Благодаря тому, что система разработана с использованием технологии TensorFlow, распознавание может быть запущено на маломощных системах (Android-смартфоны, Raspberry pi и т.д.).

На экспериментальных данных проверена точность для нейронных сетей для задач локализации – 91% и классификации – 93,2%. Описана работа программного обеспечения по распознаванию дорожных знаков.

Потенциальная область практического применения данной работы довольно широка: может быть встроена в систему автопилота автомобиля, может применяться в системах построения автодорожных карт для автоматического внесения в базу данных обозначения и местонахождения дорожных знаков или выступать в качестве ассистента водителю автотранспорта.

1. Tensorflow documentation [Электронный ресурс]. Режим доступа - <https://www.tensorflow.org/guide/keras?hl=ru>
2. Ren, S. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks/ S. Ren, K. He, R. Girshick, J. Sun // Advances in Neural Information Processing Systems. Vol. 39. 2015, P.1137–1149.
3. Szegedy, C. Rethinking the inception architecture for computer vision/ C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna// IEEE Conference on Computer Vision and Pattern Recognition. Vol.1. 2016, P.2818–2826.
4. Liu, W. SSD: Single shot multibox detector/ D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg // European Conference on Computer Vision 2016. Vol. 1. 2016, P.21–37.
5. Howard A.G. Mobilenets: Efficient convolutional neural networks for mobile vision applications./ A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam // arXiv 2017, arXiv:1704.04861.
6. Huang J. Speed/accuracy trade-offs for modern convolutional object detectors / J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama // 30th IEEE Conference on Computer Vision and Pattern Recognition. arXiv 2017, arXiv:1611.10012v3.
7. Шахуро, В.И., Конушин А.С. Российская база изображений автодорожных знаков // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 294-300.

Корнюшко В.Ф., Николаева О.М., Золотов С.Ю., Крюкова О.В., Садеков Л.В.
Интеллектуальные информационные технологии в экспериментальных
исследованиях разработки готовых лекарственных форм (ГЛФ) при производстве
лекарственных средств

МИРЭА – Российский технологический университет
(Россия, Москва)

doi: 10.18411/sr-10-02-2021-22

idsp: sciencerussia-10-02-2021-22

Аннотация

В статье рассматривается проблема применения интеллектуальных информационных систем при экспериментальных исследованиях и разработке готовых лекарственных форм при производстве лекарственных средств в твердой форме. Проведён информационный анализ жизненного цикла разработки и производства лекарственных средств и дана краткая характеристика интеллектуальных информационных систем, применяемых на различных этапах жизненного цикла. Для этапа разработки готовой лекарственной формы приведены информационная модель в нотации IDEF0 и схема интеллектуальной информационной системы. Дана краткая характеристика функциональных модулей системы. Подробно рассмотрено назначение и функционал базы данных вспомогательных и лекарственных веществ. Приведена математическая модель выбора технологической платформы разработки ГЛФ на основе

многооткликовой регрессионной модели и онлайн алгоритмов оптимизации. Приведен пример тестирования разработанной интеллектуальной информационной системы при разработке ГЛФ димедрола.

Ключевые слова: Интеллектуальные информационные системы, готовая лекарственная форма, жизненный цикл разработки и производства лекарственных средств, информационная модель в нотации IDEF0, база данных лекарственных и вспомогательных веществ, многооткликовые регрессионные модели, алгоритм online оптимизации, лекарственный препарат димедрол.

Abstract

This article discusses the problem of using intelligent information systems in experimental research and development of ready-made medicines (GLF) in the production of drugs in solid form. An informational analysis of drug development and production is carried out, and a brief description of intelligent information systems used at various stages of the life cycle is given. For the development stage of the ready-made medicine, an informational model in IDEF0 notation and an intelligent information system diagram are given. A brief description of the function modules of the system is given. The purpose and functionality of the database of excipients and medicinal substances are considered in detail. A mathematical model for choosing a technological platform for the developments of the GLF based on a multi-response regression model and online optimization algorithms is presented. An example of testing the development intelligent information system during the development of the GLF of diphenhydramine is given.

Keywords: intelligent information systems, ready-made drug form, life cycle of drug development and production, information model in IDEF0 notation, database of medicinal and auxiliary substances, multi-response regression models, online optimization algorithm, the drug diphenhydramine.

