Artificial Intelligence-based Efficacy Prediction of Phase 3 Clinical Trial for Repurposing Heart Failure Therapies

Nansu Zong, Ph.D. [1,§,*], Shaika Chowdhury, Ph.D. [1,§], Shibo Zhou, Ph.D. [1,§], Sivaraman Rajaganapathy[1], Ph.D. [1], Yue yu, Ph.D. [2], Liewei Wang, M.D., Ph.D. [3], Qiying Dai, MD[4], Suzette J. Bielinski, Ph.D., M.Ed. [2], Yongbin Chen, MD, Ph.D[5], James R. Cerhan, M.D., Ph.D.[2]

[1] Department of Artificial Intelligence and Informatics Research, Mayo Clinic, Rochester, MN, USA

[2] Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA

[3] Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA

[4] Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA

[5] Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, USA

[§]The authors contribute equally

[*] Corresponding author, Nansu Zong, Ph.D., Zong.Nansu@mayo.edu

Abstract

Introduction: *Drug repurposing involves finding new therapeutic uses for already approved drugs, which can save costs as their pharmacokinetics and pharmacodynamics are already known. Predicting efficacy based on clinical endpoints is valuable for designing phase 3 trials and making Go/No-Go decisions, given the potential for confounding effects in phase 2.*

Objectives: *This study aims to predict the efficacy of the repurposed Heart Failure (HF) drugs for the Phase 3 Clinical Trial.*

Methods: *Our study presents a comprehensive framework for predicting drug efficacy in phase 3 trials, which combines drug-target prediction using biomedical knowledgebases with statistical analysis of real-world data. We developed a novel drug-target prediction model that uses low-dimensional representations of drug chemical structures and gene sequences, and biomedical knowledgebase. Furthermore, we conducted statistical analyses of electronic health records to assess the effectiveness of repurposed drugs in relation to clinical measurements (e.g., NT-proBNP).*

Results: *We identified 24 repurposed drugs (9 with a positive effect and 15 with a non-positive) for heart failure from 266 phase 3 clinical trials. We used 25 genes related to heart failure for drug-target prediction, as well as electronic health records (EHR) from the Mayo Clinic for screening, which contained over 58,000 heart failure patients treated with various drugs and categorized by heart failure subtypes. Our proposed drug-target predictive model performed exceptionally well in all seven tests in the BETA benchmark compared to the six cutting-edge baseline methods (i.e., best performed in 266 out of 404 tasks). For the overall prediction of the 24 drugs, our model achieved an AUCROC of 82.59% and PRAUC (average precision) of 73.39%.*

Conclusion: *The study demonstrated exceptional results in predicting the efficacy of repurposed drugs for phase 3 clinical trials, highlighting the potential of this method to facilitate computational drug repurposing.*

1. Introduction

Drug repurposing is an innovative approach to drug discovery that involves finding new therapeutic uses for existing drugs. This approach has become increasingly popular and is currently used to develop 30% of newly approved drugs in the United States[1]. The appeal of drug repurposing is due to the reduced cost and time required for development[1], as safety studies have already been conducted[2,3], mitigating the risk of unexpected side effects. The concept of pharmacology also supports drug repurposing, as a single drug can act on multiple targets[4,5], allowing for the treatment of multiple diseases with one drug. Evaluating the potential efficacy of the repurposed drug candidate on patients is critical to maximizing the drug's therapeutic effects while minimizing any health hazards. Conventionally, phase 3 clinical trial tests the drug's efficacy by seeking improved therapeutic effects on a random subset of the cohort who were given the drug candidate as an intervention (i.e., the study group) against a different random subset who got a sham treatment such as a placebo or the standard treatment (i.e., the control group)[6]. Phase 3 trials have a high failure rate (e.g., 20% for heart failure, HF[7]) resulting from a high placebo response, lack of adequate power, random variability, or lack of relationship of biomarkers targeted in phase 2 and clinical efficacy endpoint in phase 3. As multiple confounding effects in phase 2 can result in a biased prediction in phase 3, it is highly valuable to predict the efficacy of the drugs with respect to the clinical endpoints for the Go/No-Go decision-making and design of the phase 3 trial. An important determinant in the mechanism of drug efficacy is the biological effect, such as the drug's binding affinity to the drug compound, known as drug-target interaction (DTI), so as to trigger the chemical reaction relevant to treating the disease[8]. On that account, predicting the novel off-target interactions[9-11] between drugs and targets (e.g., network-based machine learning methods[12-14]) is crucial in establishing reliable drug efficacy and, subsequently, in accelerating phase 3 pharmaceutical research in drug repurposing. Additionally, complementing biomedical databases with electronic health records (EHR) has also proven to be useful in computational drug repurposing[15,16]. The unique capability of the EHR offers real-world longitudinal diagnostic and pathophysiological patient data to test the hypotheses with statistical power[15,17].

Here, this study presents a comprehensive framework that proposes the prediction of phase 3 trial drug efficacy by leveraging the results from the drug-target prediction by biomedical knowledgebase-base and with statistical analysis by real-world data (i.e., EHR). Specifically, the proposed framework introduces a novel drug-target prediction model that generates low-dimensional representations (i.e., embeddings) of drugs and targets (proteins) based on the drug's chemical structure, gene sequence, and heterogeneous biomedical knowledgebase. The proposed model is tested and evaluated using a large-scale drug-target prediction benchmark, BETA[18], and is furhter used to discover hidden associations between phase 3 repurposed drugs and HF druggable genes. Additionally, EHR-based statistical analyses are conducted to determine the effectiveness of the repurposed drugs concerning HF-related clinical measurements such as natriuretic peptide tests (NT-proBNP). It should be noted that in this article, efficacy refers to a measure of drug performance under ideal conditions (e.g., trial), while effectiveness is a measure of its performance under real-world conditions (e.g., EHR).

2. Methods

In order to anticipate the efficacy of drugs being tested in phase 3 clinical trials for the treatment of HF, we utilized a combination of knowledge-based drug-target prediction and EHR-based drug-effective analysis. Our approach consisted of four main steps. Firstly, we identified the genes associated with the disease (HF) using a third-party knowledge base (DisGeNET[19]). Secondly, we used the list of disease-related genes to compute the probabilities of drug-target associations. To achieve this, we developed a novel predictive model for drug-target interactions, which was evaluated against the BETA drug-target prediction benchmark[18]. We then applied this model to predict the repurposed drugs for HF that are currently being tested in phase 3 clinical trials. Thirdly, we retrieved relevant patient cohorts from EHRs using clinical measures and inclusion/exclusion criteria. Lastly, we performed statistical analyses on the EHR data to

evaluate the effectiveness of the repurposed drugs in terms of HF-related clinical measurements. Please refer to Figure 1 (a) for a detailed overview of our study design.

