



Перспективы прогнозной аналитики на основе больших данных и машинного обучения в повышении эффективности работы здравоохранения

Гусев Александр, директор по развитию проекта Webiomed,  
ассоциация «Национальная база медицинских знаний»

## Основные цифры здравоохранения

95%

### МО ВНЕДРИЛИ ЭМК

В среднем среди развитых стран мира. Мы в России приближаемся к этому показателю

2314

### ЭКЗАБАЙТ ДАННЫХ

Было произведено в сфере здравоохранения в 2020 году

48%

### УВЕЛИЧЕНИЕ ДАННЫХ

Каждый год объем генерируемых здравоохранением данных растет на 48%

25

### ТЫС. ПЕТАБАЙТ ДАННЫХ

Производит в среднем автоматизированная на данном этапе развития МО

1 трлн

### УСТРОЙСТВ К 2035 г.

Будут собирать различную медицинскую информацию.

## Направления сбора данных

- ✓ Электронные медицинские карты
- ✓ Данные PACS и ЛИС
- ✓ Генетические данные
- ✓ Прием и движение лекарств
- ✓ Данные с персональных медицинских устройств
- ✓ Данные мобильных приложений пациентов
- ✓ Данные социальных сетей
- ✓ Данные окружающей среды

## Основные цифры здравоохранения

71%

### СМЕРТЕЙ

Вызвано хроническими неинфекционными заболеваниями (ХНИЗ)

88%

### ЗАТРАТ НА ЗДРАВООХРАНЕНИЕ

Приходится на диагностику и лечение пациентов с ХНИЗ

80%

### СЛУЧАЕВ ЗАБОЛЕВАНИЯ ХНИЗ

Можно предотвратить, если выявлять и лечить пациентов на ранней стадии ХНИЗ

8X

### ЗАТРАТЫ НА ЛЕЧЕНИЕ ХНИЗ

Выше, чем затраты на скрининг и профилактику.

55

### МЛРД. ДОЛЛ. США

Можно экономить за счет эффективного скрининга и профилактики заболеваний

## Ключевые технологии



Большие данные



Искусственный интеллект



Прогнозная аналитика



Носимые устройства



Мобильные приложения



Телемедицина



Удаленный мониторинг

Прогнозная аналитика - это использование **данных**, **математического анализа** и **моделирования**, которые помогают выявлять и понимать шаблоны ретроспективных данных, прогнозировать будущие события и принимать наилучшие решения



**Управленческая**  
прогнозная аналитика



**Клиническая**  
прогнозная аналитика



Healthcare Analytics Market by Type, Component, Delivery Mode, Application, End user and Geography- Global Forecast to 2027, <https://www.meticulousresearch.com/product/healthcare-analytics-market/>

# Условная схема прогнозирования

5



# От накопления к эффективности

6

Базовая информатизация

Управление на основе данных

Предиктивное управление

Данные

Понимание

Действия

Предсказания



Сбор цифровых  
данных




Анализ и  
процессы



Мониторинг  
и аналитика



Прогнозирование  
ситуаций

A large iceberg floats in the ocean. The visible tip is jagged and white against the blue sky. The submerged part is a massive, rectangular block of ice, illustrating the concept of hidden potential. The water is a deep blue, and the horizon is visible in the distance.

То что мы видим и используем:  
знания, инструменты

Поразительные  
возможности для решения  
самых сложных проблем, о  
которых мы уже знаем, но  
еще не применяем на  
практике

# Взрывной рост научных исследований

**83**  
**ТЫС.** | Публикаций  
В международной  
рецензируемой  
литературе посвящено  
оценке рисков ССЗ

**2012**  
**ГОД** | Активное изучение  
Перспектив создания прогнозных  
моделей на основе больших данных  
и машинного обучения для  
повышения точности предсказания  
ССЗ

**91**  
**%** | Точность  
Прогнозирования развития  
гипертонии в течении 1  
года в работе Ye C, 2018,  
Китай







IDx-DR: автономная диагностическая система на основе искусственного интеллекта для выявления диабетической ретинопатии. 2018 г. FDA впервые выдало разрешение на использование ИИ-системы. В этом же году университет Айовы стал первой МО в США, которая запустила прибор в работу.