Введение. В последние годы проблемы и задачи, которыми многие годы занимались в основном IT- специалисты вошли в число первоочередных проблем развития страны и нашли свое формальное отражение в виде директивных документов: Национальный проект «Цифровая экономика»; Указ Президента РФ № 490 от 10 октября 2019г. "О развитии искусственного интеллекта в Российской Федерации"; План фундаментальных исследований РАН на период до 2025 года: «4. Информатика: Разработка фундаментальных проблем искусственного интеллекта, распознавания образов, оптимизации, проблемно-ориентированных систем и экспертных систем, основанных на знаниях».

Таким образом проблемы цифровизации и применения систем искусственного интеллекта стали важнейшими во многих отраслях народного хозяйства, в том числе в фармацевтической отрасли. При этом актуальность применения этих направлений при разработке и производстве лекарственных средств многократно возросла в последнее время в связи с пандемией COVID-19. Глобальными проблемами, вставшими перед разработчиками лекарственных средств для вакцинирования и лечения при COVID-19, стали задачи резкого снижения общего времени производства ЛС, всемерного повышения их эффективности и снижения риска побочных явлений. Дело в том, что классический жизненный цикл разработки и производства ЛС занимает 12-15 лет (Рис.1), из них большая часть времени идет на 3 фазы клинических исследований, во время которых проверяются эффективность разработанного ЛС и его риск к возможному появлению побочных эффектов. Жесткая международная стандартизация почти всех этапов жизненного цикла, их цифровизация и применение интеллектуальных информационных систем, позволяют значительно сократить общее время разработки и производства лекарственного средства (ЛС).

Рисунок 1. Схема системы контроля качества, эффективности и безопасности медицинской продукции во время жизненного цикла создания лекарственного средства

Здесь: GLP (Good Laboratory Practice)- надлежащая лабораторная практика; GCP (Good Clinical Practice)-надлежащая клиническая практика; GMP (Good Manufacturing Practice)- надлежащая производственная практика; GSP (Good Storage Practice) – надлежащая практика хранения; GDP (Good Distribution) – надлежащая дистрибьютерская практика [1].

Конечно, глобальной задачей в век цифровизации явилась бы проблема создания единой цифровой среды жизненного цикла разработки и производства лекарственных препаратов. Но очевидно, что это задача скорее всего очень далекого будущего. Так как на сегодня весь жизненный цикл очень сильно распределен во времени и пространстве. Нормальное, не экстремальное, время жизненного цикла разработки и производства лекарственного препарата составляет 12-15 лет. В работе участвуют сотни, а порой тысячи специалистов самых разных областей знаний медики, биологи, химики, технологи, математики, специалисты по информатике и т.д. Производство разработанного лекарственного средства может находиться не только в разных городах, но и разных странах. Так, например, вакцина Спутник V, разработанная в России, будет производиться на промышленных площадках Аргентины, Венгрии, Казахстана и других стран, с использованием промышленных мощностей этих стран. Поэтому сегодня выбрано направление, не глобальной цифровизации всего жизненного цикла, а цифровизация и применение интеллектуальных информационных систем управления на его отдельных этапах. Методическую основу этого составляют Стандартизация этапов жизненного цикла производства лекарственных средств и внедрение в последние годы методологии QbD – «качество, запланированное при разработке» [2]. Внедрение методологии QbD на всех этапах жизненного цикла теоретически позволит провести разработку и организовать промышленное производство лекарственного средства в минимально допустимые сроки и на основе оптимального критерия разработки – “Максимальная эффективность применения и минимальный риск от побочных эффектов”.

Следует отметить, что сегодня интеллектуальные информационные технологии весьма успешно применяются на большинстве этапов жизненного цикла. Так, крупные фармацевтические фирмы используют в системах управления технологиями производства и логистического обращения лекарственных препаратов (доставка, хранение, дистрибуция) глобальные информационные системы типа ERP (Enterprise Resource Planning) – это, информационные системы, которая позволяет собирать, хранить, а также обрабатывать огромное количество информации, необходимой для

штатного функционирования фармацевтического предприятия. Так, ERP "1С: Управление аптечной сетью" [3], реализованное на основе типовой информационной системы 1С – отраслевое решение, которое предназначено для управления торговой деятельностью в компаниях, специализирующихся на оптово-розничной торговле лекарственными средствами. Система реализует организует информационную поддержку таким логистическим функциям, как управление закупками, управление запасами и складским хозяйством, управление продажами, маркетинг и т.д.

