

Original Paper

Hypertension Medication Recommendation via Synergistic and Selective Modeling of Heterogeneous Medical Entities: Development and Evaluation Study of a New Model

Ke Zhang¹, MS; Zhichang Zhang¹, PhD; Yali Liang¹, MS; Wei Wang¹, MS; Xia Wang², BS

¹College of Computer Science and Engineering, Northwest Normal University, Lanzhou, China

²Gansu Provincial Hospital, Lanzhou, China

Corresponding Author:

Zhichang Zhang, PhD
College of Computer Science and Engineering
Northwest Normal University
967 Anning East Road
Lanzhou 730070
China
Phone: +86 13038769329
Email: zzc@nwnu.edu.cn

Abstract

Background: Electronic health records (EHRs) contain comprehensive information regarding diagnoses, clinical procedures, and prescribed medications. This makes them a valuable resource for developing automated hypertension medication recommendation systems. Within this field, existing research has used machine learning approaches, leveraging demographic characteristics and basic clinical indicators, or deep learning techniques, which extract patterns from EHR data, to predict optimal medications or improve the accuracy of recommendations for common antihypertensive medication categories. However, these methodologies have significant limitations. They rarely adequately characterize the synergistic relationships among heterogeneous medical entities, such as the interplay between comorbid conditions, laboratory results, and specific antihypertensive agents. Furthermore, given the chronic and fluctuating nature of hypertension, effective medication recommendations require dynamic adaptation to disease progression over time. However, current approaches either lack rigorous temporal modeling of EHR data or fail to effectively integrate temporal dynamics with interentity relationships, resulting in the generation of recommendations that are not clinically appropriate due to the neglect of these critical factors.

Objective: This study aims to overcome the challenges in existing methods and introduce a novel model for hypertension medication recommendation that leverages the synergy and selectivity of heterogeneous medical entities.

Methods: First, we used patient EHR data to construct both heterogeneous and homogeneous graphs. The interentity synergies were captured using a multihead graph attention mechanism to enhance entity-level representations. Next, a bidirectional temporal selection mechanism calculated selective coefficients between current and historical visit records and aggregated them to form refined visit-level representations. Finally, medication recommendation probabilities were determined based on these comprehensive patient representations.

Results: Experimental evaluations on the real-world datasets Medical Information Mart for Intensive Care (MIMIC)-III v1.4 and MIMIC-IV v2.2 demonstrated that the proposed model achieved Jaccard similarity coefficients of 58.01% and 55.82%, respectively; areas under the curve of precision-recall of 83.56% and 80.69%, respectively; and F_1 -scores of 68.95% and 64.83%, respectively, outperforming the baseline models.

Conclusions: The findings indicate the superior efficacy of the introduced model in medication recommendation, highlighting its potential to enhance clinical decision-making in the management of hypertension. The code for the model has been released on GitHub.

JMIR Med Inform 2025;13:e74170; doi: [10.2196/74170](https://doi.org/10.2196/74170)

Keywords: hypertension medication recommendation; electronic health records; heterogeneous graph; multihead graph attention mechanism; time series selective mechanism

Introduction

Background

Hypertension represents a prevalent chronic condition and serves as a significant contributor to cardiovascular mortality, making timely pharmacological intervention for blood pressure management crucial [1]. With the increasing trend of an aging population, the growing number of patients with hypertension has placed a significant burden on health care systems [2]. Consequently, automated medication recommendation systems for hypertension have been developed.

Prior Work and Limitations

Early hypertension medication recommendation methods were primarily rule-based. For example, Wu and Xie [3] developed a hypertension ontology and reasoning rules to recommend appropriate antihypertensive medications to patients. However, these methods relied only on predefined rules and a limited set of case data, neglecting other critical patient information, which resulted in recommendations that lacked flexibility and personalization. In recent years, numerous neural network models for recommending hypertension medications based on electronic health records (EHRs) have been proposed [4,5]. These models have shown improved outcomes and effectively address many limitations inherent in earlier algorithms.

Nevertheless, the complexity of EHR data continues to present significant challenges for medication recommendation tasks, particularly in 2 critical areas: insufficient synergy among heterogeneous medical entities and neglect of temporal dynamics in the patient's condition.

Insufficient Synergy Among Heterogeneous Medical Entities

EHR data contains heterogeneous but interrelated medical entities—such as diagnoses, procedures, and medications—that jointly influence treatment outcomes. For example, a diagnosis indicates a patient's health status, which subsequently informs procedure and medication decisions. Effective modeling of such cross-entity relationships is essential for generating accurate and personalized treatment recommendations.

