

# Исследования в медицинской информатике в МНУЛ интеллектуальных систем и структурного анализа

С.О.Кузнецов

Повышение качества медицинских изображений

Анализ траекторий госпитализации

Оптимизация терапии в подгруппах пациентов

Предсказание естественной истории развития онкозаболеваний

Поддержка рандомизированных клинических исследований

# Медицинская информатика (МИ) [*Medical informatics*]

- Наиболее общее определение: **the discipline, dedicated to the systematic processing of data, information and knowledge in medicine and healthcare**
- Включает: **Клиническая информатика** [*Clinical Informatics*], **Оптимизация ухода за пациентами** [*Nursing Informatics*] и др.
- Примыкает к: **Биоинформатика** [*Bioinformatics*]
- Официальная история начинается в 1967 году с создания *International Medical Informatics Association (IMIA)*
- С конца 1990-х годов – специализации в ведущих университетах (например, Гарвардской медицинской школе)
  - В соответствии с новыми стандартами и сертификационными программами (например, *Board Certified in Clinical Informatics*)
- Число публикаций за последние пять лет – более 12000 статей только по базам *Springer Link* и *Elsevier Science Direct*!

# Основные драйверы развития МИ

- Программируемые электронно-вычислительные машины (компьютеры)
- Телекоммуникационные технологии
- Базы данных
- Онтологическое моделирование и базы знаний
- Робототехника
  - (телеприсутствие, робо-хирурги, роботизированные аптеки, ...)
- Молекулярная биология и фармацевтика
- Средства объективной диагностики пациентов с выдачей показателей в цифровом виде
- Доказательная медицина
- Формализация стандартов медицинской помощи
- Генетика и персонализированная медицина

# Медицинские информационные системы (МИС)

- **Базовые медицинские системы (собственно под МИС их обычно и понимают)** – управление историями болезни и карточками пациентов (*Electronic Healthcare Records*)
- **Лабораторные информационные системы** – автоматизация лабораторных исследований
- **Системы диагностики реального времени** – поддержка реанимационных мероприятий и др.
- **Системы поддержки принятия врачебных решений** – экспертные системы в клинической практике
- **Системы трансфузиологии** – управление донациями и компонентами крови
- **PACS-системы** – накопление и анализ мультимедийных диагностических данных
- **Клинические регистры** – сопровождение рандомизированных клинических исследований
- **Телемедицинские системы** – организация удалённой коллективной работы, телеконференций, телеприсутствия
- И многие другие ...

# История развития МИС

- До 1990 г. – исследования + *EHR*
- С 1990 гг. – резкое удешевление персоналок – локальные и клиент-серверные системы, повсеместное внедрение в США Европе
- С 2000 гг. – новое поколение МИС, интероперабельность, веб-интерфейсы, интеграция с диагностическими системами
- С 2010 гг. – резкое удешевление обработки больших данных, реальное внедрение крупномасштабной аналитики, начало поддержки персонализированной медицины
- С 2012 г. – появление полноценных интерфейсов для пациентов и методология постоянного сопровождения, носимое медицинское оборудование

# Прорыв на рубеже 2013 года

- **Открытые данные** в медицинской информатике
  - Новые проекты
  - Проекты, поменявшие стратегию развития и предоставления данных
- **Интеграция основных медицинских онтологий**
  - До этого основные результаты относят к 1998-2000 годам
  - *Nuance LinkBase* (<http://www.nuance.com/for-healthcare/resources/clinical-language-understanding/ontology/index.htm>)
- Всплеск интереса к **системам поддержки принятия врачебных решений (CDSS)** как следствие появления новых методов и инструментов интеллектуального анализа данных
- **Открытые API** для мобильных решений
- Обсуждение **изменения законодательства**
- Пристальное внимание к **качеству результатов научных исследований**, включая клинические исследования
  - Мечты о воспроизводимости (<http://biomolecula.ru/content/1729>)

# Доказательная медицина

- **Доказательная медицина** или медицина, основанная на доказательствах [*evidence-based medicine*] – совокупность методологических подходов к медицинской практике, опирающихся на проверку эффективности любых клинических действий, причём эффективность обосновывается рандомизированными клиническими исследованиями
  - В узком смысле это подразумевает, что решение о применении лечебных или профилактических мер по отношению к пациенту принимается **только исходя из существующих доказательств эффективности и безопасности** этих мероприятий
  - В современном смысле термин введён в **1990 году** введен группой ученых из Канады
  - Крупнейшая организация, специализирующаяся на доказательной медицине – The Cochrane Collaboration (<http://www.cochrane.org>) с лозунгом «**Trusted evidence. Informed Decisions. Better health**»

# Клинические исследования

- **Клиническое исследование/испытание** [*Clinical study/trial*] – исследование воздействия медицинских препаратов, приборов и методов лечения на человека (пациента) с целью выявления любых положительных или отрицательных результатов и оценки эффективности и безопасности предлагаемого *средства*
- Клиническое исследование (КИ) должно быть как минимум **рандомизированным двойным слепым (РКИ)**, но также очень желательно – с проверкой **плацебо-эффекта**
  - Такое КИ называется контролируемым или рандомизированным (по главной процедуре 😊)
  - Главная аббревиатура – РКИ
- РКИ – статистический эксперимент специального вида!



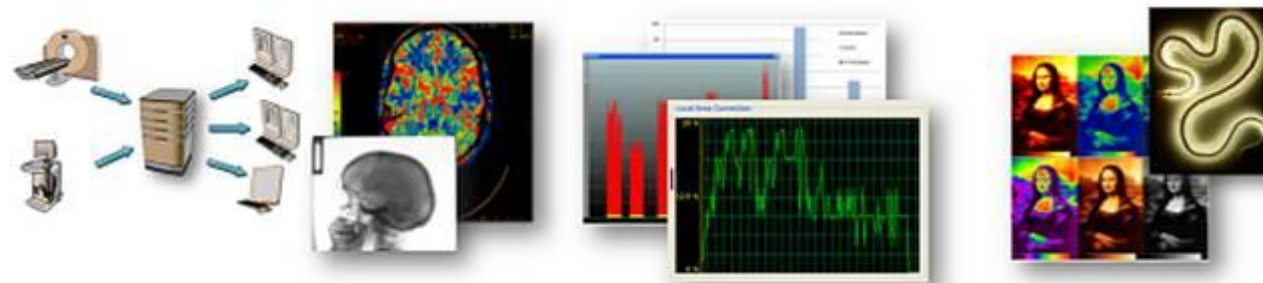
# Медицинская статистика

- Медицинская статистика – отрасль статистики, изучающая явления и процессы в области здоровья населения и здравоохранения
- Задачи медицинской статистики:
  - Разработка специальных методов исследования массовых процессов и явлений в медицине и здравоохранении
  - Выявление наиболее существенных тенденций в здоровье населения во взаимосвязи с конкретными условиями и образом жизни
- Особенности
  - Акцент на анализе временных рядов, анализе выживаемости и специальных методах подготовки данных
  - Развивалась ещё до формулировки термина «доказательная медицина» и дала много результатов, обогативших другие отрасли

