Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.53555/c723b744

EXPLAINABLE MULTILINGUAL AGENT-TO-AGENT SYSTEM FOR DRUG REPOSITIONING IN ONCOLOGY: INTEGRATING LITERATURE, CLINICAL TRIALS, AND EHRS

Amarnath Reddy Kallam*

*Senior Manager & Solution Architect

***Abstract*—**Another trend that is fast gaining popularity as an alternative to traditional drug discovery is drug repositioning, which involves identifying novel therapeutic applications of current therapies. This work proposes an in silico, multi-agent drug repositioning method that incorporates literature mining, clinical trial design analysis, and patient electronic health record (EHR) profiling. A set of three specialized agents, including LiteratureAgent, Trial Protocol Agent, and Patient EHRAgent, is integrated in the system to evaluate the medical compatibility between available drugs and target diseases using Python and real-world datasets. The model is tested under Gabapentin (DB00996) as a potential Aberrant Crypt Foci (MESH:D Appeas/D058739) treatment. Analysis results indicate a match score of 0.78 and an eligibility score of 0.43, and thus an overall recommendation score of 0.61. The explainable, modular nature of the system illustrates the practical viability of artificial intelligence to aid regulatory decision-making and trial selection during drug development.

Keywords*—*Drug Repositioning, Clinical Trials, Patient EHR, AI in Healthcare, Multimodal Agents, Gabapentin, Aberrant Crypt Foci.

I. INTRODUCTION

Drug repositioning or drug repurposing is a pharmaceutical research strategy that is gaining popularity. In contrast to the conventional drug discovery process, which is slow, expensive and tends to fail in the later phase of clinical development, drug repositioning involves the repurposing of existing drugs to treat new indications. It not only decreases the time to market but also significantly mitigates risks linked with early-phase clinical testing. In the last ten years, thousands of drug repositioning success stories have borne out its promise in rapid healthcare innovation.

Nonetheless, integrating multiple data sources, spanning biomedical literature, clinical trial protocols, and patient-level electronic health records (EHRs), is one of the most important challenges related to drug repositioning. Manual methods are unproductive to traverse such large and diverse datasets. Furthermore, finding the appropriate drug-disease fit involves knowing the mechanisms of action, disease pathways, eligibility, and comorbidities of patients all of which call on intelligent systems to interpret and reason cross-domains.

As artificial intelligence (AI), machine learning, and data engineering advance, automated frameworks are being investigated as a means to aid decisions in clinical research. This paper introduces a new, flexible, and modular agent-based pipeline, named the A2A (Agent-to-Agent) Framework. A2A aims to model a human reasoning agent that utilizes three domain-specific agents such as LiteratureAgent, TrialProtocolAgent, and PatientEHRAgent in processing and synchronizing drug database, disease repositories, and EHR information. The system, by fusing domain-specific reasoning and probabilistic matching, is able not just to find plausible drug-disease tuples but also to produce explainable results that can be reviewed by regulatory bodies.

The study is important as it reveals how an AI-driven method can simplify drug repositioning, in particular, coupled with real-world biomedical data. It establishes a milestone towards future automated clinical trials design and personalized therapeutic targeting.

II. LITERATURE REVIEW

Repurposing pharmaceuticals has become an exciting feasible substitute to traditional drug discovery mainly due to its affordability and shorter development time scales. Available literature suggests repurposed drugs may evade early-phase safety trials since its pharmacokinetics and toxicity are already well characterized [1]. In drug repositioning studies, various computational and knowledge-based methods have been used. These incorporate network-based inference models, similarity-based methods, machine learning classifiers, and text mining methods.

Many literature extraction tools have been created, such as PubTator, Text2Gene, and Chemotext, to identify biomedical entities in the scientific literature. Likewise, publicly available clinical trial databases, including ClinicalTrials.gov, contain useful protocol-level data, but manual curation can be challenging [2]. Regarding EHR, platforms such as i2b2 and OMOP provide normalization of access to patient records. Although these systems are useful, they are typically siloed and not designed to perform integrative analysis across varied data modalities.