Existing approaches like Multilevel Medical Embedding (MiME) [6] and graph convolutional transformer (GCT) [7] attempt to model medical concepts and their causal relations using homogeneous graph structures. However, these structures fail to capture the inherent heterogeneity across EHR entities. The Heterogeneous Information Network for Medical Diagnosis (HeteroMed) [8] introduces semantic associations via metapaths, but it does not model dynamic interentity interactions. CausalMed [9] explores causal inferences between treatment elements but overlooks the reinforcement between medications and procedures. Graph transformers of bidirectional encoder representations from

transformers on EHRs (GT-BEHRT) [10] integrates graph transformers with temporal modeling yet does not incorporate contextual synergy among entities.

Further, current models inadequately balance efficacy and safety. For instance, Graph-Augmented Memory Networks (GAMENet) [11] uses a drug-drug interaction (DDI) graph to reduce adverse effects but does not consider diagnostic and procedural inputs. Graph-augmented bidirectional encoder representations from transformers (G-BERT) [12] extends to multientity modeling yet omits DDI-aware safety metrics. Competitive neural network (CompNet) [13] faces issues with computational efficiency and model stability on large-scale datasets. Recent work like that by Yang et al [14] (using medication molecular structure graphs) and Li et al [15] (using contrastive learning) enhance representation power, but they still lack a unified mechanism for modeling both “multientity synergy” and “safety constraints.”

Neglect of Temporal Dynamics in the Patient's Condition

As a chronic disease, hypertension evolves over time. In clinical practice, treatment strategies are often adjusted based on both current and historical visit data. Hence, capturing the temporal dynamics of the patient's condition is essential for effective medication recommendation.

Several studies have addressed sequential modeling. Yang et al [16] used dual medical sequences to represent medication history but did not account for intersequence relationships. Liu et al [17] developed 3 long short-term memory (LSTM) variants to model correlations across medical sequences but failed to consider the immediate influence of current clinical status. Although models like the Reverse Time Attention model (RETAIN) [18] use attention mechanisms to prioritize relevant historical visits and An et al [19] introduced hierarchical temporal modeling, these methods primarily focus on general temporal behavior without consideration for hypertension-specific patterns like long-term management or phased treatment adjustment.

Advanced designs have attempted to improve modeling capabilities. Le et al [20] used a memory-enhanced neural network to represent long-term dependencies, while Yang et al [21] adopted a residual mechanism to capture patient status transitions. Wu et al [22] introduced a transformer with a “copy-generate” mechanism to decide whether to reuse previous prescriptions. However, none of these methods fully addresses the task-specific temporal dynamics required for hypertension medication recommendations.

Objective and Contributions

To address the aforementioned limitations, this study proposed CSRec, a novel hypertension medication recommendation framework that integrates the synergistic interactions of heterogeneous medical entities with selective modeling of temporal progression. CSRec is designed to

capture cross-entity synergies through both heterogeneous and homogeneous graph construction based on EHR data, model time-aware patient representations using a temporal selectivity module that weighs current versus historical visits, and enhance safety by incorporating DDI information in the recommendation process.

The primary contributions of this paper are summarized in the following paragraphs.

We propose a novel medication recommendation model, named CSRec, specifically designed for hypertension treatment. CSRec effectively integrates the synergistic interactions and selective characteristics among diverse medical entities. By constructing a heterogeneous medical entity graph derived from EHRs, our model utilizes a graph attention mechanism to generate enhanced collaborative embeddings among medical entities. Additionally, a temporal selection mechanism was incorporated to simulate hypertension progression, thereby producing a comprehensive patient representation to facilitate accurate medication recommendations.

To better capture synergistic relationships between medical entities, we innovatively modified the traditional graph attention network (GAT) to focus more on neighboring node information, thus obtaining a more aggregated representation of the principal nodes.

Extensive experiments on the publicly available datasets Medical Information Mart for Intensive Care (MIMIC)-III