- Международная научно-учебная лаборатория **Интеллектуальных систем и структурного анализа**
  - Сайт: <https://cs.hse.ru/ai/issa>
  - Руководитель – д.ф.-м.н. С.О. Кузнецов
  - Ведущий научный сотрудник – PhD, prof. Андре Щедров
- Основные направления исследований
  - Майнинг данных, обнаружение знаний и машинное обучение
  - Вычислительная логика
  - Анализ формальных понятий
  - Базы знаний и онтологическое моделирование
  - Мультимодальная кластеризация
  - Структурный анализ и прикладная теория графов
  - Автоматическая обработка текста
  - Рекомендательные системы и коллаборативные технологии
  - Анализ социальных сетей
  - Обработка больших данных

# Образовательные аспекты МИ: области знания

- **Математика** (медицинская статистика, методы оптимизации, анализ временных рядов, ...)
- **Анализ данных** и методы искусственного интеллекта (майнинг данных, визуализация данных, машинное обучение, компьютерная лингвистика, рекомендательные системы, онтологическое моделирование)
- Программная инженерия (МИС)
- Робототехника
- Менеджмент и оптимизация бизнес-процессов



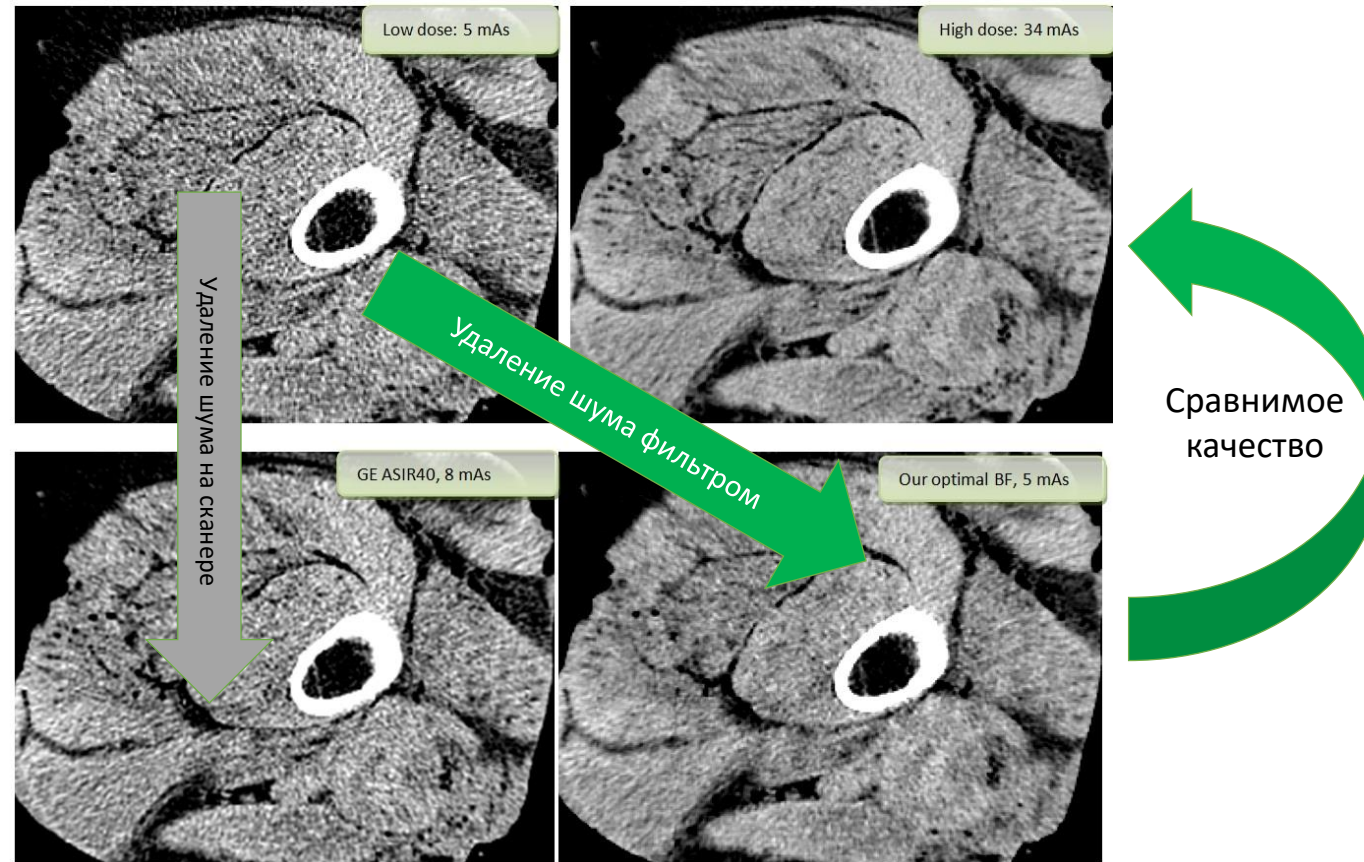
# Проект 1: Улучшение качества изображений

- Совместно с Медицинской школой Гарварда (О.С. Пьяных)
- Улучшение качества включает в себя:
  - Минимизацию побочных эффектов для пациента
  - Создание алгоритмов для повышения «веса» наиболее диагностически-важных факторов в снимке
  - Анализ факторов восприятия качества снимка человеком



# Проект 1: Улучшение качества изображений

- Удаление шума для понижения дозы радиации при сканировании пациента:



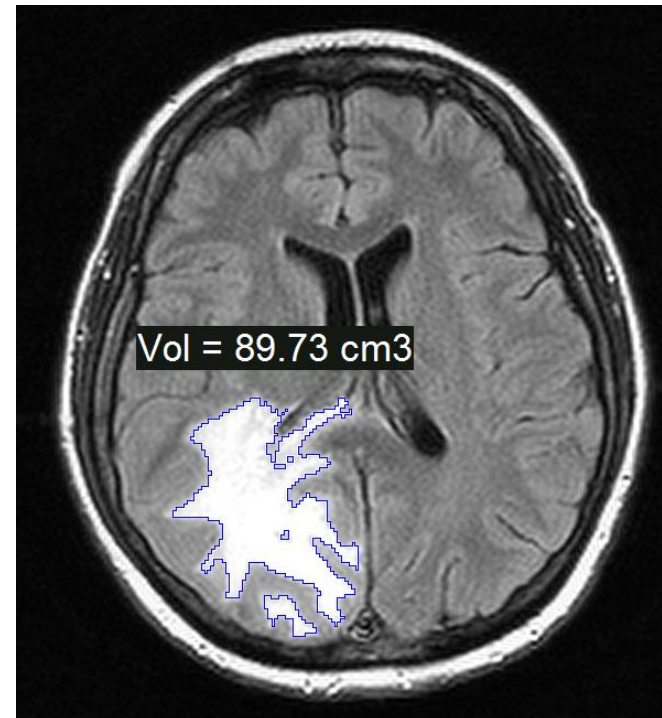
# Проект 1: Улучшение качества изображений

