VIII СЪЕЗД МЕДИЦИНСКИХ СТАТИСТИКОВ ГОРОДА МОСКВЫ

МОСКОВСКАЯ МЕДИЦИНА

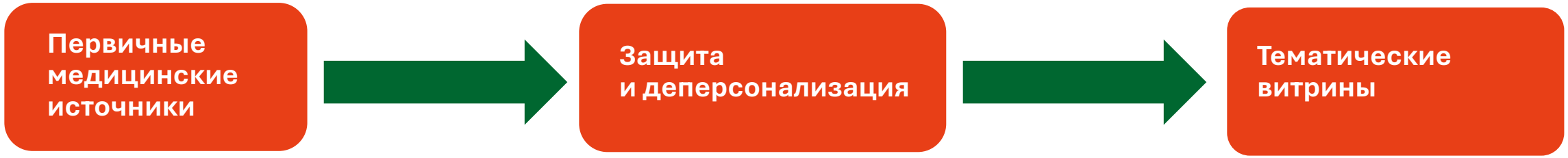
НИИ ОРГАНИЗАЦИИ ЗДРАВООХРАНЕНИЯ И МЕДИЦИНСКОГО МЕНЕДЖМЕНТА

Витрины данных: слой между медицинской информационной **системой** и статистической **формой**

Носов Алексей Евгеньевич

Заместитель начальника отдела методологии информационных ресурсов и баз данных Центра медицинской статистики НИИ организации здравоохранения и медицинского менеджмента

Что такое витрины данных?



«Витрина — структурированный набор фактов под свою предметную область»

- Возраст
- Пол
- Даты начала и окончания
- Коды диагнозов
- и т. д.

«Витрина — это не просто таблица. Это бизнес-сущность с владельцем, описанием и правилами обновления»

От правил врачей-методистов к автоматической форме

Методология
Получаем чёткие правила консолидации данных



Перевод в код
Оптимизация для миллиардов строк



Сверка
Сравниваем с расчётами врачей-методистов



Автоматизация
Запуск по расписанию, инкрементальное обновление



ФЛК
Форматно-логический контроль перед выдачей



МЕТОДОЛОГИЯ СБОРКИ
Совокупность правил и требований, определяющих, какие данные брать, как их фильтровать, связывать и агрегировать для получения корректного показателя формы

ПРЕДСТАВЛЕНИЕ (VIEW)
Виртуальная таблица, которая не хранит данные физически, а формирует их «на лету» из базовых витрин по заданной логике

ВАЛИДАЦИЯ ДАННЫХ
Процесс проверки соответствия собранных данных эталонным расчётам методистов для выявления расхождений и ошибок

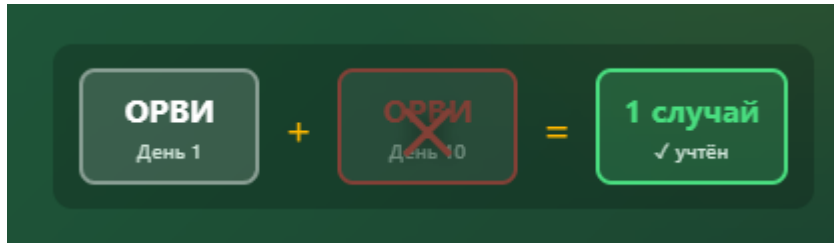
ИНКРЕМЕНТАЛЬНОЕ ОБНОВЛЕНИЕ
Обработка только новых или изменённых записей с момента последнего запуска, без пересчёта всего массива данных

ФЛК
Автоматическая проверка бизнес-правил (временные последовательности, справочная валидность, логические связи) перед формированием формы