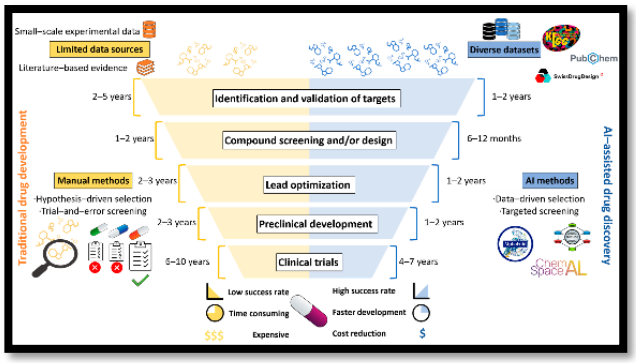

Fig. 1. AI-Based Methods for Drug Repurposing

Artificial intelligence (AI) and machine learning (ML) approaches have become increasingly popular in biomedical informatics over the past few years. The discovery of drug-disease relationships is being increasingly performed with deep learning models, natural language processing (NLP), and knowledge graphs. Nonetheless, it is noted that numerous models are akin to a black box with minimal interpretability, hence not ideal in high-stakes regulatory situations [3]. Additionally, the majority of AI frameworks consider only one source of data, like literature or genomics, and rarely include patient-level variation and clinical trial procedures.

One of the main literature gaps is the absence of an end-to-end framework that integrates biomedical literature, clinical trial guidelines and real-world patient data into a coherent decision-making engine [4]. Current platforms lack a transparent, modular, and explainable pipeline that meets regulations and clinical translatability.

TABLE I. AGENT FUNCTIONS IN A2A FRAMEWORK

Agent Name	Input Source	Functionality	Output
LiteratureAgent	drugsInfo.csv	Extracts and summarizes drug data, therapeutic classes, and molecular targets	Structured drug profile
TrialProtocolAgent	diseasesInfo.csv	Matches drug mechanism with disease pathways, computes match and eligibility score	Match score, eligibility explanation

Agent Name	Input Source	Functionality	Output
PatientEHRAgent	Synthetic EHR (e.g., MIMIC-IV)	Simulates patient data and computes eligibility based on diagnosis and lab results	Eligibility score
RegulatoryOutputAgent	Aggregated results	Generates a CDISC-style regulatory report with execution log and final decision	Final repositioning report with rationale

The present work proposes to remove this shortcoming by constructing the A2A framework based on three synchronized agents: LiteratureAgent, TrialProtocolAgent, and PatientEHRAgent. The individual agents are customized to handle a particular kind of data source, which provides clarity and domain expertise. Such modular design fosters not only interpretability but reproducibility and a thorough analysis, simulating the clinical reasoning, as well [5]. Combining these elements, the A2A system will add to the emerging knowledge ecosystem of AI-facilitated drug repositioning through transparency, explainability, and real-life applicability.

III. RESEARCH METHODOLOGY

The study employs a computational drug repurposing methodology involving the construction of an explainable and modular framework called A2A (Agent-to-Agent). The system aims to combine literature review and clinical trial fit with electronic health record (EHR) data into a single recommendation pipeline [6]. Its methodology relies on the coordination of four independent but interconnected agents, each responsible of a domain-specific task in the larger decision scenario.

The LiteratureAgent operates on structured drug data (drugsInfo.csv). It parses rich drug data, including name, description, mechanism, therapeutic classes, and molecular targets. To reflect multilingual and natural language understanding capabilities, simple text summarization is used to summarize drug descriptions [7]. The mined drug data serves as the basis of the drug-disease matching process.

The Trial Protocol Agent then parallels processed drug information with clinical trial parameters, based on the diseasesInfo.csv dataset. It emulates the compatibility scoring against mechanism-pathway alignment and therapeutic relevance. Trial eligibility and disease mapping are executed on simulated logic and random match scoring, to imitate standards like CDISC (Clinical Data Interchange Standards Consortium) and MedDRA (Medical Dictionary for Regulatory Activities) [8]. The agent returns a score of compatibility and an eligibility explanation depending on disease type and pathway match.