ИИ классифицирует пациента: имеется ли у него ретинопатия средней или высокой тяжести.

Эффекты такой ИИ-системы:

- Работу с пациентом может провести любой средний медперсонал (нет врача -> теперь не проблема)
- Оценку мед.данных осуществляет ИИ-система, дорогостоящий и редкий врач-офтальмолог для рутинной «расшифровки» теперь не требуется
- Очень удобно пациенту: не надо ехать / ждать приема офтальмолога



В США впервые разрешили ИИ проводить диагностику людей без участия врачей  
<https://habr.com/ru/post/411689/>

ARTICLE OPEN

## Scalable and accurate deep learning with electronic health records

Alvin Rajkumar<sup>1,2</sup>, Eyal Oren<sup>1</sup>, Kai Chen<sup>1</sup>, Andrew M. Dai<sup>1</sup>, Nissan Hajaj<sup>1</sup>, Michaela Hardt<sup>1</sup>, Peter J. Liu<sup>1</sup>, Xiaobing Liu<sup>1</sup>, Jake Marcus<sup>1</sup>, Mimi Sun<sup>1</sup>, Patrik Sundberg<sup>1</sup>, Hector Yee<sup>1</sup>, Kun Zhang<sup>1</sup>, Yi Zhang<sup>1</sup>, Gerardo Flores<sup>1</sup>, Gavin E. Duggan<sup>1</sup>, Jamie Irvine<sup>1</sup>, Quoc Le<sup>1</sup>, Kurt Litsch<sup>1</sup>, Alexander Mossin<sup>1</sup>, Justin Tansuwan<sup>1</sup>, De Wang<sup>1</sup>, James Wexler<sup>1</sup>, Jimbo Wilson<sup>1</sup>, Dana Ludwig<sup>2</sup>, Samuel L. Volchenboum<sup>3</sup>, Katherine Chou<sup>1</sup>, Michael Pearson<sup>1</sup>, Srinivasan Madabushi<sup>1</sup>, Nigam H. Shah<sup>4</sup>, Atul J. Butte<sup>2</sup>, Michael D. Howell<sup>1</sup>, Claire Cui<sup>1</sup>, Greg S. Corrado<sup>1</sup> and Jeffrey Dean

Predictive modeling with electronic health record (EHR) data is anticipated to drive personalized medicine and improve healthcare quality. Constructing predictive statistical models typically requires extraction of curated predictor variables from normalized EHR data, a labor-intensive process that discards the vast majority of information in each patient's record. We propose a representation of patients' entire raw EHR records based on the Fast Healthcare Interoperability Resources (FHIR) format. We demonstrate that deep learning methods using this representation are capable of accurately predicting multiple medical events from multiple centers without site-specific data harmonization. We validated our approach using de-identified EHR data from two US academic medical centers with 216,221 adult patients hospitalized for at least 24 h. In the sequential format we propose, this volume of EHR data unrolled into a total of 46,864,534,945 data points, including clinical notes. Deep learning models achieved high accuracy for tasks such as predicting: in-hospital mortality (area under the receiver operator curve [AUROC] across sites 0.93–0.94), 30-day unplanned readmission (AUROC 0.75–0.76), prolonged length of stay (AUROC 0.85–0.86), and all of a patient's final discharge diagnoses (frequency-weighted AUROC 0.90). These models outperformed traditional, clinically-used predictive models in all cases. We believe that this approach can be used to create accurate and scalable predictions for a variety of clinical scenarios. In a case study of a particular prediction, we demonstrate that neural networks can be used to identify relevant information from the patient's chart.

npj Digital Medicine (2018)1:18; doi:10.1038/s41746-018-0029-1

### Материалы исследования:

- Выборка из ЭМК 216 тыс. госпитализаций (114 тыс. уникальных пациентов) из 2х медицинских центров США
- Извлекли 46 млрд. «токенов» (признаков)
- Прогнозное моделирование с использованием данных электронных медицинских карт (EHR)