На этапе клинических исследований, где речь идет об испытаниях разрабатываемого лекарственного средства на десятках тысяч людей, широко применяются популярные в настоящее время облачные сервисы Интернета интеллектуального анализа, прогнозирования и кластеризации данных "Data Mining" [4].

На этапе скрининга широко используются системы искусственного интеллекта.

Скринингом называется оптимизированная процедура, в результате которой большое количество химических соединений (> 10000) проверяется на активность по отношению к специальной тестовой (имитирующей биологическую) системе. В основе компьютеризированного скрининга лежат компьютерные программы, где вещества и белки представлены описанием их структур в определенном формате (обычно применяемые для этого базы данных содержат миллионы соединений).

Этот этап осуществляется с помощью высокопроизводительного скрининга (*in vitro*) или его компьютерного (*in silico*) анализа – высокопроизводительного докинга. В результате высокопроизводительного скрининга обычно просматривается более 1000000 образцов. По данным прессы в США при синтезе вакцины от COVID-19 с помощью суперкомпьютеров было проанализировано более 500 млн соединений. Здесь хочется отметить, что, используя естественный интеллект, наряду с искусственным, опыт, накопленный при разработке и массовом применении вакцины от лихорадки Эбола, первыми успеха добились ученые нашей страны при регистрации вакцины СПУТНИК V, разработанной в ФГБУ НИЦЭМ им. Н.Ф. Гамалеи Минздрава России.

Основная часть. Этапом, который в настоящее время в наименьшей степени, по сравнению с остальными, оснащен интеллектуальными системами поддержки экспериментальных исследований, является этап фармацевтической разработки лекарственных средств. Этот этап состоит из двух достаточно автономных подэтапов. Это синтез активной фармацевтической субстанции (АФС), целью которого является получение лидерного (лекарственного) вещества. На 2-м подэтапе получают готовую лекарственную форму (ГЛФ). Кроме этого, именно на 2-м подэтапе происходит выбор технологической платформы получения ЛС, выбор технологического и контрольно-измерительного оборудования. На Рис. 2 на диаграмме в нотации *IDEF0* приведена информационная модель разработки ГЛФ [5].

Рисунок 2. Диаграмма информационной модели в нотации IDEF0 фармацевтической разработки готовых лекарственных форм.

Для управления исследованиями на этом этапе предлагается интеллектуальная система информационной поддержки экспериментальных исследований. Система включает в себя несколько баз данных, ряд подсистем подготовки управленческих решений, подсистемы моделирования и оптимизации и экспертного оценивания качества получаемого продукта.

Подсистема (П/С) контроля подготовки производства, по существу, представляет собой формализованный набор инструкций по подготовке оборудования, воздуха, воды и персонала (Важно отметить, что все работы выполняются в изолированных помещениях класса D).

Рисунок 3. Схема интеллектуальной информационной системы управления экспериментальными исследованиями при разработке ГЛФ

Базы экспериментальных данных и данных Фармакопея XIII предназначены для решения следующих задач:

1. Ввод и хранение исходных данных:

- Данные от всех методов контроля для каждой стадии проводимых исследований с учетом возможной вариативности применяемых методов;
- Данные Государственной фармакопеи для оценки полученных результатов контроля. При организации постадийного контроля разработки твердых лекарственных форм применяются методы биологического, химического и технологического контроля.
- Данные об использованных методах контроля;
- Информацию о времени, месте и контактные данные исследователей, осуществлявших контроль, и аналитиков, обрабатывавших данные.

Базы данных лекарственных и вспомогательных веществ включают набор лекарственных веществ, с которыми работает данная лаборатория или центр, и

вспомогательные вещества. Вспомогательные вещества служат для придания разрабатываемым твердым ЛП определенных свойств и подбираются исследователем [6]. Наполнение базы вспомогательных веществ сильно отличается в различных лабораториях, но в общем виде эта база должна включать наполнители (типа крахмала, глюкозы, сахарозы, лактозы и др); связующие и разрыхляющие вещества; корригенты, отвечающие за вкус, запах, цвет и т.д.

Базы данных лекарственных и вспомогательных веществ непрерывно пополняются и развиваются за счет как новых активных молекул, так и необходимых для них вспомогательных веществ.

Остановимся подробнее на построении базы данных вспомогательных веществ. В процессе их подбора исследователю необходимо производить поиск фармацевтических субстанций и вспомогательных веществ по ряду критериев, таких как класс, способ применения и лекарственная форма, для получения более подробной информации о веществе и поставщиках, а также сохранять результаты оценки показателей контроля после исследования, и добавлять новые вещества.