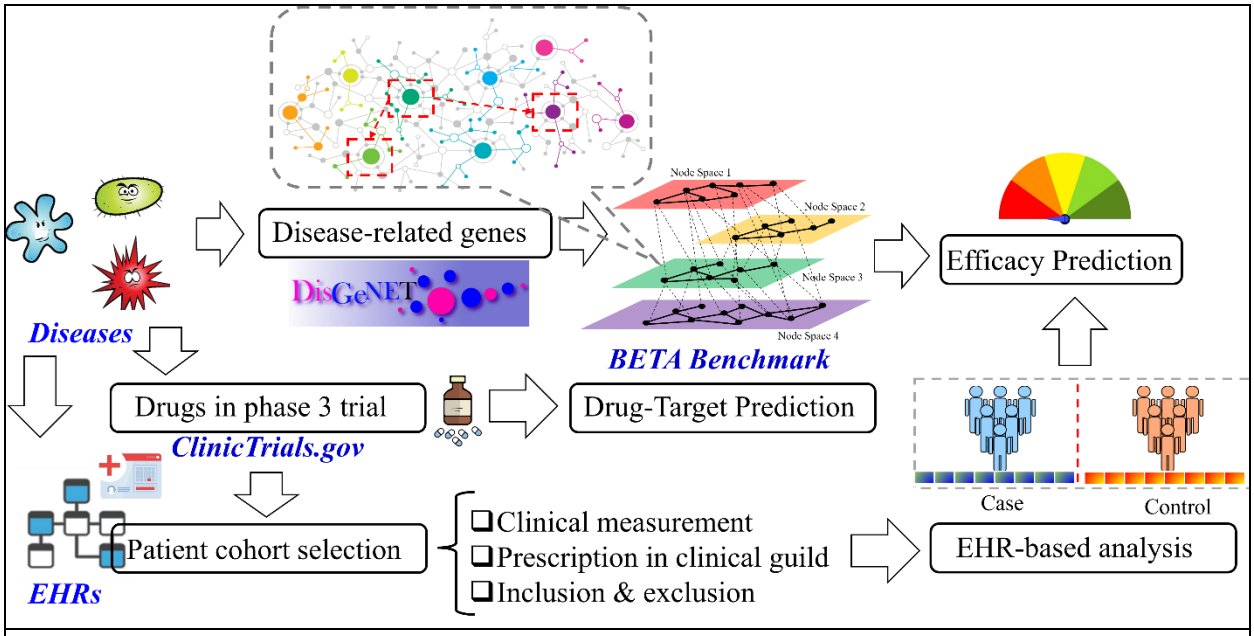

Figure 1, An overview of the proposed study, which aims to predict the efficacy of phase 3 clinical trial drugs for a specific disease (HF in this case). **(a)** shows the general pipeline of the study, which starts with a search for the relevant clinical trial drugs on ClinicTrials.gov and the identification of druggable genes using the DisGeNET[19] database. Using a list of repurposed trial drugs and the druggable genes, the proposed drug-target prediction model is used to predict the probabilities of drug-gene associations. The effectiveness of the repurposed drug is then analyzed based on EHR data, where the case (HF patients prescribed with the repurposed drugs) and control groups (HF patients prescribed without the repurposed drugs) are compared against a list of clinical measures related to HF. The overall prediction results are obtained by combining the results from drug-target prediction and EHR-based analysis.

2.1. Candidate for repurposed drugs for heart failure in phase 3 clinical trial

We conducted a search on ClinicalTrials.gov using the keywords "heart failure" and filtered by "completed" status and "Phase 3" phase in order to acquire drugs for our phase 3 clinical trial. From the 266 clinical trials obtained, two experts with medical degrees and experience in the clinical practice of Cardiology, Drs. Chen and Dai, reviewed each trial to determine if the drugs being studied were repurposed for heart failure treatment and if they showed clear efficacy. For the determination of drug efficacy, we reviewed results published in the trial or article with a clear indication of the trial number. After this review, we narrowed down the list to 80 clinical trials for 53 drugs. We then further filtered out drugs, not in the BETA benchmark and arrived at a final list of 24 drugs, with 9 showing a positive effect on heart failure and 15 showing no effect. Please see Supplementary Table 1 for the list of drugs.

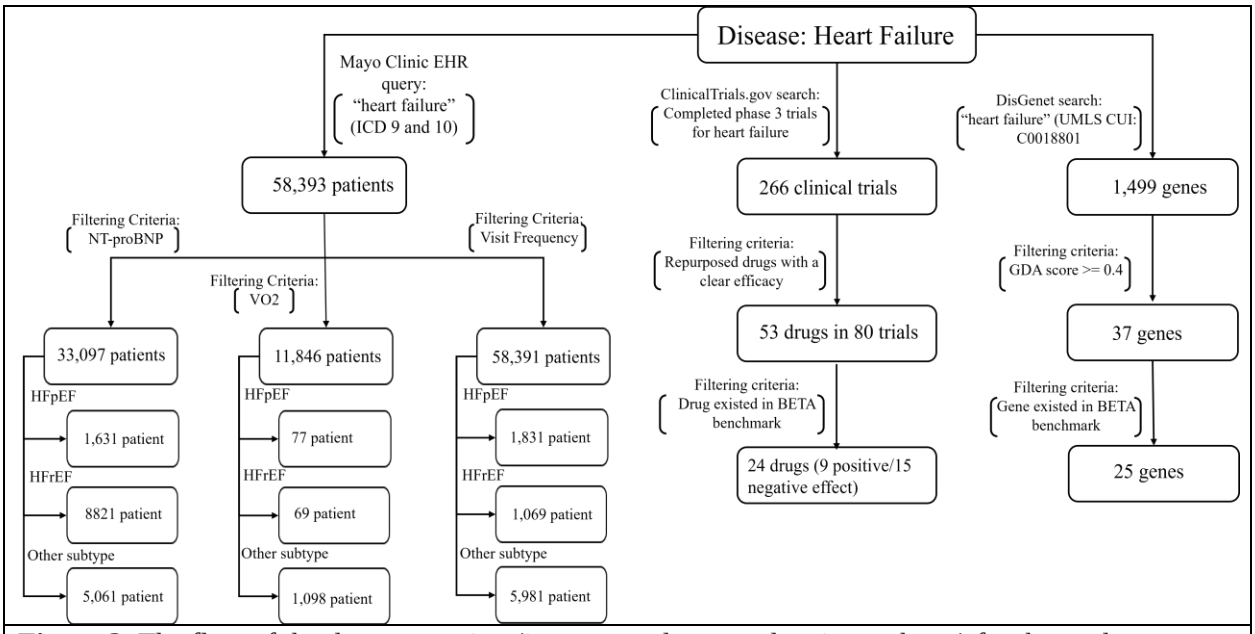

Figure 2. The flow of the data generation (i.e., genes, drugs, and patient cohorts) for the study.