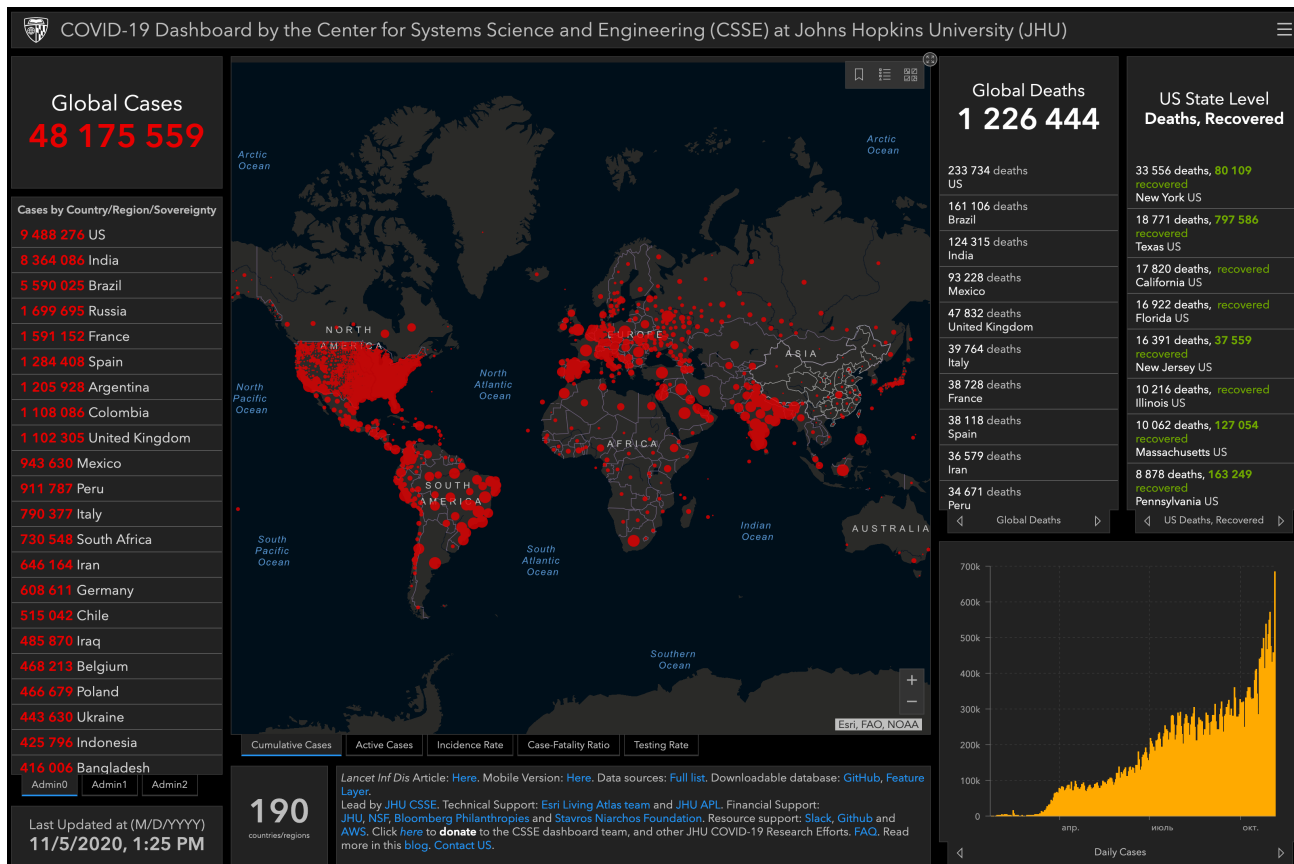
### Точность предсказания [ROC\_AUC] :

- Внутрибольничная смерть в течение 24 часов: **95%**
- Постановка диагноза: **90%**
- Осложнение заболевания после выписки в течение 30 дней: **76%**
- Длительность пребывания в МО: **86%**



# Кейс: прогнозирование пандемии COVID-19

11



Кейсы, точность, эффективность :

- Китайская модель Yu-Feng Zhao и соав. предсказывала число заболевших на 7, 8 и 9 дней с ошибкой (MAPE) 1,65-4,75%
- Vayat и соав. представили модель, предсказывающую диагноз COVID-19, AUC 86,4%
- Yan и соав. представили модель, предсказывающую смерть в стационаре, AUC 90%



Технологии прогнозной аналитики в борьбе с пандемией COVID-19,

<https://webiomed.ai/publikacii/tekhnologii-prognoznoi-analitiki-v-borbe-s-pandemiei-covid-19/>

16

МЛРД. ДОЛЛ. США

Составил размер глобального рынка медицинской аналитики в 2019 г.

