

## Реальные данные встречаются с машинным обучением: что нужно для обеспечения актуальности генерации доказательств

**Источник:** Arbiter

**Оригинал:** <https://medcitynews.com/2026/03/real-world-evidence-meets-machine-learning-what-it-takes-to-future-proof-evidence-generation/>

FDA

клинические исследования

машинное обучение

реальные данные

регуляторика

фармацевтика

Реальные данные (RWE) стали важным драйвером фармацевтической и биотехнологической отраслей в последние годы. Фактически, RWE теперь включено в 70 процентов новых регуляторных заявок на лекарственные препараты и биологические препараты в Управление по санитарному надзору за качеством пищевых продуктов и медикаментов США (FDA).

По мере того как RWE продолжает стимулировать прорывы, машинное обучение (ML) стало необходимым инструментом для обеспечения устойчивости генерации доказательств. Организации в сфере наук о жизни теперь имеют доступ к большему количеству реальных данных (RWD), чем когда-либо прежде, что делает ручной, управляемый человеком подход к анализу через скрипты SAS и разовые запросы нереалистичным и неустойчивым.

Использование моделей машинного обучения — наиболее эффективный способ обработки и осмысления RWD в масштабе. Это позволяет организациям в сфере наук о жизни анализировать больше данных из

большого количества источников, проводить постоянные анализы и выполнять более сложные анализы с помощью высокопроизводительных вычислений (HPC) и облачных вычислений.

Механика более связанной экосистемы здравоохранения [Видео]

Анджали Джеймсон из Arbiter о согласовании между больницами и плательщиками.

Регуляторные органы обратили на это внимание. FDA недавно выпустило руководство по использованию искусственного интеллекта (ИИ) в поддержку регуляторных одобрений и обновило дискуссионный документ в начале этого года, чтобы охватить роль ИИ и ML в жизненном цикле разработки лекарств.

Ничего удивительного, что регуляторы уделяют этому пристальное внимание — обещание RWE в сочетании с ML огромно. С RWE на базе ML ученые могут создавать синтетические контрольные группы, которые ускоряют сроки исследований, значительно снижают затраты и сокращают потребность в наборе пациентов на 20-50 процентов.

Кроме того, модели ML могут выявлять субпопуляции пациентов, которые с наибольшей вероятностью получают пользу от определенных методов лечения, или тех, кто находится в группе наивысшего риска неблагоприятных событий. Они также могут предсказывать, как определенные пациенты отреагируют на различные терапии на основе их уникальных клинических и геномных профилей. Кроме того, модели ML могут постоянно просматривать систему FDA Sentinel для выявления потенциальных неблагоприятных событий экспоненциально быстрее, чем при ручных запросах.

Каждая организация в сфере наук о жизни должна рассмотреть, как лучше всего применять ML к RWE для улучшения результатов лечения пациентов. Вот как они могут обеспечить, чтобы RWE была «готовой к ML» и внедрила ML для использования в средах, соответствующих требованиям GxP.

За пределами аналитики: как компания Dorsey закладывает ценность в политику Medicaid [Видео]

Как превратить аналитику в реальные результаты политики.

Получение RWE, готовой к ML

Чтобы сделать RWE готовой к ML, организации в сфере наук о жизни должны относиться к ней как к продукту: планировать ее, развертывать и предпринимать соответствующие шаги для обеспечения «удовлетворенности клиентов».

Вот как это выглядит:

### 1. Планирование

Организации должны создать четкую структуру управления данными для RWD, которая определяет владение, контроль доступа и политики использования. Это способствует доверию и помогает поддерживать соответствие таким регуляторным требованиям, как HIPAA и GDPR. Они также должны создать инфраструктуру данных, определив общие модели данных (например, OMOP, FHIR, SDTM) и структуры качества данных (например, полнота, соответствие, правдоподобие, своевременность).