2.2. Retrieve of the druggable genes

Using the publicly available knowledgebase, DisGeNET[19], we conducted a search for druggable genes associated with a particular disease. DisGeNET is a comprehensive database that contains genes and variants linked to human diseases, sourced from GWAS catalogs, scientific literature, and animal models. By entering the search term "heart failure" (UMLS CUI: C0018801) into the DisGeNET platform, we obtained a list of 1,499 results. We filtered them based on the GDA[19] score, which is a confidence score calculated by the number and type of sources (level of curation, organisms). We identified 25 genes related to "heart failure" that were also present in the BETA benchmark. For more information on these genes, please refer to Supplementary Table 2.

2.3. Drug-target prediction

Given a heterogeneous network $G(V, E)$, a set of vertices V (i.e. biomedical entities), and a set of edges E (i.e. known associations), where V are multiple types of vertices (e.g., drugs or targets) and E are multiple types of edges that connect the vertices (e.g., drug-target associations), our objective is to predict the potential new associations among V. Specifically, we have four types of vertices, which are drugs D, targets T, diseases Ds, and side effects S, where $D \in V, T \in V, Ds \in V, S \in V$. For each drug d, a chemical structure of such a drug is given as str_d. For each target t, a gene sequence of such a target is given as seq_t. When there is a linking edge existing in the G, the pair of vertices is defined as $\{(u, v) | u \in V \wedge v \in V \wedge (d, t) \in E\}$, while there is an unknown linking edge, the pair of vertices is defined as $\{(u, v) | u \in V \wedge v \in V \wedge (d, t) \notin E\}$. Therefore, the problem of drug-target prediction can be defined as, given a pair of a drug d and a target t, predicting whether the pair $\{(d, t) | d \in V \wedge t \in V\}$ is an existent association (refer to as positive) or nonexistent association (refer to as negative).

We have developed a novel hybrid model for drug-target prediction with the incorporation of two encoding networks (as depicted in Figure 1(b)), the D-network that learns the biological features from drug chemical structure and gene sequence, the N-network that learns the topological features from a set of bipartite knowledge graphs. Specifically, for the D-network, we get the latent vectors (i.e., embeddings) generated from the encoder $\mathcal{R}_{drug}(str_d)$ for the drug's chemical structure. Similarly, the latent vectors for gene sequences are obtained from the encoder $\mathcal{R}_{target}(seq_t)$. In practice, we adopted the encoder layers from

DeepPurpose[20]. For the N-network, we get the embedding of a node v from the t-th layer as $f^t(v) = \sum_{e(v,v')\in E} f^{t-1}(v')$, $f^{t-1}(v')$ embeddings of the neighborhood v'. In practice, $f^t(v)$ is implemented based on NeoDTI[12]. We further update the embedding as $f^t(v) \leftarrow Concat(f^{t-1}(v), \mathcal{R})$, where $\mathcal{R}$ is an embedding learned from D-network. Specifically, $\mathcal{R} = \mathcal{R}_{drug}$ when v is a drug and $\mathcal{R} = \mathcal{R}_{target}$ when v is a target. The output of the designed network is $\mathbb{C}$ which corresponds to the set of bipartite graphs. The loss function is $\mathbb{C} = \lambda_1\mathbb{C}_1 + \lambda_2\mathbb{C}_q \ldots + \lambda_k\mathbb{C}_k$, where $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$. $\mathbb{C}_k$ the objective function for the subgraph k, which is calculated: $\mathbb{C}_k = \sum_{e(u,v)\in E} \|s(e) - f(u)\mathbb{Q}\mathbb{N}^T f(v)\|^2$, where $s(e)$ is the weight of an edge e and $s(e) = 1$ if there is a link between the nodes u and v. $f(u)$ is the embedding of node u, $\mathbb{Q}$ and $\mathbb{N}$ are the edge projection function. For a detailed description of the proposed model, please refer to Supplementary File 1.

2.4. EHR-based statistical analysis

We conducted a comparison test to determine if the study cohort would experience better outcomes than the control cohort for HF-related measurements in response to repurposed drugs (refer to Table 1). We used three specific measurements: natriuretic peptide tests (NT-proBNP), maximum rate of oxygen consumption (VO2), and visit frequency. Our control group comprised HF patients (See Supplementary Table 3 for the identification of HF patients) who were prescribed medication from any of the following five categories defined by the AHA/ACC/HFSA guideline[21]: angiotensin-converting-enzyme inhibitors (ACEI), beta-blockers (BB), angiotensin II receptor blockers (ARB), statins, and loop diuretics. In contrast, our case group consisted of HF patients who also take repurposed drugs before completing their medication regimen from any of the above categories.

Table 1. Cohorts statistics of the Mayo Clinic EHR

	NT-proBNP	VO2	Visit Frequency
# patient in total	33,097	11,846	58,391
Female	14,654	4,337	25,014
Male	18,443	7,509	33,377
Average Age (SD)	72.27 (±14.47)	62.54 (±15.22)	71.52 (±14.67)
African American	885	531	1979
Hispanic or Latino	108	87	264
Non-Hispanic White	30,310	10,018	51,881
Others	1,794	1,210	4,267
ACEI	16,972	4,475	26,166
Beta Blocker	27,160	9,705	45,240
ARB	10,809	3,883	17,230
Statin	24,661	8,355	40,890
Loop Diuretic	31,304	11,021	52,486
# average visit (SD)	216.16 (±174.46)	157.59 (±163.40)	168.61(±162.17)
HFpEF	1,631	77	1,831
HFrEF	882	69	1,069
Other subtypes	5,061	1,098	5,981

To perform our analysis, we utilized three subtypes of HF based on the diagnosis tag: HF with reduced ejection fraction (HFrEF), HF with preserved ejection fraction (HFpEF), and other types of HF. Please refer to Supplementary Table 4 for additional information regarding the specific drug categories used.