84

МЛРД. ДОЛЛ. США

По прогнозу Meticulous Research составит рост рынка к 2027 г.

26%

РОСТ РЫНКА В ГОД

Один из самых динамично развивающихся секторов цифрового здравоохранения

## Самые перспективные тренды:

- ✓ Предписательная аналитика (рекомендательные системы)
- ✓ Облачные сервисы / SaaS
- ✓ Интеграция с системами ведения ЭМК и автоматический back-end анализ
- ✓ Системы поддержки принятия клинических решений

## Основные эффекты прогнозной аналитики

- ✓ Сократить неэффективные потери благодаря прогнозированию событий, например пустые койки, потери просроченных лекарств и т.д.
- ✓ Оперативно перераспределять ресурсы при изменении нагрузки на МО, например при эпидемиях
- ✓ Выявлять пациентов высокого риска до момента манифеста ХНИЗ и тем самым существенно повысить эффективность профилактики
- ✓ Сократить ручную обработку и анализ больших данных, получать не текущие показатели (врачу, руководителю), а сразу опасные ситуации

Компания развивает платформу TAVHealth для оказания медицинской помощи на дому, используя прогнозную аналитику. Заказчики: страховщики, правительство, работодатели. SGFY работает в двух сегментах: оказание услуг в отдельных эпизодах медицинской помощи и оценка состояния здоровья на дому. Signify Health обслуживает 26 из 50 планов Medicare Advantage (медицинское страхование) и увеличила количество ежегодных оценок на дому с примерно 390 тыс. в 2015 году до почти 1,1 млн в 2019 году.