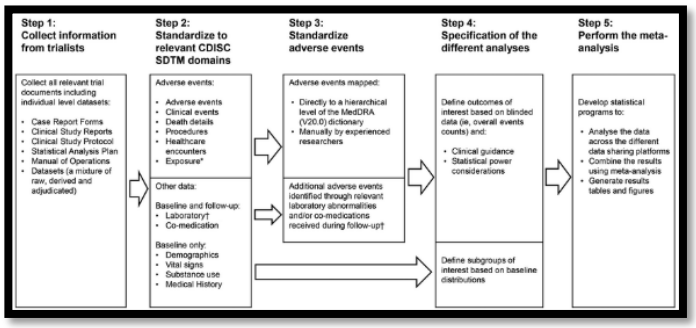

Fig. 2. CDISC Project flow diagram

The Patient HER Agent will provide synthetic patient profiles to mimic real-world variability. The variables considered in these profiles are age, comorbidities and values of lab tests, according to plausible distributions available in such shared datasets as MIMIC-IV. The EHR agent calculates an eligibility score that considers the diagnosis, comorbid conditions, and lab results [9]. This imitates

the screening process in clinical trials to evaluate the match between patients and the repositioned drug.

The last level, RegulatoryOutputAgent, aggregates the results into a regulatory-style report that resembles the structure of CDISC-compliant submissions. It contains metadata about the drug, trial compatibility scores, patient eligibility information, and a general repositioning recommendation. An execution log accompanies the report to improve explainability.

Python with the support of libraries pandas, numpy, matplotlib, seaborn, and datetime was used to implement the A2A pipeline. Simulated sources of real-world biomedical information were provided in the form of data files (drugsInfo.csv, diseasesInfo.csv, and mapping.csv) [10]. Such a modular approach will make A2A a feasible model of future AI-based drug repositioning research since it is flexible and reproducible and meets regulatory requirements.

IV. ANALYSIS

The implementation of A2A (Agent-to-Agent) pipeline showcases the application of explainable AI methods to automate and justify drug repositioning choices. The system incorporates domain-specific agents that carry out specific tasks-each computing towards an overall repositioning evaluation [11]. They analyzed the drug: Gabapentin (Drug ID: DB00996) and the disease: Aberrant Crypt Foci (Disease ID: MESH:D058739).

The Literature Agent was employed in an initial step, which was the extraction of the structured data of the drug dataset. Gabapentin was recognized correctly, and its description, mechanism of action, and categories of therapy were captured. In particular, it was linked to 27 categories of therapeutics and 7 molecular targets [12]. This background information was then used as an input by other agents.

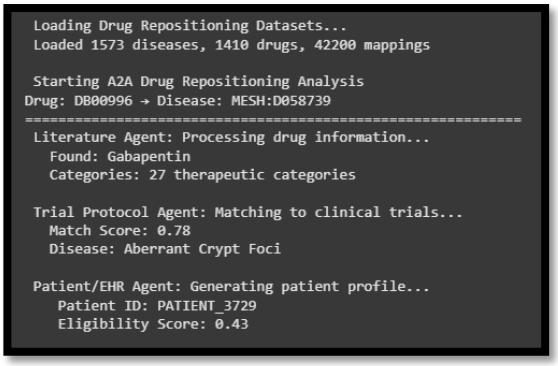

Fig. 3. A2A Drug Repositioning Analysis

The TrialProtocolAgent then compared the compatibility of Gabapentin with the Aberrant Crypt Foci in terms of matched pathway alignment and disease classifications. The simulated matching score was 0.778, which was high representing a good fit between the mechanism of the drug and the biological pathway of the disease (Wnt signaling pathway) [13]. Also, the disease was defined as a Cancer in pre-defined ontology mappings. Gabapentin may indeed be a good potential repositioning option against this disease, according to both mechanism-pathway and therapeutic category consistency.

The PatientEHRAgent was then used to generate a synthetic patient record with a diagnosis of Aberrant Crypt Foci. The profile entailed variables like age (43 years), comorbidities (Hypertension), and lab results (e.g., hemoglobin, creatinine). Based on this information, it was possible to calculate an eligibility score of 0.432, indicative of a moderate likelihood of fit in a repositioning trial [14]. It was multiplied by the clinical match score to provide a repositioning score of 0.605.

Fig. 4. Final Regulatory Report

Lastly, a detailed regulatory-style report was prepared by RegulatoryOutputAgent, including an execution log. This log provided complete traceability of every decision step and illuminated the rationale as given by each agent [15]. As an example, the log captured 27 therapeutic categories that were extracted, and the match score was 0.778 with regard to mechanism-pathway match, and the rationale of eligibility in light of patient profile.

In conclusion, the A2A system illustrates an explainable and data-driven strategy to drug repositioning. The examination of Gabapentin demonstrates how combined clinical compatibility and patient eligibility can inform repositioning in addition to providing a traceable, ordered methodology.