2.5. Statistical analysis for the critical genes

To identify which druggable drugs could contribute more to the discovery of repurposed drugs, we conducted a statistical analysis to examine the relationship between gene prediction and drug efficacy. Specifically, we used a proposed method to predict the druggable genes related to heart failure (HF) that were identified by DisGeNET, and assigned a value of 1 or 0 to indicate positive or negative predictions, respectively.

3. Results

Our approach that predicts the efficacy of drugs in phase 3 clinical trials involves a combination of knowledge-based drug-target prediction and EHR-based drug-effective analysis. Firstly, we developed a unique drug-target predictive model and validated it using the BETA drug-target prediction benchmark. Using this model, we predicted potential repurposed drugs for HF identified in phase 3 clinical trials. Additionally, we conducted statistical analyses using EHR data to measure the effectiveness of these repurposed drugs with respect to clinical measurements related to HF. A detailed overview of our study design can be found in Figure 1 (a).

3.1. Development of the drug-target prediction model and the evaluation with BETA benchmark

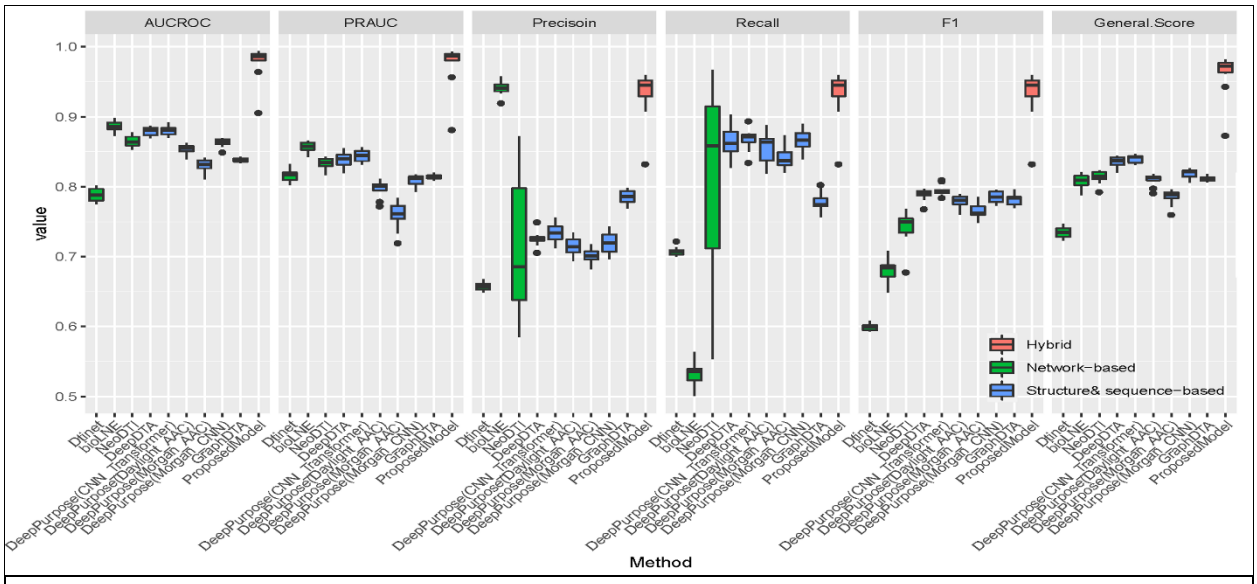

Figure 2. Performance of the proposed drug-target prediction method compared to other baseline methods in the evaluation based on the BETA benchmark. **(a)** a breakdown of the detailed results of the overall test in BETA evaluation tasks (i.e., Test 0). Overall, the proposed method outperforms all other methods across all evaluation measurements.

Our proposed model was put to the test using the BETA benchmark, which comprises of seven tests and 344 tasks designed to evaluate drug-target prediction models. These tests evaluate the models from various perspectives, including general screening, target and drug screening based on categories, searching for specific drugs and targets, and drug repurposing for specific diseases. To compare our model's performance, we used six state-of-the-art predictive models (DTINet[13], bioLNE[22], NeoDTI[12], DeepPurpose[20], DeepDTA[23], GraphDTA[24]) as baseline methods. The proposed method performed exceptionally well with an average AUCROC of 97.73%, PRAUC of 97.41%, F1 score of 93.13%, and General Score of 96.09%, outperforming the best-performing baseline methods, DeepPurpose (average AUCROC: 88.01%, PRAUC: 84.44%, F1: 79.49%, General Score: 83.98%), and NeoDTI (average AUCROC: 86.52%, PRAUC: 83.32%, F1: 74.19%, General Score: 81.34%). Our proposed method was outstanding, with 266 out of 404 tasks (65.84%) outperforming other methods across all seven tests. It's important to note that for test 4, we counted the average precision for the top 10, 20, 50, and 100, which made 80 tasks counted rather than 20 tasks counted in the BETA. For each test, our proposed method achieved excellent results with 10/10 (100.00%), 61/90 (67.78%), 66/72(91.67%), 62/72(86.11%), 42/80 (52.50%), 17/40 (42.50%), 8/40 (20.00%). Please refer to the Supplementary Figures 1-11for more detailed results.

3.2. Drug-target prediction for drug repurposing for heart failure in phase 3 clinical trial

We utilized the proposed drug-target prediction model to predict the efficacy of drugs in phase 3 clinical trials. Our search on ClinicalTrials.gov yielded 266 unique completed phase 3 clinical trials, which were designed to test repurposed drugs for HF. After multiple stages of filtering, we selected 24 repurposed drugs for our experiment. Out of these, 9 drugs exhibited positive effects (i.e., yes), while 15 drugs had no effect (i.e., no). We identified 25 druggable genes associated with HF using DisGeNET[19].