# 15%

# 10%

**РОСТ ЛЕЧЕНИЯ НА ДОМУ**  
Взамен госпитализации в стационар

**СОКРАЩЕНИЕ**  
Повторных госпитализаций за счет более эффективного лечения

# 417

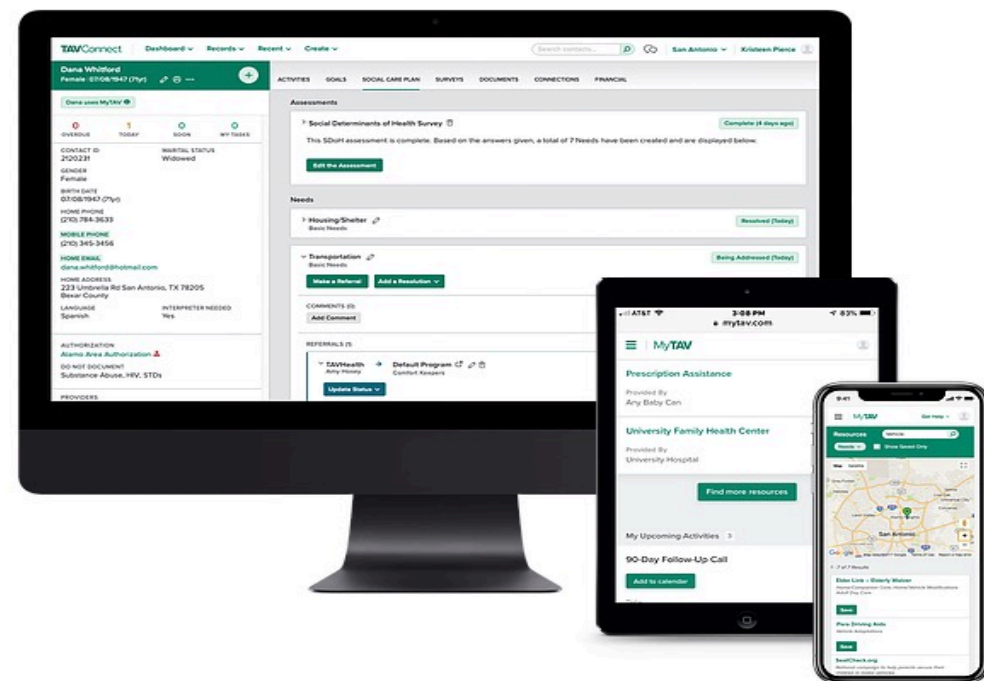
# 7.12

# 564

**МЛН. ДОЛЛ. США**  
Выручка компании в 2020 г.

**МЛРД. ДОЛЛ. США**  
Капитализация компании после IPO

**МЛНД. ДОЛЛ. США**  
Инвестиции в компанию после выхода на IPO



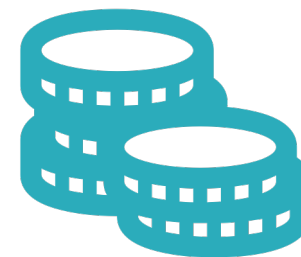
## Драйверы развития технологий

- ✓ **Нехватка ресурсов** из-за роста численности населения и распространенности ХНИЗ (финансовое обеспечение, кадры и т.д.)
- ✓ Растущие требования обеспечить **качество и доступность** медицинской помощи
- ✓ Острая необходимость **контролировать расходы** и эффективность
- ✓ Рост **больших данных**
- ✓ **Сокращение стоимости** хранения и обработки информации
- ✓ Развитие **искусственного интеллекта**