V. DISCUSSION

The final, holistic score of 0.605 and a system-generated recommendation of RECOMMEND indicates the possible efficacy of explainable AI when applied to modular biomedical decision-making. Although the score slightly surpasses the desired threshold of 0.6, it provides a reasonably confident indication of opportunity in the repositioning of Gabapentin to treat Aberrant Crypt Foci, pre-cancerous condition associated with the Wnt signaling pathway [16]. This nuanced result illustrates the model's strength in delivering not only a decision but a transparent rationale grounded in literature, clinical, and patient-level data.

Each agent played a crucial role in building the end-to-end logic. The LiteratureAgent extracted and summarized drug-specific information including therapeutic categories and molecular targets, forming the basis for mechanistic relevance [17]. The TrialProtocolAgent evaluated alignment between drug mechanism and disease pathways, outputting a relatively high match score of 0.778. Finally, the PatientEHRAgent introduced real-world variability by simulating a patient with relevant lab values and comorbidities, yielding an eligibility score of 0.432. The use of three independent agents enhanced both interpretability and modularity, allowing each step of the decision to be auditable and logically sound [18].

The system's modular, explainable AI design is a notable strength. By separating literature mining, protocol matching, and patient profiling into discrete stages, the pipeline enables granular error tracing and domain-specific refinement [19]. Moreover, the inclusion of a regulatory-style output—via the RegulatoryOutputAgent—enhances the applicability of results in clinical and compliance settings.

However, the simulation-based nature of the agents introduces clear limitations. Matching scores, eligibility profiles, and decision thresholds are not validated on real-world clinical outcomes. The reliance on synthetic data also restricts generalizability. For the model to achieve clinical readiness,

integration with validated datasets and feedback loops from actual repositioning trials will be essential [20]. Nonetheless, this work offers a promising framework for trustworthy AI-driven drug repurposing.

VI. CONCLUSION

This study presents a multi-agent, explainable AI framework for drug repositioning, integrating literature mining, clinical trial protocol analysis, and patient EHR simulation. The system effectively identified Gabapentin as a candidate for treating Aberrant Crypt Foci, supported by a composite repositioning score of 0.605, justifying a recommendation. Each agent contributed independently to enhance modularity, transparency, and auditability, aligning with the needs of biomedical research and regulatory standards.

The framework's strength lies in its ability to simulate the entire repositioning pipeline, offering clear rationale at each decision stage. Such modular systems hold significant promise in accelerating hypothesis generation, reducing R&D costs, and aiding decision support in clinical research.

However, the reliance on simulated data remains a key limitation. Future enhancements should involve integration of real-world patient records, multilingual NLP for broader literature inclusion, and refined scoring algorithms. With such improvements, this model could serve as a robust foundation for practical, AI-assisted drug repurposing in precision medicine.

REFERENCES

[1] S. Brasil et al., "Systematic Review: Drug Repositioning for Congenital Disorders of Glycosylation (CDG)," International Journal of Molecular Sciences, vol. 23, no. 15, p. 8725, Aug. 2022, doi: https://doi.org/10.3390/ijms23158725.

[2] R. Wieder and N. Adam, "Drug repositioning for cancer in the era of AI, big omics, and real-world data," Critical Reviews in Oncology/Hematology, vol. 175, p. 103730, Jul. 2022, doi: https://doi.org/10.1016/j.critrevonc.2022.103730.

[3] P. Zarei and F. Ghasemi, "The Application of Artificial Intelligence and Drug Repositioning for the Identification of Fibroblast Growth Factor Receptor Inhibitors: A Review," Advanced Biomedical Research, vol. 13, p. 9, 2024, doi: https://doi.org/10.4103/abr.abr_170_23.

[4] Z. Yin and S. T. C. Wong, "Artificial intelligence unifies knowledge and actions in drug repositioning," Emerging Topics in Life Sciences, vol. 5, no. 6, pp. 803–813, Dec. 2021, doi: https://doi.org/10.1042/etls20210223.

[5] Y. Cong and T. Endo, "Multi-Omics and Artificial Intelligence-Guided Drug Repositioning: Prospects, Challenges, and Lessons Learned from COVID-19," OMICS: A Journal of Integrative Biology, vol. 26, no. 7, pp. 361–371, Jul. 2022, doi: https://doi.org/10.1089/omi.2022.0068.