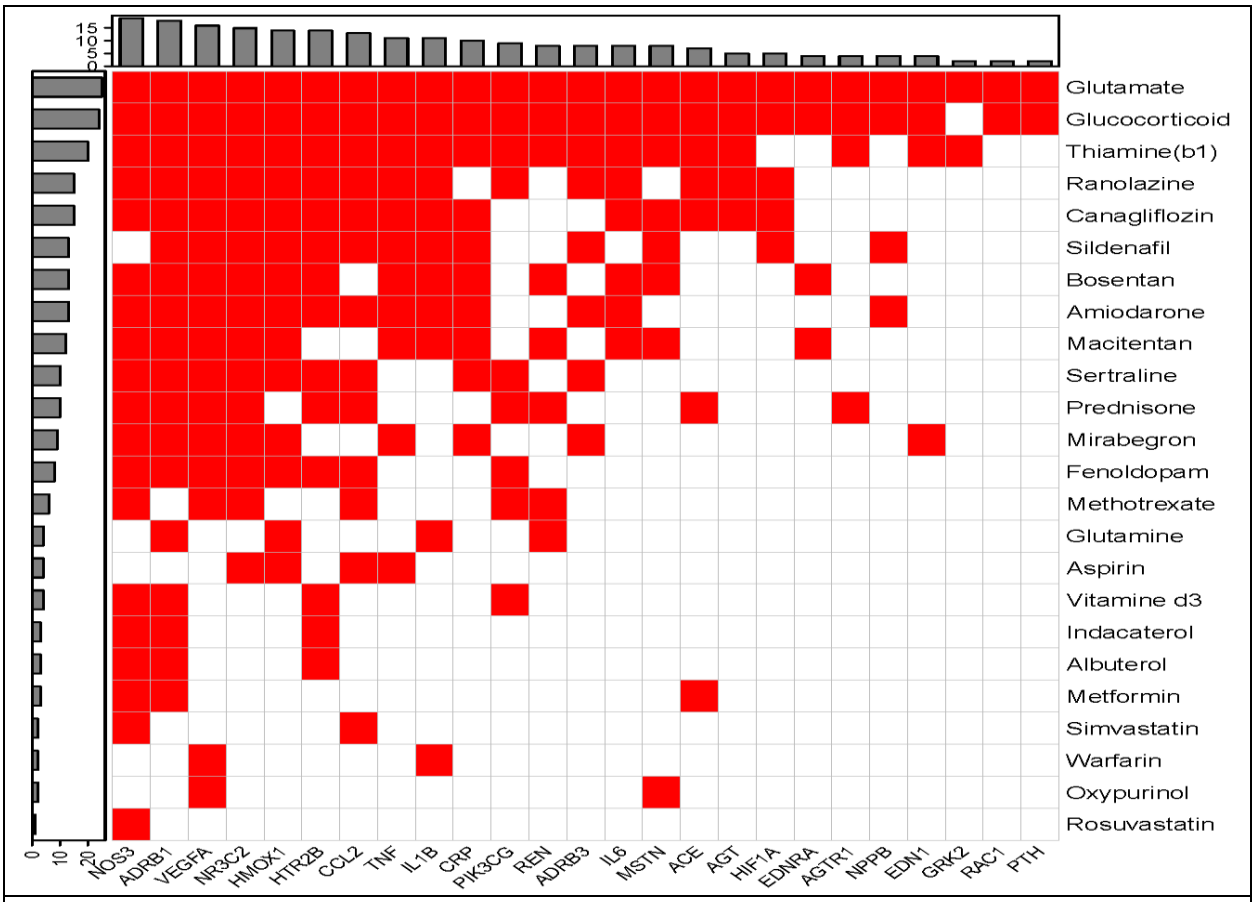

Figure 3. Predicted drug-target associations of the 24 phase 3 trial drugs.

The predictions of 24 drugs tested against 25 druggable genes related to HF are displayed in Figure 3. Glutamate, Glucocorticoid, and Thiamine B1 are the top drugs predicted to have the most HF-related genes, with 25/25, 24/25, and 20/25 genes, respectively. On the other hand, Rosuvastatin, Oxypurinol, Warfarin, and Simvastatin have the least number of HF-related genes, with only 1/25, 2/25, 2/25, and 2/25 genes, respectively. It is worth noting that all these drugs are used for treating or managing medical conditions related to the cardiovascular system. Specifically, Rosuvastatin and Simvastatin work by reducing cholesterol production and increasing clearance of LDL-cholesterol from the bloodstream by inhibiting HMG-CoA reductase, an enzyme involved in cholesterol synthesis. Regarding genes, NOS3, ADRB1, and VEGFA have the most HF-repurposed drugs, with 19/24, 18/24, and 16/24 drugs, respectively. Meanwhile, PTH, RAC 1, and GRK2 are the genes with the least number of repurposed drugs, with only 2/24 drugs for each.

3.3. EHR-based screening

We conducted an investigation to determine whether patients who were prescribed repurposed drugs, compared to those who were not, showed improvements in clinical measures recorded in the Mayo Clinic EHRs. Specifically, we learned that, among all the three clinical measurements related to heart failure (i.e.,

NT-proBNP, VO2, and visit frequency), NT-proBNP provides the best results. We divided patients into different groups based on their prescribed heart failure drugs[21], including ACEI, BB, ARB, statin, and loop diuretic, as well as subtypes of heart failure[25], such as HFrEF, HFpEF, and others (See Table 1 for more details).

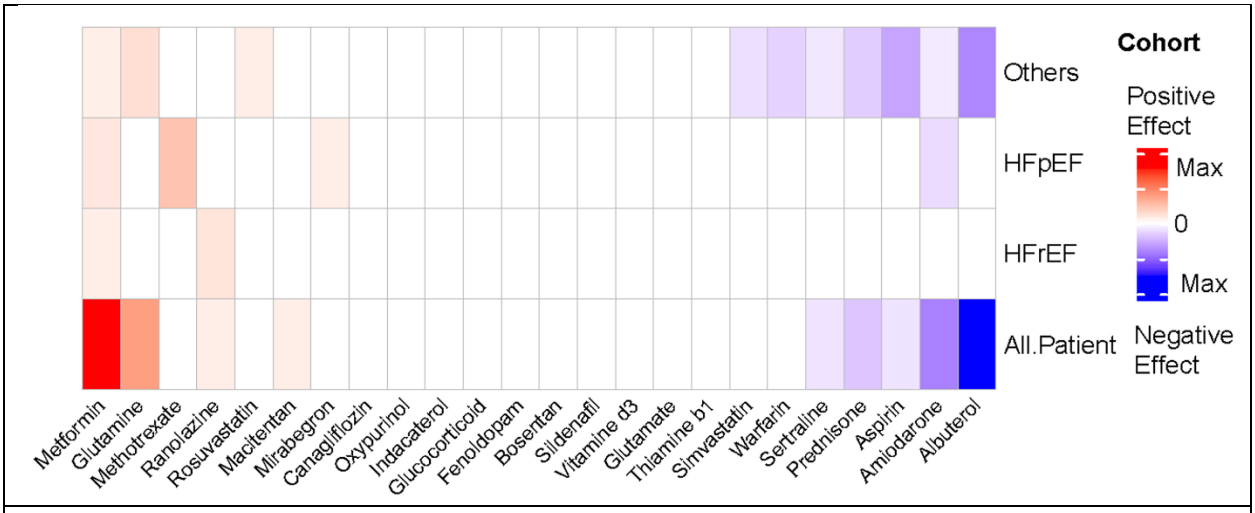

Figure 4. Effectiveness analysis of the 24 phase 3 trial drugs compared to the baseline HF drugs based on EHR. (a) Effectiveness of the repurposed drugs based on the comparison of the patient cohorts (e.g., All HF patients, HFrEF, HFpEF, and Others) with any of the HF drugs.