- Улучшение диагностического качества изображений



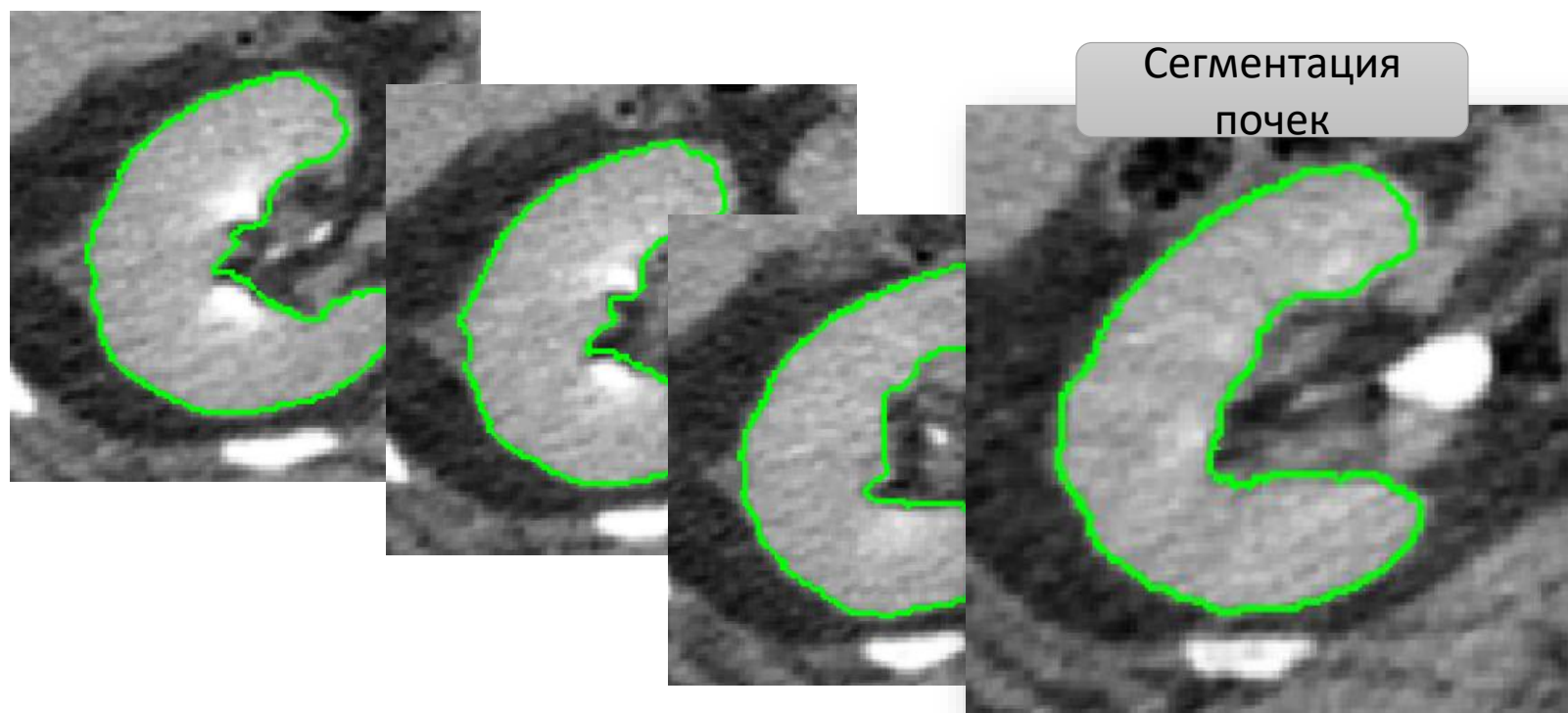
## Проект 2: Сегментация изображений

- Совместно с Медицинской школой Гарварда (О.С. Пьяных)
- Большой объем медицинских данных требует автоматических алгоритмов сегментации – разделения данных на области и анатомические органы
- Задача должна решаться в реальном времени, устойчиво к шуму и артефактам



## Проект 2: Сегментация изображений

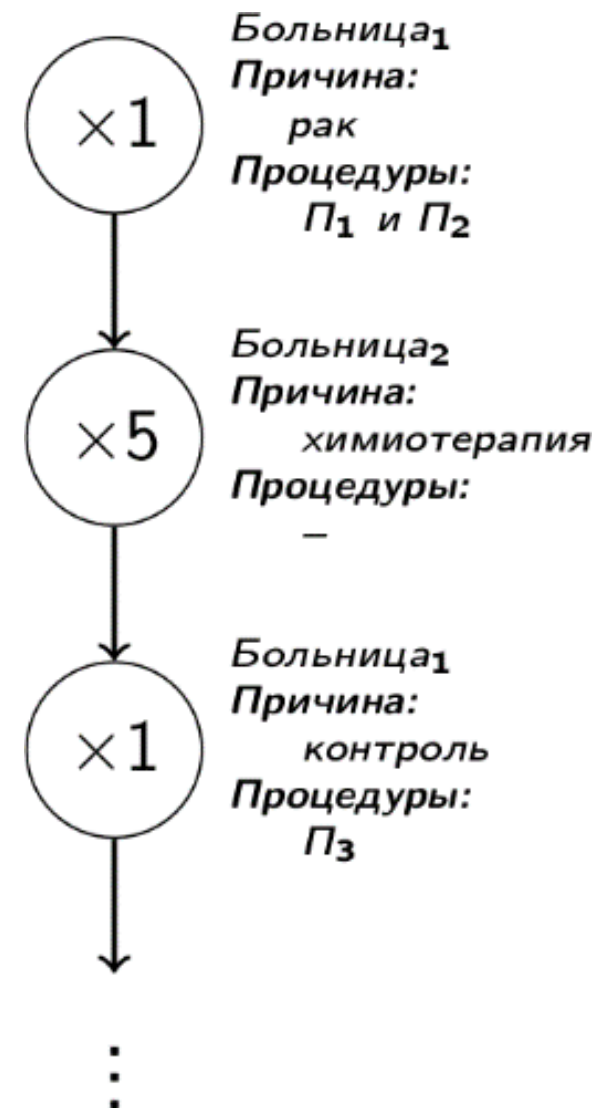
- Интерактивные алгоритмы сегментации помогают врачу быстро и надежно определить 3D-границы интересующей области:



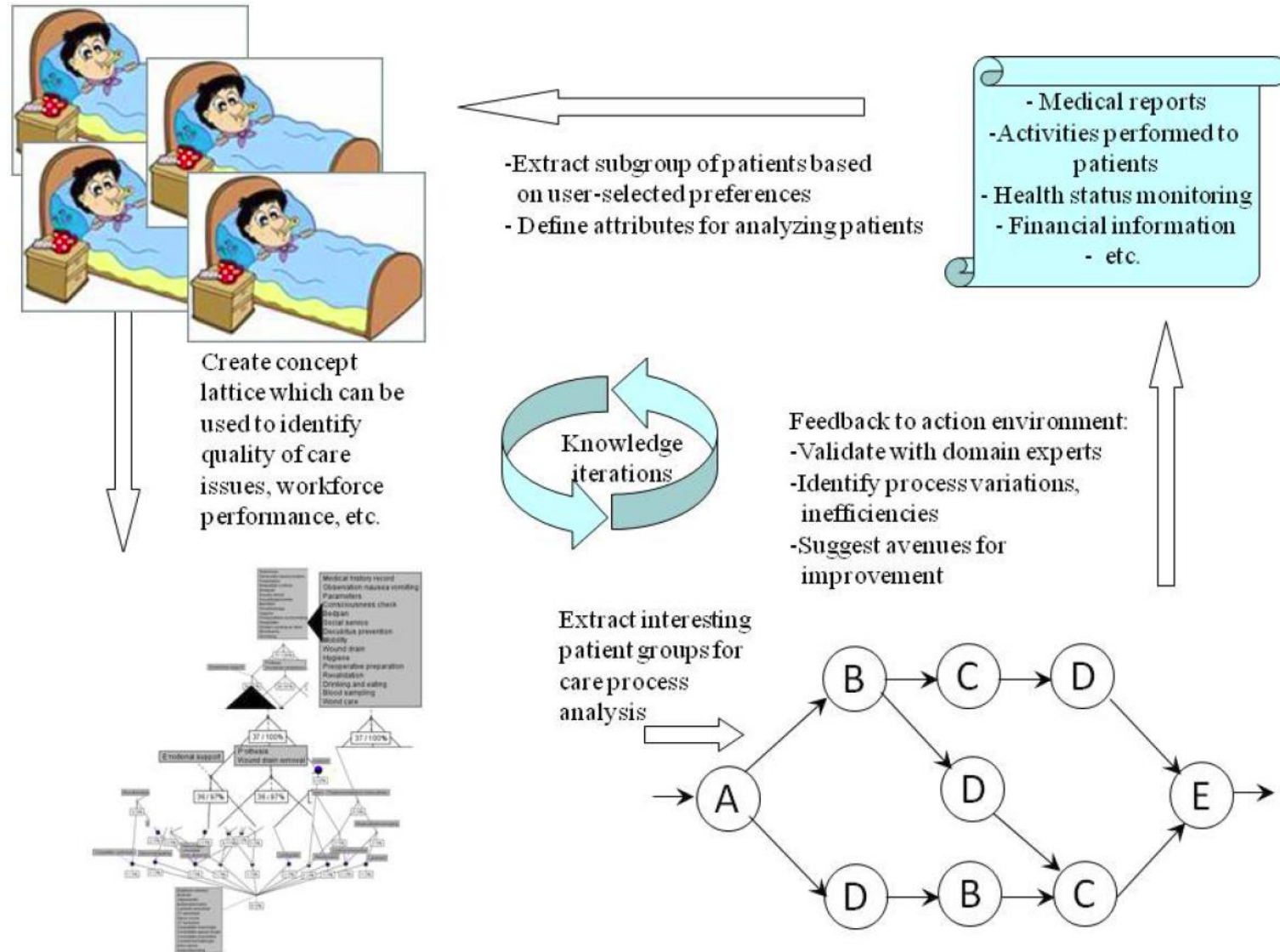


# Проект 3. Оптимизация госпитализаций

- Существует много процессов, которые могут быть оптимизированы только при участии эксперта
- Эксперт заинтересован в моделях анализа, которые:
  - Просто интерпретируются
  - Уточняются (детализируются) на нескольких уровнях
  - На верхнем уровне имеют небольшой размер (обозримы)
  - Строятся за разумное время
- Пример – модель траекторий госпитализаций пациентов
  - Большое число параметров и скрытого знания



# Проект 3. Анализ траекторий госпитализации



## Проект 3. Узорные структуры

- Траектории госпитализаций изучаются с помощью «**узорных структур**» [Pattern Structures] – формализма теории анализа формальных понятий

Пациент	Траектория госпитализаций
$p_1$	...
...	...
$p_i$	$\left\langle \dots; \left[ \begin{array}{c} \text{Болница} \\ \text{Причина} \end{array} \right], \left\{ \text{Мед.Проц.}_1, \text{Мед.Проц.}_2, \dots \right\}, [\text{Повторение}] ; \dots \right\rangle$
...	...
$p_j$	$\langle [H_1, \text{Рак}, \{P_1, P_2\}]; [H_1, \text{Хим.Подг.}, \{P_3\}]; [H_2, \text{Химио}, \{\}] * [18] \rangle$
...	...
$p_n$	...