В тексте словесного описания мы можем выделить основные понятия предметной области, которые представляют собой независимые сущности.

Сущность оценки связывает показатель контроля и вспомогательное вещество. Эти характеристики далее были использованы для построения модели “сущность-связь” и ее нормализации при проектировании реляционной модели.

В основу построения модели положены следующие соображения:

- Каждое вспомогательное вещество может относиться к нескольким классам;
- Каждая лекарственная форма имеет один способ применения, но каждое вспомогательное вещество или фармацевтическая субстанция могут применяться несколькими способами;
- Фармацевтические субстанции имеют множество возможных синонимов;
- Как фармацевтические субстанции, так и вспомогательные вещества могут иметь много разных поставщиков.

Отсюда вытекают основные сущности, положенные в основу модели “сущность-связь” (рис. 4): вспомогательное вещество; фармацевтическая субстанция; лекарственная форма; способ применения; класс вспомогательного вещества; поставщик; синоним; показатель контроля.

Рисунок 4. Модель «Сущность-связь» базы данных вспомогательных веществ при разработке ГЛФ лекарственных соединений.

Выше было отмечено, что на подэтапе фармацевтической разработки состава ГЛФ производится выбор технологической платформы, включая выбор оптимальной технологии, контрольно-измерительной аппаратуры, алгоритмов моделирования и оптимизации. Ключевым переделом в этой технологии является процесс прессования(таблетирования), схема которого зависит от выбора вспомогательного вещества. Обычно процесс прессования предусматривает наличие компьютерной системы управления, алгоритмы для которой разрабатываются исследователем.

П/с цифровизации оперативного управления (маршрутные листы)

Задачей этой подсистемы является цифровая обработка и хранение всей оперативной информации об управлении процессом разработки ГЛФ, выполняемой исследователем. Эта информация записывается в маршрутные листы и используется в дальнейшем не только самими исследователями, но и экспертами для контроля за соблюдением требований Стандартов и Фармакопеи и необходима инспектирующим органам в случае выхода разрабатываемого ЛП за допустимые требования на стадии клинических испытаний (иногда, не смотря на затраченные годы работы и огромные средства, при нарушении условий эффективности или появления неучтенных побочных эффектов весь процесс разработки приходится начинать заново).

П/с экспертного оценивания служит для информационной и алгоритмической поддержки экспертного оценивания процесса разработки ГЛФ. Методические основы алгоритма работы экспертов заложены в **приказе Министерства здравоохранения РФ от 24 августа 2017 г. N 558н "Об утверждении Правил проведения экспертизы лекарственных средств для медицинского применения, форм заключений комиссии экспертов"**

Ключевым переделом процесса разработки ГЛФ, на который влияет выбор вспомогательного вещества и дальнейшее формирование технологической платформы, является процесс прессования(таблетирования), управление которым осуществляется с помощью компьютерной системы.

Для моделирования процесса прессования был применен регрессионный анализ и многооткликовые уравнения регрессии [7]. Общее уравнение решения задачи оптимизации имеет следующий вид:

$$G = b_1 * x_1 + b_2 * x_2 + \dots + b_m * x_m \rightarrow \text{optim} \quad (2)$$

$$Y_j^k = Y_j(x_1, x_2 \dots x_m) \leq \text{abs}(Y_{\text{доп}}^k) \quad (3)$$

где **G** – многокритериальный показатель качества полученной смеси, определяемый стандартом GMP, а **x1, x2,..., xm** - управляемые параметры процесса прессования (давление, время прессования и др.), а также параметры смеси, зависящие от ее предварительной обработки и подбора соответствующего вспомогательного вещества.

При этом контролируемые параметры таблеточной смеси, такие как размер частиц и степень их однородности, текучести смеси и т.д., определяемые на основе химического и технологического контроля, имеют вид ограничений (3), которые находятся по данным химического и технологического контроля и ограничены нормативными требованиями $Y_{\text{доп}}^k$.

В работе выбран двухступенчатый подход к решению задачи оптимизации с помощью математической модели: на первом - выполняется пошаговое движение к оптимуму, на втором - для найденного плана управляемых переменных выполняется проверка допустимых значений состава. Такой подход позволяет экономить время и ресурсы за счет отказа от таких планов, при которых возможно нарушение ограничений по составу. Для решения задачи оптимизации выбран один из методов on-line оптимизации - модифицированный метод Хука-Дживса.