Figure 4 illustrates the predictions for 25 drugs tested across four cohorts, yielding three types of results: positive effect (where the case group has a significantly lower NT-proBNP value than the control group), unknown (where there is no significant difference in NT-proBNP values between case and control groups, or insufficient data to conduct the test), and negative effect (where the case group has a significantly higher NT-proBNP value than the control group). Among these drugs, metformin, glutamine, methotrexate, ranolazine, rosuvastatin, macitentan, and mirabegron exhibit a positive effect for HF patients in at least one of the four cohorts, with metformin (p value=1.20E-16), glutamine (1.28E-08), ranolazine (3.90E-02), and macitentan (3.82E-02) showing significance across all HF patients. Conversely, albuterol, amiodarone, atorvastatin, aspirin, prednisone, sertraline, warfarin, and simvastatin display a negative effect for HF patients in at least one of the four cohorts, with albuterol (5.64E-16), amiodarone (3.95E-09), atorvastatin (1.77E-05), aspirin (1.74E-02), prednisone (1.76E-04), and sertraline (1.54E-02) demonstrating significance across all HF patients. Moreover, the significance of a drug is contingent on the number of the sub-cohort (i.e., HF drug category), with metformin being the most significant positive drug and having a significant p-value across all HF drug categories (7.29E-14 for ACEI, 1.27E-16 for ARB, 6.25E-07 for BB, 9.79E-05 for STATIN, and 7.58E-09 for LD), and albuterol is the most significant negative drug, also with a significant p-value across all HF drug categories (1.30E-10 for ACEI, 2.01E-08 for ARB, 1.94E-18 for BB, 6.63E-14 for STATIN, and-2.88E-10 for LD).

3.4. Overall prediction

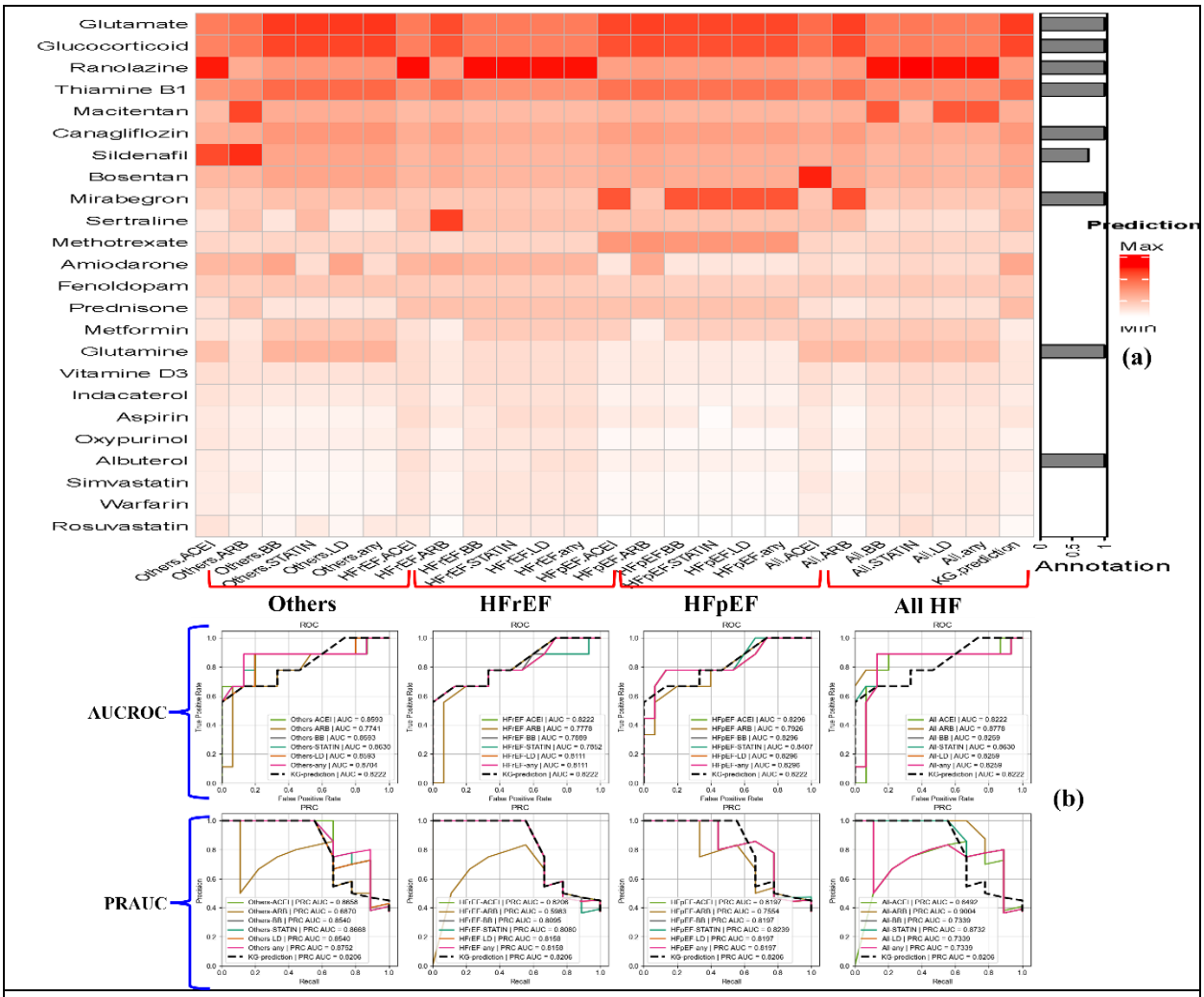

Figure 5. Overall prediction of 25 phases 3 trial drugs across different cohorts. (a) prediction for each repurposed drug across different cohorts (e.g., All HF patients, HFrEF, HFpEF, and Others). (b) AUC ROC and PRAUC (i.e, Average Precision) of the over prediction for each cohort.