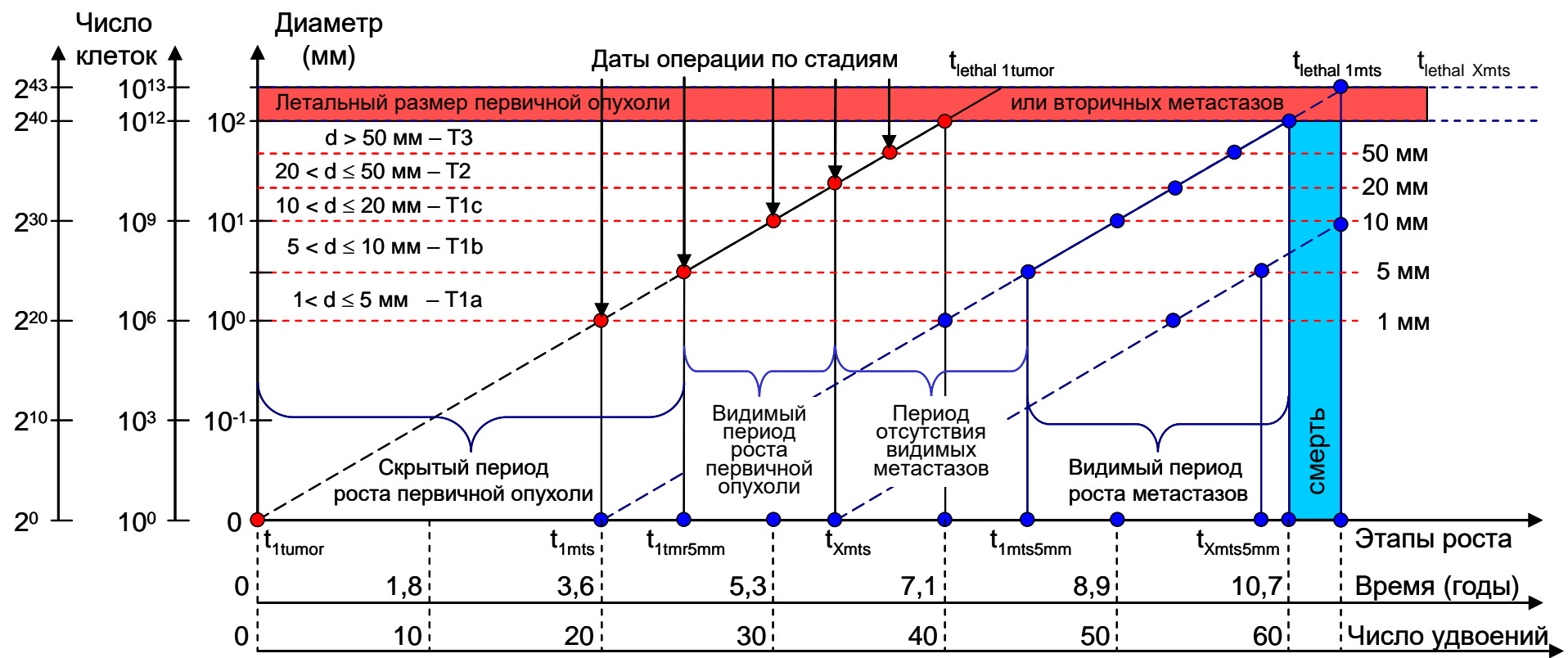
## Проект 4. Математическое сопровождение рандомизированных клинических исследований

- Основа **доказательной медицины** и значимый раздел **клинической информатики**
- Базируются на **медицинской статистике** и различных **регистрах**
- Стандартизованы и обладают развитой **методологией**  
(<http://acto-russia.org>)
- Поддерживаются развитыми **базами данных и знаний**
  - Например: <http://www.controlled-trials.com>, <http://www.clinicaltrials.gov>
- Актуальные задачи
  - Эффективное сопровождение мультицентровых исследований
  - Оптимизация дизайна исследования
  - Повышение эффективности анонимизации и ослепления
  - Новые методы анализа результатов
  - Мета-анализ

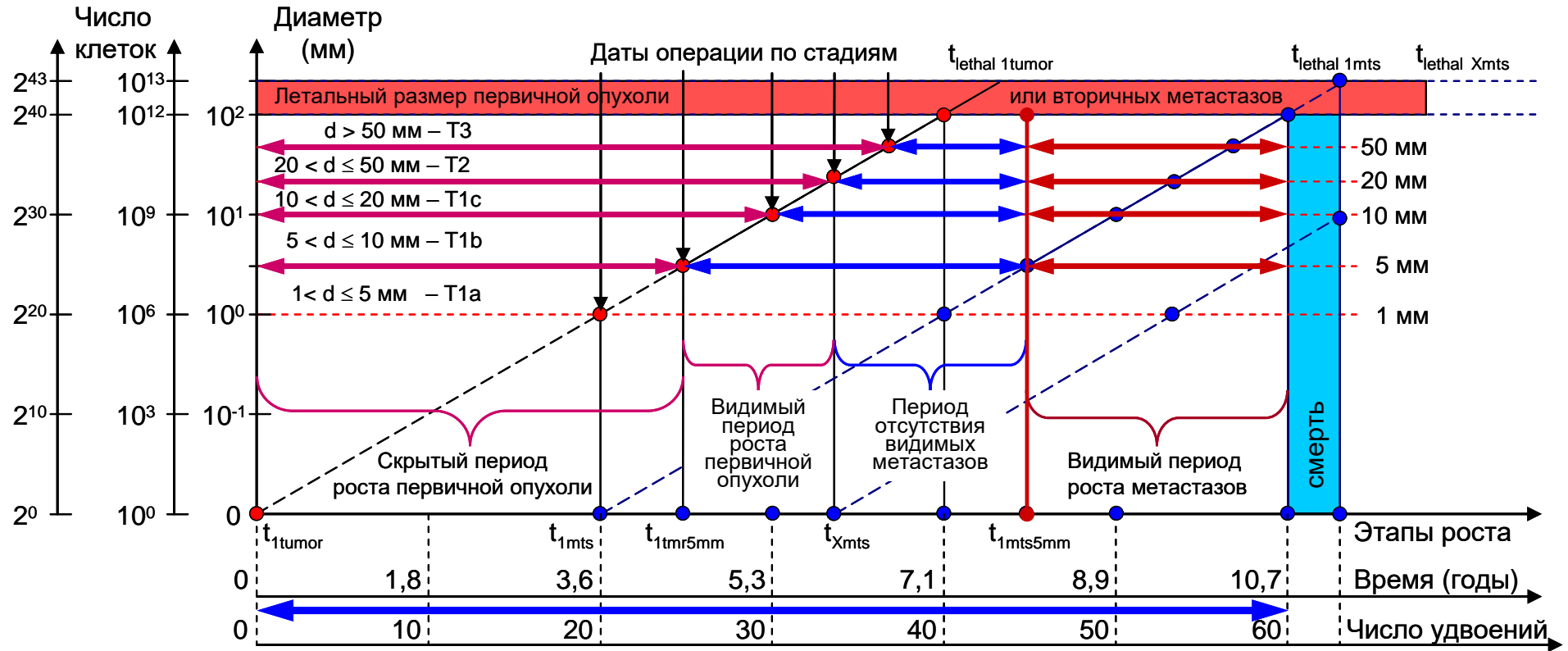
# Проект 5. Предиктор естественной истории РМЖ

- Пример относительно простой, но удобной и хорошо интерпретируемой модели
- Рак молочной железы – один из самых распространённых вариантов раковых заболеваний у женщин
  - Хорошо изучен с точки зрения статистики и основных механизмов
  - Но до сих пор накопленные данные не были сведены в единую модель
- «Объединенная математическая модель роста рака молочной железы» (*CoMPaS*) и её программная реализация
  - Тюрюмина Э.Я., Незнанов А.А. Объединенная математическая модель роста первичной опухоли и первичных метастазов рака молочной железы (IV стадия РМЖ) // 40-я междисциплинарная школа-конференция «Информационные технологии и системы», 2016. С. 13-21.