Пример операций алгоритма

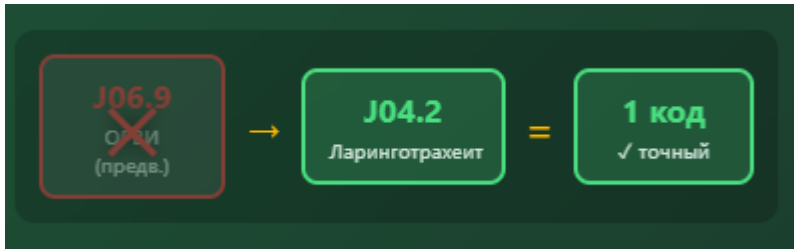
Исключение дублей
10 дней

Пациент с ОРВИ → через 10 дней снова ОРВИ = один случай (правило 21 дня). Раньше: тысячи строк в Excel вручную. Сейчас: автоматически за секунды."



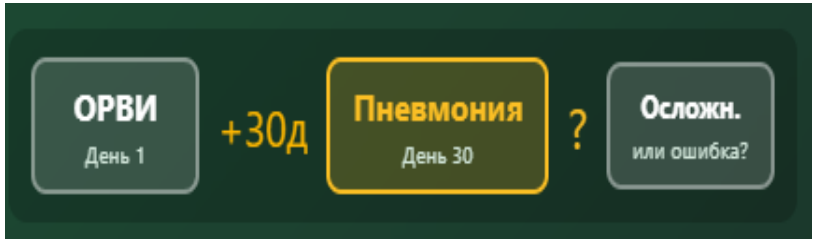
Верификация диагнозов
Предварительный → уточнённый

Предварительный «ОРВИ» (J06.9) уточнён до «острый ларинготрахеит» (J04.2). Алгоритм оставляет только уточнённый код — без ложного завышения заболеваемости.



Контроль осложнений
ОРВИ → пневмония (30 дней)

Пневмония в течение 30 дней после ОРВИ: алгоритм помечает случай для проверки. Осложнение или ошибка первичной диагностики? От этого зависит учёт в форме.



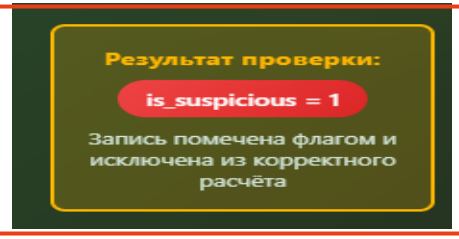
Результат: готовая таблица с показателями, которая обновляется каждую ночь. Врач-статистик видит не пустую форму, а уже рассчитанные данные для проверки.

Форматно-логический контроль как страж качества

ФЛК работает не постфактум, а в процессе самой сборки

ОШИБКА ФЛК

Врач по ошибке указал **дату завершения этапа раньше, чем дату её начала.**
Или указал код диагноза, который не соответствует полу или возрасту пациента.



Маршрутизация уведомлений по типам ошибок:

Ошибки первичных мед. документов
↓
Уведомление уходит **врачу-статистику** в медицинскую организацию

Ошибки хранения в витринах
↓
Уведомление получает **команда инженеров данных**

Ошибки в логике сборки форм
↓
Разбираемся совместно с **врачами-методистами**

Панель мониторинга врача-статистика

"По пяти записям сработало правило: ***дата завершения раньше даты начала***

Форматно-логический контроль как страж качества

Традиционный подход

- Сбор формы раз за период (месяц, квартал, год)
- Дни на ручную сверку и исправление ошибок
- Стресс при сборке годовых форм (обработка годового объема за несколько дней)
- Роль: врач-статистик как **«сборщик данных»**

Современный подход

- Доступность данных в режиме, близком к реальному времени
- Ежедневное обновление витрин и постоянный контроль качества
- Время тратится на анализ: почему возникла ошибка и как её предотвратить
- Роль: врач-статистик как **эксперт по качеству**, работающий с исключениями

Главное преимущество слоя витрин: скорость реакции на изменения

Новая форма или мониторинг

Существующие витрины
данные уже собраны и очищены

Новый запрос
написание алгоритма

Готовая форма
За считанные часы, а не месяцы