By integrating the outcomes of drug-target prediction and EHR analysis, we can obtain a comprehensive estimate of the over-prediction of phase 3 trial drugs. This estimation is computed using a $score\ (d) = k_d e^{s(p_k)}$, where k_d represents the number of predicted associations based on drug-target prediction, and $s(p_k)$ reflects the effect of a drug based on the statistical significance of its p-value (1 for a positive effect, -1 for a negative effect with a significant p-value, and 0 for other scenarios) in practical terms.

The prediction achieved an AUCROC of 82.59% and PRAUC (Average Precision) of 73.39% overall. Among the five drug categories, the highest accuracy was observed for ARB drugs, with an AUCROC of 87.78% and PRAUC of 90.04%. Regarding the cohorts, the Others category exhibited the best performance across all cohorts, with an AUCROC of 87.04% for Others, 81.1% for HFrEF, and 82.96% for HFpEF. Notably, we observed a significant improvement in screening accuracy when utilizing EHR and drug-target-based prediction (i.e., knowledge graph-based prediction) as opposed to only using drug-target-based

prediction, with the best AUCROC of 87.04% for Others, 82.22% for HFrEF, 84.07% for HFpEF, and 87.78% across All HF patients, compared to 82.22% for knowledge graph-based prediction only.

The high-efficacy drugs successfully predicted by our method include Glutamate, Glucocorticoid, Ranolazine, and Thiamine B1. Glutamate exhibited strong support with 25 out of 25 predicted genes, and Glucocorticoid and Thiamine B1 had 24 out of 25 genes and 20 out of 25 genes predicted, respectively. Ranolazine had moderate support from predicted genes (15 out of 25) but strong evidence of a significant positive effect from EHR data in 10 out of 24 cohorts. Additionally, Rosuvastatin, Warfarin, and Simvastatin were successfully predicted to have low efficacy. Warfarin and Simvastatin negative effects were observed in EHR data in 8 out of 24 cohorts and 5 out of 24 cohorts, respectively, with a few druggable genes (2 out of 25 genes) associated. Despite some positive effects observed in EHR data for Rosuvastatin in 4 out of 24 cohorts, its low support in terms of the number of druggable genes (1 out of 25 genes) resulted in its classification as low efficacy.

4. Discussion

This study aimed to predict the efficacy of repurposed drugs for phase 3 trials using a novel method that combines biological and clinical perspectives. From a biological standpoint, the study developed a new drug-target predictive model that incorporates drug chemical structure, gene sequences, and heterogeneous biomedical knowledge from eight knowledge graphs. The proposed model was evaluated using an open-sourced benchmark, BETA[18], and demonstrated superior results compared to all baseline methods across seven tests. From a clinical standpoint, the study generated various sub-cohorts based on drug category and disease subtypes and conducted a statistical analysis using EHR data to evaluate the effectiveness of repurposed drugs in terms of different clinical measurements. To illustrate the proposed method, the study examined 24 repurposed drugs with phase 3 trials for HF and successfully predicted high efficacy for Glutamate, Glucocorticoid, Ranolazine, and Thiamine B1, while predicting low efficacy for Rosuvastatin, Warfarin, and Simvastatin.

Although three measurements were taken from the EHR to assess the effectiveness of repurposed drugs, only the results for NT-proBNP are presented as the other two measures were found to be inadequate in testing the drug's effectiveness (See Supplementary Table 5-6 for the drug screening based on VO2 and visit frequency). Additionally, a unified method was used to pull patient cohorts (e.g., same control groups) for statistical testing, which may result in reduced power compared to a more tailored method that matches real trial eligibility criteria for the patient cohort. However, building a customized cohort in EHR-based automatic analysis that accurately reflects a real trial can be challenging due to dataset limitations, such as insufficient sample size or missing variables. This limitation may introduce bias into the study. The study's inclusion and exclusion criteria, clinical measurements, and eligibility criteria were simplified compared to real phase 3 trials. As this is a proof-of-concept study, a leverage of the tradeoff between availability, complexity, bias, and accuracy led to the development of a proposed method that simplifies factors for EHR-based analysis. The authors acknowledge that a more customized analysis can provide more accurate predictions for some drugs in practice. Another potential solution to increase power in the study is to utilize EHR data from multiple sources. The authors propose two potential directions: Using integrated EHR databases (e.g., All of Us[26]) can provide a feasible solution to conduct the analysis over the large patient cohort while a potential bias can be embedded as the factors used to pull the cohort may vary in each participant hospital. Another solution is sharing models (e.g., federated learning[27,28]) that are built separately with a selected set of hospitals available in the existing research network (e.g., PCORnet[29]) but with the potential of more effort in the data processing and common challenges (e.g., the culture of keeping an advantage for competitive) in the multi-site study[30,31].

The process of drug-target prediction is used to assess the feasibility of repurposing a drug to treat a particular disease. To prioritize potentially druggable genes, while the GDA scores[19], calculated using multiple scientific sources such as the number of supporting publications, are used in the gene selection in

the proposed method, they are considered equally important in predicting drug-target associations. Our preliminary analysis, detailed in Supplementary Table 7, indicates that certain genes (e.g., AGT, HIF1A) are significantly associated with repurposed drugs for treating HF. Previous studies have shown that the AGT and HIF1A increase the risk of HFpEF for patients with chronic kidney disease[32] and patients with obesity and metabolic syndrome[33]. Although the significance is based on the prediction of the potential novel drug-target associations, it may be beneficial to evaluate these genes more comprehensively as they are discovered to play a more critical role in drug development (e.g, AGT plays a central role in the regulation of blood pressure[34], and HIF1A has been implicated in the development of cardiac hypertrophy and fibrosis in a low oxygen level[35]).

The focus of this study was to use HF as a case study for evaluating the proposed method's ability to predict repurposed drugs for phase 3 trials. We chose HF as it is a complex disease with limited treatment options and a high unmet medical need. Although the initial results demonstrate promising predictive performance, it may be premature to draw conclusions about the method's effectiveness. Notably, while a more advanced ad-hoc linear model can be used in the selection of the combination of the cohort and potentially achieve better performance in prediction, it was not included in the results section due to concerns about overfitting and lack of generalizability. Therefore, it is crucial to test the proposed method across a broad range of diseases to determine which diseases perform better and identify any commonalities among them. This information is necessary to optimize the method's performance and facilitate its use in drug repurposing for precision medicine in clinical practice.