# Модель естественной истории рака молочной железы

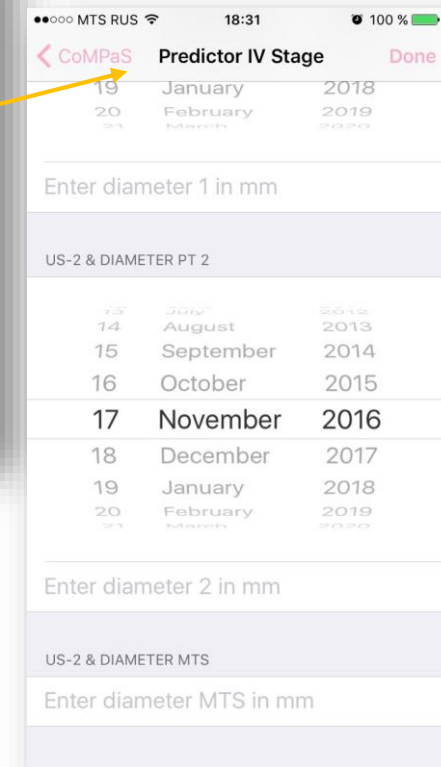
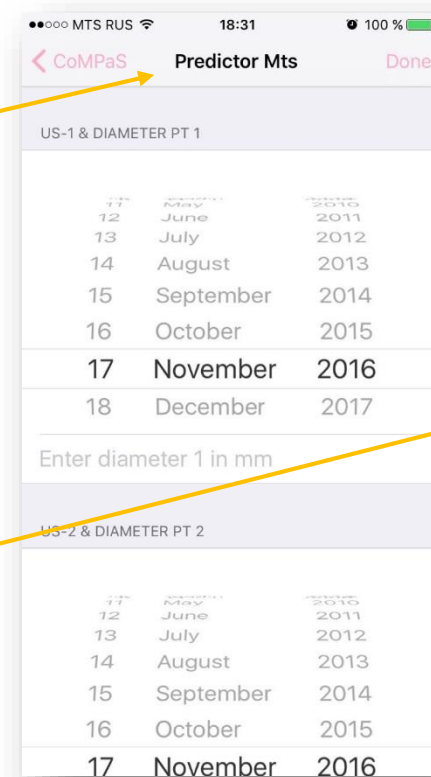
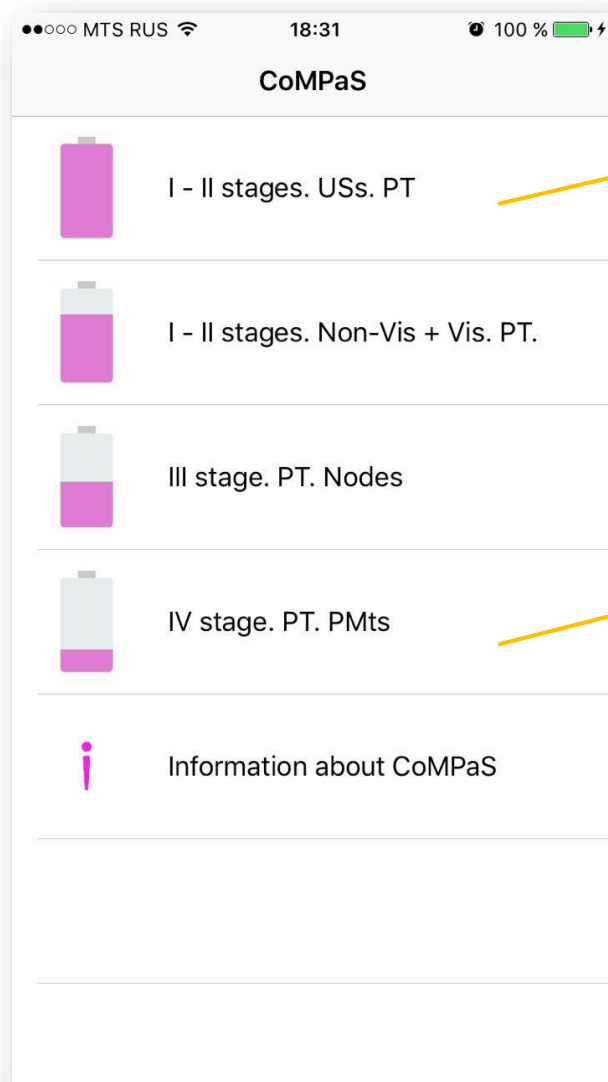


# Определены критические периоды роста опухоли и метастазов



- «скрытый» период роста первичной опухоли
- «видимый» период роста первичной опухоли
- «скрытый» период роста вторичных отдаленных метастазов
- «видимый» период роста вторичных отдаленных метастазов

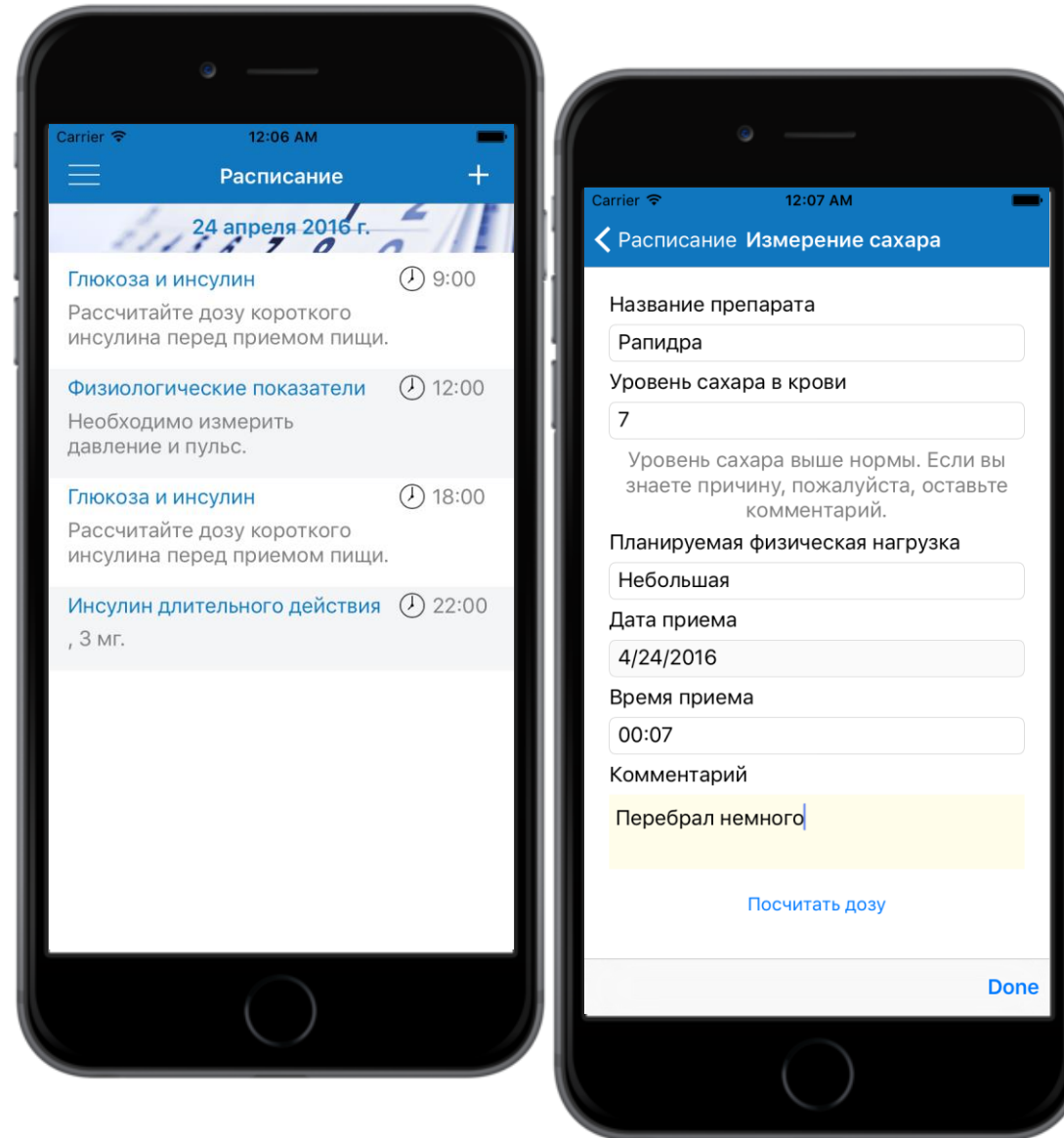
# Программная реализация (мобильный вариант)





