_{расч} = \frac{D_{ад}}{D_B}.$$

Критическое значение $F_{кр}$ для уровня значимости $P_{ош}$ и чисел степеней свободы $f_{ад}$ и f_B , берется из таблиц F-распределения. Эти таблицы удовлетворяют условию $\text{Вер}[F \leq F_{кр}] = 1 - P_{ош}$.

Гипотеза об адекватности модели принимается, если $F_{расч} < F_{кр}$. В противном случае необходимо функцию отклика искать в виде полинома более высокого порядка.

Наконец, если после этого не удалось получить адекватную модель, то необходимо разбить область варьирования факторов на несколько подобластей и находить функцию отклика отдельно для каждой из них. Это приводит к уменьшению интервалов варьирования факторов Δk_j в каждой подобласти, а чем меньше интервал варьирования, тем проще обеспечить адекватность функции отклика.

Адекватность на основе F-критерия устанавливается при условии $l < N$, если же $l = N$, то нельзя найти оценку дисперсии адекватности $D_{ад}$. Однако в этом случае об адекватности модели можно судить по отклонению значения \bar{w} от его действительной величины в других точках области эксперимента, не совпадающих с опытными точками.

Обычно такой критической точкой является центр плана, где кодированные параметры $k_1 = k_2 = k_3, \dots, k_\alpha = 0$. Предсказанное значение отклика в центре плана равно

свободному члену b_0 . Поставив в центре плана m_0 параллельных опытов, находят среднее значение $\bar{w}^{(0)}$ и оценку дисперсии отклика $D_{ц.п.} = D_B$ в центре плана (последняя имеет $m_0 - 1$ степеней свободы), после чего проверяется статистическая значимость отклонения $\varepsilon = \bar{w}^{(0)} - b_0$. Если величина ε окажется статистически незначимой, модель признается адекватной. В противном случае гипотеза об адекватности отклоняется.

Выводы по главе 11: Наблюдается полное совпадение результатов расчётов баз данных демонстрационных примеров методик РА технических систем кафедры 703 МАИ с расчётами в программном обеспечении научного аналитического комплекса РФА «ФАКТОР-ПРОГНОЗ» («PLANETA») с дополнительными возможностями:

1) оптимизации по нелинейным весовым функциям РФА n -го порядка (см. рисунки в этом разделе 11.2 – 2D и 3D оптимизация) – такой оптимизации нет в методиках кафедры 703 МАИ ввиду сложности их реализации при отсутствии специализированного инструмента РФА типа «PLANETA» («ФАКТОР-ПРОГНОЗ»);

2) возможность быстрой работы с большими объёмами данных с наглядными графическими формами представления результатов, что позволяет за малый промежуток времени просмотреть большое количество вариантов для выбора наилучших решений. В онкологии это позволяет быстро разбивать общий поток на однородные группы по гистологии, морфологии, стадиям и подстадиям с целью оценки наилучших воздействий по всем критериям;

3) возможность записи окончательных весовых функций РФА в анкетную форму регистрации данных для автоматического прогноза по всему перечню критериев (результатирующих параметров) процесса, что особенно актуально в медицине (онкологии) для индивидуальной диагностики и подбора наилучших схем лечения.

12. Используемые материалы

1. Министерство образования Российской Федерации, Министерство по атомной энергии Российской Федерации, Министерство промышленности, науки и технологий Российской Федерации. ISBN 5-7262-0354-2. Научная сессия МИФИ-2001. Том 1. УДК 621.382(06) Электронные измерительные системы. Колосов Е.А., д. м. н., Сусулева Н.А.
2. «Экспресс-методика оценки эффективности изделий МО на государственных испытаниях», КБТОЧМАШ им. А.Э. Нудельмана, отдел эффективности, исп. Колосов Е.А., сов. секр.
3. Материалы научной работы на защиту учёной степени доктора медицинских наук «Стратегия лечения лимфогранулематоза у детей», д. м. н. Сусулева Н.А., Колосов Е.А.
4. По материалам методики МАИ, 1986 год, «Применение методов планирования экспериментов в процессе формирования облика объекта», раздел 2. «Методы планирования эксперимента». 2.1. «Основы регрессионного анализа», Москва, 1986 год,

Московский ордена Ленина и ордена Октябрьской революции авиационный институт имени Серго Орджоникидзе, к. т. н., доцент В.Б. Петров.

5. «ФАКТОР-ПРОГНОЗ – Программа РФА и создания систем высокоточного прогнозирования во всех областях науки, техники и экономики», 2006, Свидетельство Федеральной службы по интеллектуальной собственности, патентам и товарным знакам № 2006613748, автор и правообладатель: Колосов Е.А.
6. Материалы НМИЦ им. Н.Н. Блохина, РАМН, Москва, РФ, научная статья «Магнитно-резонансная томография в диагностике метастатического поражения печени и оценка эффективности лечения больных нейроэндокринными опухолями», дисс. к. м. н. (по параметрам последствий), Москва, 2021. – 121 с. НМИЦ онкологии им. Н.Н. Блохина.
7. «ФАКТОР-ПРОГНОЗ – эндометрия – Система индивидуального прогнозирования состояния эндометрия в постменопаузе по биологическим маркерам», 2014 год, Свидетельство Федеральной службы по интеллектуальной собственности, патентам и товарным знакам № 2015620269, авторы и правообладатели: академик Стрижаков А.Н., к. м. н. Исаева Э.А., Шахламова М.Н., Колосов Е.А.
8. «ФАКТОР-ПРОГНОЗ – клинико-функциональные-предикторы – Система прогнозирования переносимости хирургического лечения», 2006 год, Свидетельство Федеральной службы по интеллектуальной собственности, патентам и товарным знакам № 2007610002, авторы и правообладатели: д. м. н. Шестопалова И.М., Колосов Е.А.
9. Материалы НМИЦ им. Н.Н. Блохина, РАМН, Москва, РФ, научная статья «Магнитно-резонансная томография в диагностике метастатического поражения печени и оценка эффективности лечения больных нейроэндокринными опухолями», акад. РАН, д. м. н., проф. Б.И. Долгушин, д. м. н., проф. В.А. Горбунова, М.Г. Лаптева. Дисс., Москва, 2021. – 121 с. НМИЦ онкологии им. Н.Н. Блохина.
10. «PLANETA – Медицинская информация в клинической и экспериментальной радиологии», 2019 год, Свидетельство Федеральной службы по интеллектуальной собственности, патентам и товарным знакам № 2019617905, авторы и правообладатели: Долгушин Б.И., Трофимов И.А., Виршке Э.Р., Стилиди И.С., Лактионов К.К., Колосов Е.А.
11. Материалы исследовательских работ по определению оптимальных доз (ОФЛ, офлоксацин – опыты на мышах) на защиту учёной степени д. м. н. ФГБНУ «ЦНИИТ» – Федеральное государственное бюджетное научное учреждение «Центральный научно-исследовательский институт туберкулеза».
12. В экспресс-методике использованы примеры научных работ по РФА других медицинских организаций и все видео © Колосова Е.А.