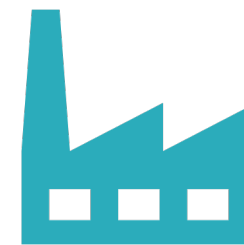
## Главные выгодоприобретатели



Государственные органы



Страховые компании



Работодатели



Амбулаторные медицинские организации



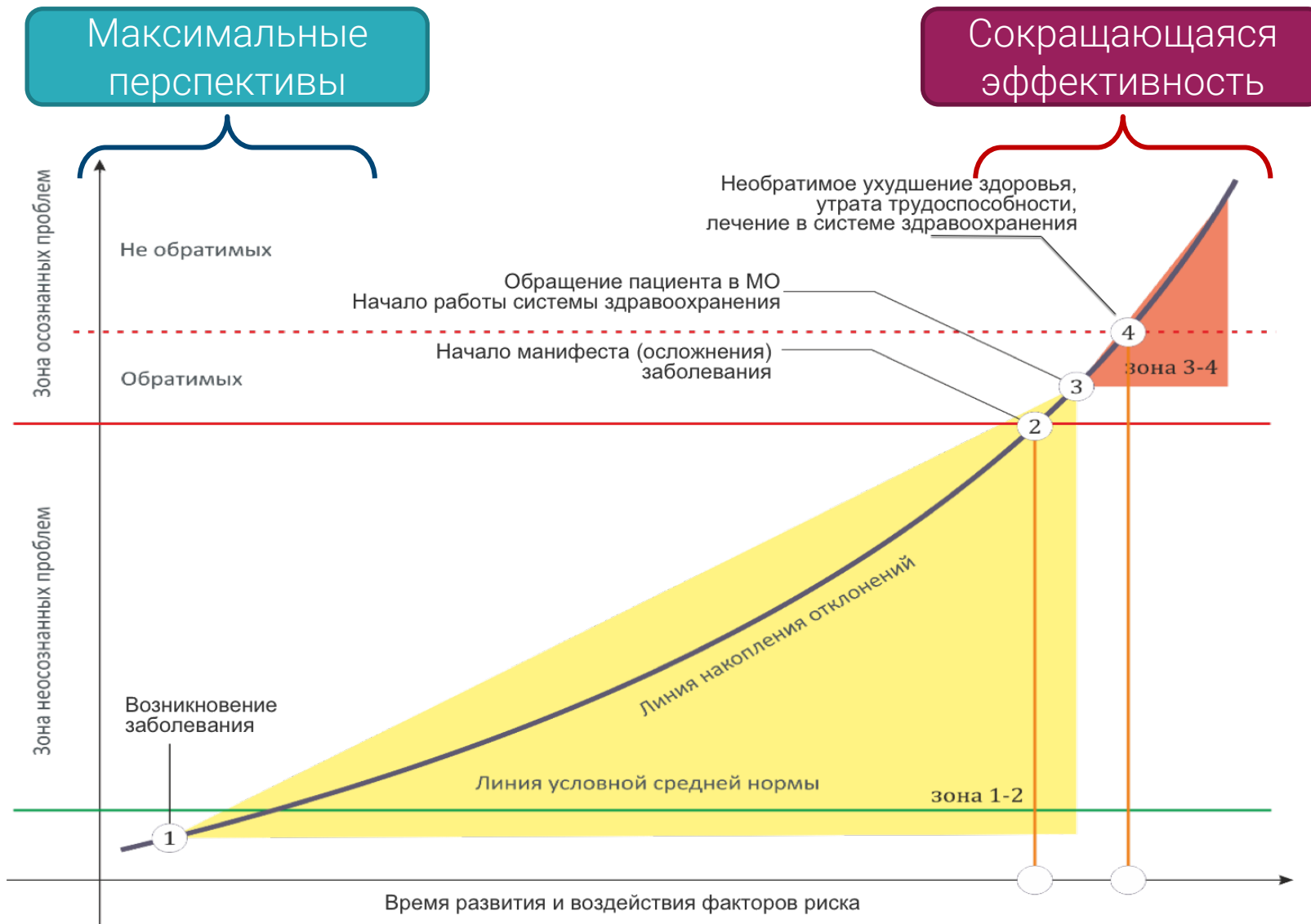
Фармацевтические компании



Технологические компании

# Самое интересное: ранний (бессимптомный) этап

15



Очень интересно предсказать не смерть пациента, у которого уже есть ССЗ, а начало развития заболевания на его бессимптомной фазе.

Что это дает:

1. Профилактика развития заболеваний, которая требует в 8 раз меньше затрат, чем лечение
2. Сокращение до 80% распространенности ХНИЗ



## АНАЛИЗ ОБЕЗЛИЧЕННОЙ ЭЛЕКТРОННОЙ МЕДИЦИНСКОЙ КАРТЫ

Автоматический анализ медицинских данных, включая извлечение информации из неструктурированных врачебных записей с помощью NLP-технологий



## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Анализ больших данных и машинное обучение для создания прогнозных моделей на основании информации из ЭМК. Поддержка принятия управленческих и врачебных решений на основе персональной оценки риска и точных прогнозов возможного ухудшения здоровья пациента в будущем



## РЕКОМЕНДАЦИИ ВРАЧУ И ПАЦИЕНТУ

Персональные советы врачу и пациенту по профилактике заболеваний, сформированные на основе утвержденных клинических рекомендаций





## Прогнозная аналитика для руководителя в сфере здравоохранения

- ✓ Увидеть будущую заболеваемость и смертность в вашем регионе / медицинской организации на ваших данных. Возможность заранее выделить самые проблемные места и отработать на опережение
- ✓ Увидеть спрос на лекарства и изделия медицинского назначения и сократить неэффективные затраты на ненужные закупки или сбои в поставках в случае недостатка лекарств/изделий

## Прогнозная аналитика для врачей

- ✓ Увидеть точный прогноз возможных ухудшений здоровья пациента в ближайшем будущем и предпринять точечные меры
- ✓ Увидеть прогноз возможной диагностической находки и назначить то, что вероятнее всего даст результат
- ✓ Обратить внимание на пациента, нуждающегося в дополнительном обследовании и/или лечении



<https://webiomed.ai/>



185031, РФ, Республика Карелия,  
г.Петрозаводск,  
набережная Варкауса, д. 17



+7 (814-2) 28-08-18



[info@webiomed.ai](mailto:info@webiomed.ai)



ВКонтакте

<https://vk.com/webiomed>



Facebook

<https://www.facebook.com/webiomed/>



Twitter

<https://twitter.com/webiomed>



Telegram

<https://t.me/webiomed>



YouTube

<https://www.youtube.com/>



Спасибо за внимание!