# Проект 7. Мониторинг действий пациентов

- Амбулаторное лечение под контролем врача
- Гибкая настройка терапии
- Прогнозирование течения заболевания
- Уведомления
- Рекомендации врача



Основные разработчики: Р.А. Бородин, А.А. Незнанов

# Проект 7. Мониторинг действий пациентов

Врач: Доктор Доктор  
Докторова  
Докторова

Настройка терапии диабета  
Пациент: Назик Аркадий Валерьевич

Индикаторы гликемии

Целевая ГП Мин: 5 Макс: 6  
Дополнительная ГП Мин: 2 Макс: 10

Курсовой инсулин

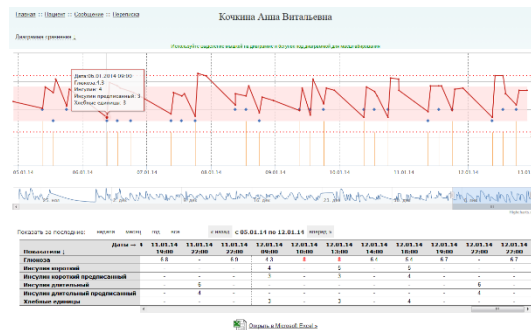
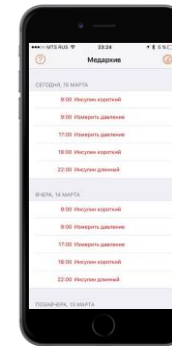
Препарат короткого действия Карбогемин

Увеличенный коэффициент	Часы приема
1 0 6:00 - 1,2	12:00
2 0 13:00 - 1,8	16:00
3 0 18:00 - 2,0	
4 0 23:00 - 3	

Панкреатический коэффициент: 3,3

Настройка терапии  
врачом

Использование  
мобильного  
приложения  
пациента для  
контроля гликемии



Доктор  
использует  
отчеты для  
дистанционной  
коррекции  
терапии

## Внедрение в практику клиницистов?

- **Отражаются ли вышеперечисленные результаты на распространённых МИС?**
- **Насколько «средний» клиницист задействует вышеперечисленное в своей практике (хотя бы и опосредовано)?**
- **Может ли клиницист потрогать рекламируемые методы и инструменты руками?**
- **Существует ли в России реальный проект по сбору, сертифицированной анонимизации и открытию первичных данных РКИ?**

- Клиническая информатика всё более востребована в здравоохранении как **прикладная дисциплина**, ориентирующая на достижение реального результата
  - **Интересные приложения** математики и информатики, ...
- Новые технологии открывают новые возможности
  - Изменения методологии, методов и инструментов



## Проект 8. Анализ ответа на терапию

- Совместно с ФНКЦ ДГОИ им. Д. Рогачева (А.И. Карачунский)
- Выявление подгрупп пациентов
  - Задача – персонализация лечения на основе детальных данных рандомизированных контролируемых исследований
- Разнообразные методы – от деревьев решений до нейронных сетей и ленивой классификации
- Последние достижения:
  - Хорошо интерпретируемый метод на основе замкнутых описаний
    - Среди описаний, которым соответствует одно и то же подмножество пациентов, замкнутое описание является наиболее полным и позволяет не упустить ни один из признаков, который на самом деле является прогностическим
  - Оптимизация пространства разнородных сложноструктурированных признаков пациентов

# ФНКЦ ДГОИ им. Дмитрия Рогачева

- **Федеральный научно-клинический центр детской ГЕМАТОЛОГИИ, ОНКОЛОГИИ И ИММУНОЛОГИИ им. Д. Рогачева**
  - Крупнейший в Европе центр по лечению и исследованию онкологических заболеваний детей и подростков
  - Под руководством ведущих специалистов ФНКЦ ДГОИ ведётся разработка и внедрение эффективных протоколов терапии заболеваний крови, злокачественных новообразований, патологий иммунной системы и других тяжелых заболеваний детского возраста
  - Для повышения эффективности лечения и организации научных исследований в ФНКЦ ДГОИ активно разрабатываются и внедряются современные информационные системы и технологии