Contributors

NZ conceptualized the study. NZ designed the drug-target predictive model, SZ and SC developed and optimized the model. SC prepared data and NZ conducted the experiments. NZ, SC, and YY conducted the statistical analysis and data visualization. The initial draft of the report was composed by NZ and SC, and the other authors evaluated and provided feedback on it. All of the authors collectively assumed the ultimate responsibility of deciding to submit the report for publication and committed to being answerable for all aspects of the work, ensuring that any concerns regarding the precision or honesty of any part of the work are adequately addressed and resolved. NZ supervised and administered the study.

Declaration of interests

None

Acknowledgment

This work is supported by a grant from the National Institute of Health (NIH) NIGMS (R00GM135488).

Reference

1. Park K. A review of computational drug repurposing. *Translational and clinical pharmacology* 2019; **27**(2): 59-63.
2. Wang X, Guan Y. COVID-19 drug repurposing: a review of computational screening methods, clinical trials, and protein interaction assays. *Medicinal research reviews* 2021; **41**(1): 5-28.
3. Talevi A, Bellera CL. Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics. *Expert Opinion on Drug Discovery* 2020; **15**(4): 397-401.
4. Wu Z, Li W, Liu G, Tang Y. Network-based methods for prediction of drug-target interactions. *Frontiers in pharmacology* 2018; **9**: 1134.

5. Cheng F, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS computational biology* 2012; **8**(5): e1002503.
6. bookS Metaphor S, StatiSticS A. Time for one-person trials. *Nature* 2015; **520**: 30.
7. Bayes-Genis A, Aimo A, Jhund P, et al. Biomarkers in heart failure clinical trials. A review from the Biomarkers Working Group of the Heart Failure Association of the European Society of Cardiology. *European Journal of Heart Failure* 2022; **24**(10): 1767-77.
8. Lambrinidis G, Tsantili-Kakoulidou A. Multi-objective optimization methods in novel drug design. *Expert Opinion on Drug Discovery* 2021; **16**(6): 647-58.
9. Bagherian M, Sabeti E, Wang K, Sartor MA, Nikolovska-Coleska Z, Najarian K. Machine learning approaches and databases for prediction of drug–target interaction: a survey paper. *Briefings in bioinformatics* 2021; **22**(1): 247-69.
10. Ezzat A, Wu M, Li X-L, Kwoh C-K. Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey. *Briefings in bioinformatics* 2019; **20**(4): 1337-57.
11. Chen R, Liu X, Jin S, Lin J, Liu J. Machine learning for drug-target interaction prediction. *Molecules* 2018; **23**(9): 2208.
12. Wan F, Hong L, Xiao A, Jiang T, Zeng J. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions. *Bioinformatics* 2019; **35**(1): 104-11.
13. Luo Y, Zhao X, Zhou J, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nature communications* 2017; 8(1): 573.
14. Li J, Lu Z. Pathway-based drug repositioning using causal inference. *BMC bioinformatics* 2013; 14(16): 1-10.
15. Zhou M, Zheng C, Xu R. Combining phenome-driven drug-target interaction prediction with patients' electronic health records-based clinical corroboration toward drug discovery. *Bioinformatics* 2020; 36(Supplement_1): i436-i44.
16. Hodos RA, Kidd BA, Khader S, Readhead BP, Dudley JT. Computational approaches to drug repurposing and pharmacology. *Wiley interdisciplinary reviews Systems biology and medicine* 2016; **8**(3): 186.
17. Zhou M, Wang Q, Zheng C, Rush AJ, Volkow ND, Xu R. Drug repurposing for opioid use disorders: integration of computational prediction, clinical corroboration, and mechanism of action analyses. *Molecular psychiatry* 2021: 1-11.
18. Zong N, Li N, Wen A, et al. BETA: a comprehensive benchmark for computational drug-target prediction. *Briefings in Bioinformatics* 2022; **bbac199**.
19. Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic acids research* 2016: gkw943.
20. Huang K, Fu T, Glass LM, Zitnik M, Xiao C, Sun J. DeepPurpose: a deep learning library for drug–target interaction prediction. *Bioinformatics* 2020.
21. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Journal of the American College of Cardiology* 2022; **79**(17): e263-e421.
22. Zong N, Wong RSN, Yu Y, Wen A, Huang M, Li N. Drug–target prediction utilizing heterogeneous bio-linked network embeddings. *Briefings in bioinformatics* 2021; **22**(1): 568-80.
23. Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug–target binding affinity prediction. *Bioinformatics* 2018; **34**(17): i821-i9.
24. Nguyen T, Le H, Venkatesh S. GraphDTA: prediction of drug–target binding affinity using graph convolutional networks. *BioRxiv* 2019: 684662.
25. Shah SJ. Precision medicine for heart failure with preserved ejection fraction: an overview. *Journal of cardiovascular translational research* 2017; **10**(3): 233-44.
26. Investigators AoURP. The "All of Us" research program. *New England Journal of Medicine* 2019; **381**(7): 668-76.

27. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research* 2021; **5**(1): 1-19.
28. Yang Q, Liu Y, Cheng Y, Kang Y, Chen T, Yu H. Federated learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 2019; **13**(3): 1-207.
29. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association* 2014; **21**(4): 578-82.
30. Cifuentes M, Davis M, Fernald D, Gunn R, Dickinson P, Cohen DJ. Electronic health record challenges, workarounds, and solutions observed in practices integrating behavioral health and primary care. *The Journal of the American Board of Family Medicine* 2015; **28**(Supplement 1): S63-S72.
31. Reisman M. EHRs: the challenge of making electronic data usable and interoperable. *Pharmacy and Therapeutics* 2017; **42**(9): 572.
32. Xu Y, Rong J, Zhang Z. The emerging role of angiotensinogen in cardiovascular diseases. *Journal of Cellular Physiology* 2021; **236**(1): 68-78.
33. Warbrick I, Rabkin SW. Hypoxia-inducible factor 1-alpha (HIF-1α) as a factor mediating the relationship between obesity and heart failure with preserved ejection fraction. *Obesity Reviews* 2019; **20**(5): 701-12.
34. McMurray JJ, Packer M, Desai AS, et al. Angiotensin–neprilysin inhibition versus enalapril in heart failure. *n engl j med* 2014; **371**: 993-1004.
35. Krishnan J, Suter M, Windak R, et al. Activation of a HIF1α-PPARγ axis underlies the integration of glycolytic and lipid anabolic pathways in pathologic cardiac hypertrophy. *Cell metabolism* 2009; **9**(6): 512-24.