### 1. Развертывание

Для успешного развертывания RWE организациям нужна среда, которая позволяет проводить анализ правильным образом (более подробно об инструментах, которые это обеспечивают, ниже). Прежде всего, эта среда должна быть стандартизирована для запуска валидации. Она должна поддерживать воспроизводимость и иметь встроенные протоколы соответствия и управления. Важно, что она также должна легко подключаться к RWD и преобразовывать их для эффективного и надежного анализа, не ставя под угрозу целостность данных или регуляторное соответствие. Наконец, эта среда должна обеспечивать бесшовное сотрудничество между заинтересованными сторонами (например, отделом обеспечения качества, наукой о данных и клиническими командами) в процессе разработки.

### 1. Обеспечение «удовлетворенности клиентов»

Обеспечение «удовлетворенности клиентов» в этом контексте означает, что среда должна быть построена не только для ученых по данным, но и для команд RWE и бизнеса, которые будут принимать важные бизнес- и клинические решения на основе моделей ML. Эти команды будут нуждаться в механизмах для просмотра моделей и их результатов, предоставления инсайтов и обратной связи, а также управления моделями для предотвращения дрейфа, смещения и т.д., поскольку данные, питающие их, неизбежно эволюционируют со временем.

## Внедрение ML для RWE

Исторически внедрение ML для RWE вызывало проблемы, когда организации в сфере наук о жизни пытались использовать устаревшие подходы в средах GxP. Это связано с тем, что команды часто работают в невалидированных средах с неконтролируемыми кодовыми базами и пытаются «обернуть» валидацию вокруг финального результата постфактум. Этот подход не выдерживает стандартов GxP и не обеспечивает воспроизводимость.

Сегодня платформы науки о данных, готовые к GxP, имеют соответствие и управление, встроенные непосредственно в саму платформу для удовлетворения регуляторных ожиданий. Они поддерживают воспроизводимость, управление и аудит, а также мониторинг и управление моделями, чтобы организации в сфере наук о жизни могли выдержать регуляторный надзор. Эти платформы поддерживают весь жизненный цикл модели, от исследовательского анализа до развертывания и мониторинга. Они также могут подключаться к внешним источникам данных, таким как API, без необходимости в таком большом количестве инженерии конвейеров, как при традиционном анализе.

Кроме того, они могут создавать визуализации или приложения, которые позволяют бизнес-пользователям видеть и понимать инсайты, которые извлекают их модели ML. Результат? Организации в сфере наук о жизни могут безопасно и эффективно применять ML к RWE для получения более глубоких инсайтов при сохранении соответствия и контроля над своими моделями.

RWE на базе ML больше не просто интересный эксперимент — это необходимость, движущая критическими достижениями. Но прежде чем организации в сфере наук о жизни смогут использовать всю мощь ML, они должны сделать свою RWE готовой к ML и принять инструменты, поддерживающие соответствие и воспроизводимость в средах GxP. Приоритизируя эти инициативы, они могут обеспечить устойчивость генерации доказательств и проложить путь для следующего поколения инноваций, движимых ML, в сфере наук о жизни.

Фото: claudenakagawa, Getty Images

Кристофер МакСпиритт — вице-президент по стратегии в сфере наук о жизни в Domino Data Lab. Он способствует пониманию потребностей клиентов и работает с командами управления продуктом и маркетинга для продвижения подходов к выводу на рынок в секторе наук о жизни.

Кристофер начал фокусироваться на индустрии наук о жизни, когда присоединился к небольшой eClinical стартап в 2005 году. С тех пор у него была возможность работать как в консалтинговых фирмах, так и в ведущих программных компаниях в качестве менеджера проектов, бизнес-аналитика, консультанта, менеджера продукта и стратега.

Этот пост появляется через программу MedCity Influencers. Любой может опубликовать свою точку зрения на бизнес и инновации в здравоохранении на MedCity News через MedCity Influencers. Нажмите здесь, чтобы узнать, как.

---

---

Перевод выполнен: 20.03.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.