# Что такое ФНКЦ для специалиста в ИТ?



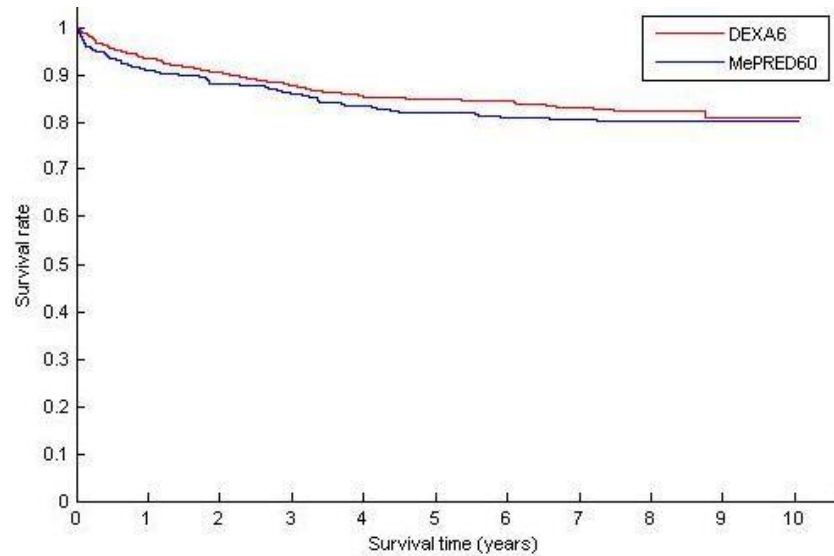
- Исключительно сложные больные с опасными заболеваниями
  - Их судьба зависит от согласованности и эффективности работы всех врачей и сотрудников разных подразделений
  - Необходимость адекватного сопровождения больных и грамотной организации работы персонала (оптимизация основных и вспомогательных бизнес-процессов)
- С точки зрения информатизации, оптимизации и анализа данных
  - Сложность структуры (> 100 подразделений, > 1200 сотрудников)
  - Уникальность многих элементов для России
  - Совершенно новые информационные системы и пилотные проекты
  - Участие в федеральных проектах в области здравоохранения
  - Активная научно-исследовательская и образовательная деятельность

# Обнаружение знаний в задаче сравнения эффективности стратегий лечения

- Клинические исследования в детской онкологии
  - Острая лимфобластная лейкемия – протоколы серии Москва-Берлин
- Суть РКИ: для детей, больных лейкозом, случайным образом выбирается один из двух вариантов протокола лечения
  - В среднем эффект от лечения одинаков
  - Можно ли выделить и описать подгруппы пациентов, в которых один из видов лечения значительно эффективнее другого?
- Суть предложения – использование специальных видов кластеризации и порождения гипотез на основе различных методов машинного обучения

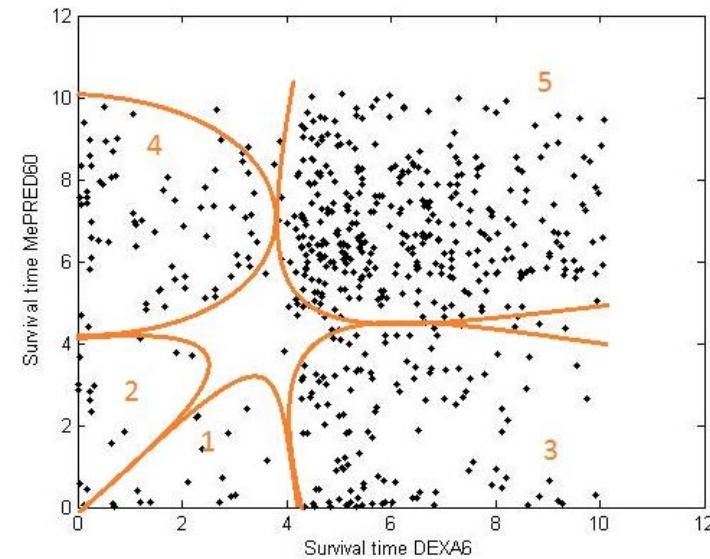


# Идея



Стандартные подходы к оценке выживаемости:

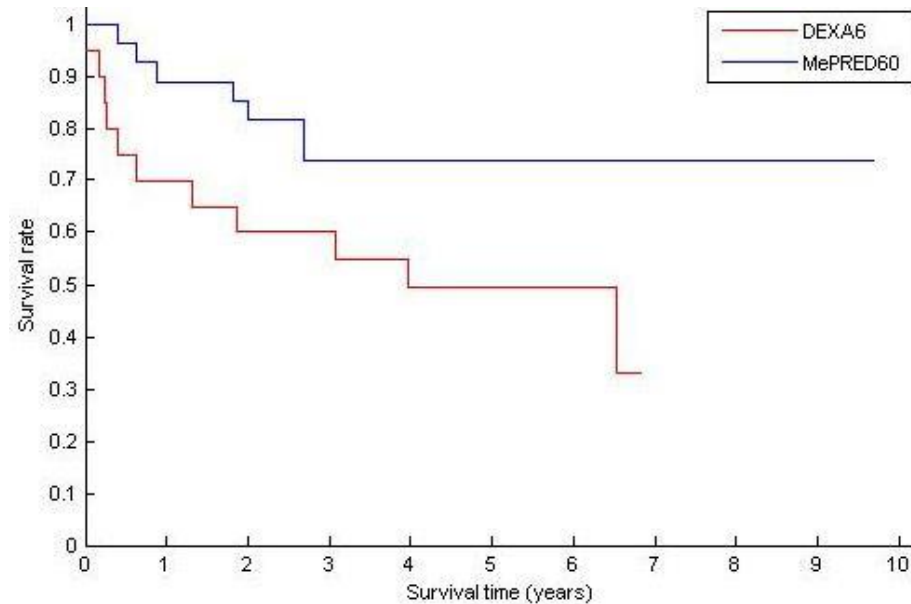
- Кривые выживаемости Каплан-Майера (на рисунке)
- Логранговый критерий
- Модель Кокса
- и тд.



Предложенный подход:

- Найти пары физиологически схожих пациентов
- Разбить пары на группы согласно эффективности каждого из двух видов лечения
- Попытаться описать эти группы методами машинного обучения
- Проверить полученные гипотезы стандартными инструментами анализа выживаемости

# Пример подтверждённой гипотезы



Показаны графики выживаемости в 2 группах

- Для пациентов с пальпируемым размером селезенки  $\geq 3,5$  см в возрасте старше 6,6 лет и *pre-pre-B* или *pre-B* иммуно-фенотипом метилпреднизолон в дозе 60 мг/м<sup>2</sup>/сут (*MePRED*) значительно эффективнее, чем дексаметазон в дозе 6 мг/м<sup>2</sup>/сут (*DEXA*)
- Всего таких пациентов **47**
- При лечении всех пациентов *MePRED*, **возможно**, удалось бы спасти около **8** пациентов, т.е. порядка 1/6 всей группы (в отсутствие других факторов)
  - С учётом максимальной поправки на множественные сравнения – **2** пациентов

*Спасибо за внимание!*

- **Контакты С.О.Кузнецова**

- International Laboratory for Intelligent Systems and Structural Analysis, Faculty of Computer Science, NRU HSE, Moscow, Russia
- E-mail: [skuznetsov@hse.ru](mailto:skuznetsov@hse.ru)
- Web-site: <http://hse.ru/staff/skuznetsov>