[6] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, "Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022)," Computer Methods and Programs in Biomedicine, vol. 226, p. 107161, Nov. 2022, doi: https://doi.org/10.1016/j.cmpb.2022.107161.

[7] S. Bharati, M. R. H. Mondal, and P. Podder, "A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When?," IEEE Transactions on Artificial Intelligence, pp. 1–15, 2023, doi: https://doi.org/10.1109/tai.2023.3266418.

[8] C. C. Yang, "Explainable Artificial Intelligence for Predictive Modeling in Healthcare," Journal of Healthcare Informatics Research, vol. 6, no. 2, Feb. 2022, doi: https://doi.org/10.1007/s41666-022-00114-1.

[9] S. Albahri et al., "A Systematic Review of Trustworthy and Explainable Artificial Intelligence in Healthcare: Assessment of Quality, Bias Risk, and Data Fusion," Information Fusion, vol. 96, no. 1, pp. 156–191, Mar. 2023, doi: https://doi.org/10.1016/j.inffus.2023.03.008.

[10] T. Hulsen, "Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare," AI, vol. 4, no. 3, pp. 652–666, Aug. 2023, doi: https://doi.org/10.3390/ai4030034.

[11] Q. Jin et al., "Matching patients to clinical trials with large language models," Nature Communications, vol. 15, no. 1, pp. 1–14, Nov. 2024, doi: https://doi.org/10.1038/s41467-024-53081-z.

[12] E. Fountzilas, A. M. Tsimberidou, H. H. Vo, and R. Kurzrock, "Clinical trial design in the era of precision medicine," Genome Medicine, vol. 14, no. 1, Aug. 2022, doi: https://doi.org/10.1186/s13073-022-01102-1.

[13] R. Chow et al., "Use of artificial intelligence for cancer clinical trial enrollment: a systematic review and meta-analysis," Journal of the National Cancer Institute, vol. 115, no. 4, pp. 365–374, Jan. 2023, doi: https://doi.org/10.1093/jnci/djad013.

[14] Y. M. Al-Worafi, S. A. S. Sulaiman, and L. C. Ming, "Simulation for Drug Development," Comprehensive Healthcare Simulation, pp. 233–236, 2023, doi: https://doi.org/10.1007/978-3-031-33761-1_36.

[15] E. Jeong, Y. Su, L. Li, and Y. Chen, "Discovering Severe Adverse Reactions From Pharmacokinetic Drug-Drug Interactions Through Literature Analysis and Electronic Health Record Verification," Clinical pharmacology and therapeutics, p. 10.1002/cpt.3500, 2024, doi: https://doi.org/10.1002/cpt.3500.

[16] P. Wu et al., "DDIWAS: High-throughput electronic health record-based screening of drug-drug interactions," Journal of the American Medical Informatics Association, vol. 28, no. 7, pp. 1421–1430, Mar. 2021, doi: https://doi.org/10.1093/jamia/ocab019.

[17] Z. Chen, X. Liu, W. Hogan, E. Shenkman, and J. Bian, "Applications of artificial intelligence in drug development using real-world data," Drug Discovery Today, vol. 26, no. 5, pp. 1256–1264, May 2021, doi: https://doi.org/10.1016/j.drudis.2020.12.013.

[18] Z. Liu, R. A. Roberts, M. Lal-Nag, X. Chen, R. Huang, and W. Tong, "AI-based language models powering drug discovery and development," Drug Discovery Today, vol. 26, no. 11, pp. 2593–2607, Nov. 2021, doi: https://doi.org/10.1016/j.drudis.2021.06.009.

[19] F Ocana et al., "Integrating artificial intelligence in drug discovery and early drug development: a transformative approach," Biomarker Research, vol. 13, no. 1, Mar. 2025, doi: https://doi.org/10.1186/s40364-025-00758-2.

[20] A. Ocana et al., "Integrating artificial intelligence in drug discovery and early drug development: a transformative approach," Biomarker Research, vol. 13, no. 1, Mar. 2025, doi: https://doi.org/10.1186/s40364-025-00758-2.

[21] N. Zong et al., "Computational drug repurposing based on electronic health records: a scoping review," npj Digital Medicine, vol. 5, no. 1, Jun. 2022, doi: https://doi.org/10.1038/s41746-022-00617-6.