Замечания:

1. Предлагаемый алгоритм предполагает, что для каждого выбранного вспомогательного вещества производится адаптация математических моделей (2,3);

2. В случае, если не удастся найти решение системы (2,3) для выбранного вспомогательного вещества (ВВ), то оно заменяется на другое из базы данных ВВ;

3. Окончательное решение о выборе подходящего вспомогательного вещества и настройке технологической платформы принимается группой экспертов.

Апробация разработанных алгоритмов разработке ГЛФ была проведена на тестовом примере хорошо известного лекарственного средства димедрола [8].

Для разработки состава при получении ГЛФ димедрола на основе технического задания, подготовленного фармакологом и с помощью интеллектуальной системы управления, были подобраны вспомогательные вещества, а также выбрана технология прессования твердой лекарственной формы.

Все составы, разработанные в процессе исследования приведены в таблице.

Таблица 1.

Составы ГЛФ димедрола, разработанные с применением алгоритма

Название вещества	1 состав	2 состав	3 состав	4 состав	5 состав
Димедрол	50	50	50	50	50
Лактоза	44	48	75	40	-
Тальк	0,5	-	-	-	-
Крахмал картофельный	25	10	-	-	25
Стеарат кальция	0,5	1,0	0,7	2	-
МКЦ	30	10	0,6	13	50
Аэросил	-	21,0	0,7	15	2
Крахмал кукурузный	-	10	23	30	23
Масса ГЛФ	150 мг	150 мг	150 мг	150 мг	150 мг

Составы были проверены в соответствии с фармакопейной статьей ОФС.1.4.1.0015.15 таблетки. В результате испытаний состав №3 оказался оптимальным.

Заключение

1. В статье рассмотрены некоторые особенности построения интеллектуальных информационных систем для отдельных этапов жизненного цикла разработки и производства лекарственных средств.

2. Рассмотрена структура интеллектуальной информационной системы разработки ГЛФ и приведены функционалы отдельных модулей системы.

3. Подробно рассмотрена методика формирования базы данных вспомогательных веществ.

4. Приведены многооткликовые регрессионные модели и алгоритмы онлайн оптимизации для управления исследованиями при разработке ГЛФ с возможной вариативностью вспомогательных веществ.

5. Проведена апробация разработанного алгоритмического и информационного обеспечения на тестовом примере разработки известного лекарственного средства димедрола, полученные результаты получили положительную оценку экспертов.

1. Быков В.А., Береговых В.В., Швец В.И., Самылина И.А., Пятигорская Н.В., Мешковский А.П., Топников И.В. О специальности «Промышленная фармация» // Фармация. 2007. № 1. С. 45–47.
2. Энде Д. Дж. Производство лекарственных средств. Химическая технология от R&D до производства: пер. с англ. СПб.: ЦОП «Профессия», 2015. 1280 с.
3. Отраслевые и специализированные решения. 1С: Предприятие. URL: <https://solutions.1c.ru>.
4. Баргесян А.А. Анализ данных и процессов. Учебное пособие-СПб: БХВ-Петербург, 2009, 512 с.
5. Николаева О.М., Корнюшко В.Ф., Панов А.В., Богунова И.В., Флид А.А., Применение системного подхода для построения информационного пространства вокруг фармацевтической разработки

-
- технологии производства готовых лекарственных средств // Прикладная информатика. – 2018. – Т. 13. № 3 (75). – С. 83-100.
6. V.N. Taptunov, E.V. Guseva, S.E. Batin, N.V. Menshutina, D.J. Zhukov, A.V. Matasov. Information intellectual system for conceptual design of technological schemes for solid dosage forms production. // 19th International Congress of Chemical and Process 33 Engineering (CHISA'2010). – Prague, Czech Republic, 2010 – P.132.
 7. Корнюшко В. Ф., Николаева О.М., Богунова И.В., Кузнецов А.С., Панов А.В. Программно-аналитическая поддержка интеллектуальной системы управления синтезом активных фармацевтических субстанций, Программные продукты и системы // Software & Systems. 2020. – Т.33. № 1. – С. 561-566.
 8. Туешов В.И., Гладух Е.В., Ляпунова О.А., Сайко И.В., Сичкарь А.А., Рубан Е.А., Крутских Т.В. Промышленная технология лекарств. 2010. URL: <http://ztl.nuph.edu.ua/html/medication/index.html>.
-