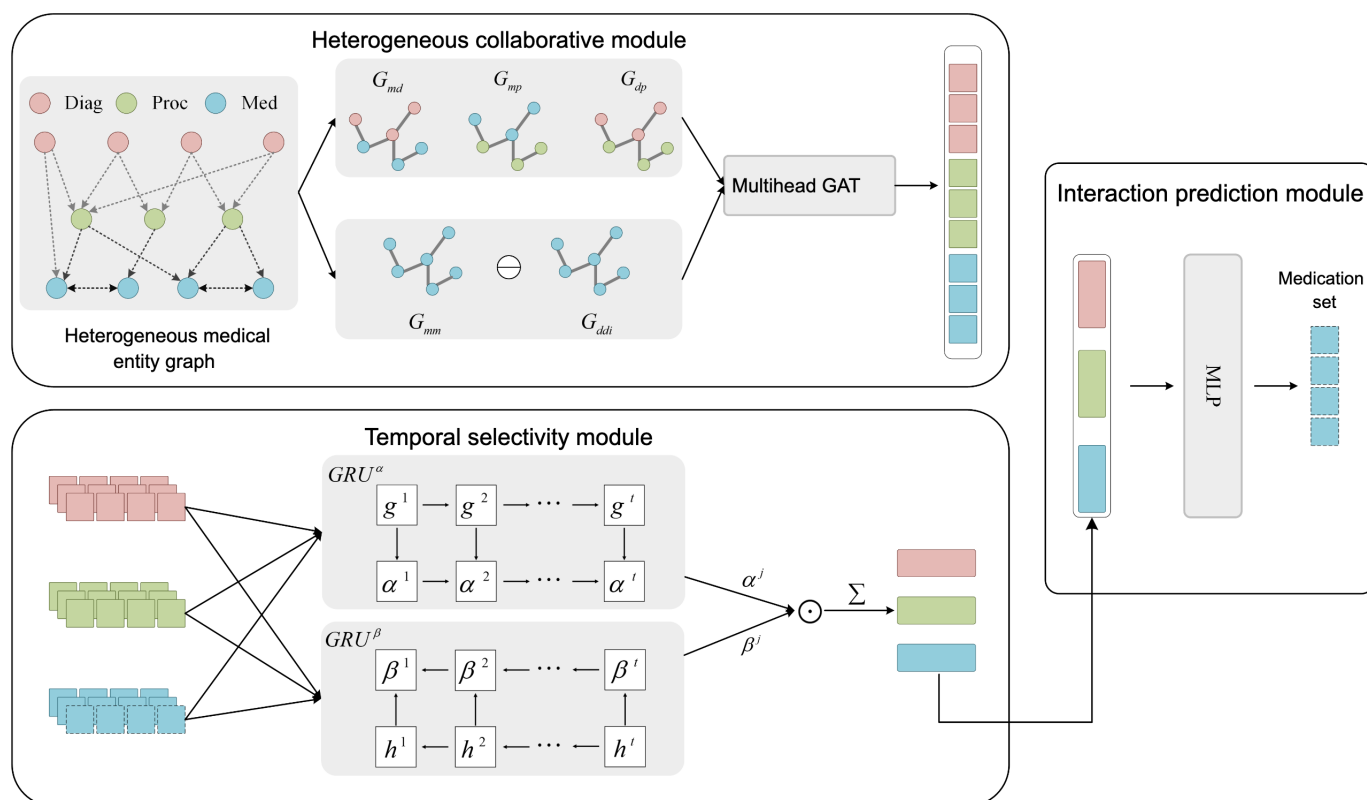
[23] and MIMIC-IV [24] were conducted to validate the superiority and effectiveness of our proposed method.

Methods

Overview

In this section, we present a comprehensive description of the CSRec model's structure. As depicted in Figure 1, our model comprises 3 core modules for end-to-end hypertension medication recommendations: (1) The Heterogeneous Collaborative Module, with heterogeneous and homogeneous graph networks, learns entity-level representations by aggregating collaboration patterns between medical entities, which are then passed to the Temporal Selectivity Module; (2) the Temporal Selectivity Module, using a bidirectional selection mechanism, processes these entity-level representations to calculate relevance coefficients between current and previous visits, generating visit-level entity representations that are transmitted to the Interaction Prediction Module; and (3) the Interaction Prediction Module concatenates these visit-level entity representations to enrich entity information, forming a patient representation. This is converted into medication recommendation probabilities, with medications exceeding a threshold output as results.

Figure 1. The framework of CSRec. Diag: diagnosis; GAT: graph attention network; GRU: gated recurrent unit; Med: medication; MLP: multilayer perceptron; Proc: procedures.



Problem Formulation

Electronic Health Records

EHRs encompass a variety of medical visit information collected from patients. For a specific patient, EHRs can be structured into a sequence comprising multiple clinical visit records, represented as $V = [V^1, V^2, \dots, V^T]$, where V^t represents the t -th visit and T indicates the total number of visits for that patient. Specifically, each clinical visit V^t can be represented as $V^t = [V_d^t, V_p^t, V_m^t]$, in which $V_d^t \in \{0, 1\}^{|D|}$, $V_p^t \in \{0, 1\}^{|P|}$, and $V_m^t \in \{0, 1\}^{|M|}$ represent multihot encoded vectors corresponding to diagnosis, procedures, and medications, respectively. Here, the notation $||$ indicates the total number of distinct categories within each respective medical entity type.

Heterogeneous Medical Entities

In this paper, distinct medications were modeled as medication entities. Each medication recorded in the EHR corresponds to a unique medication entity, identified by a specific identifier denoted as $\{m_1, m_2, \dots\}$, and each entity is independently embedded within the model using embeddings of identical dimensionality. Likewise, diagnoses and procedures were categorized into diagnostic entities and procedural entities, respectively. Collectively, these 3 categories of entities were termed heterogeneous medical entities.

Hypertension Medication Recommendation

Based on the patient's current diagnostic information V_d^t , procedural data V_p^t , historical visit sequence

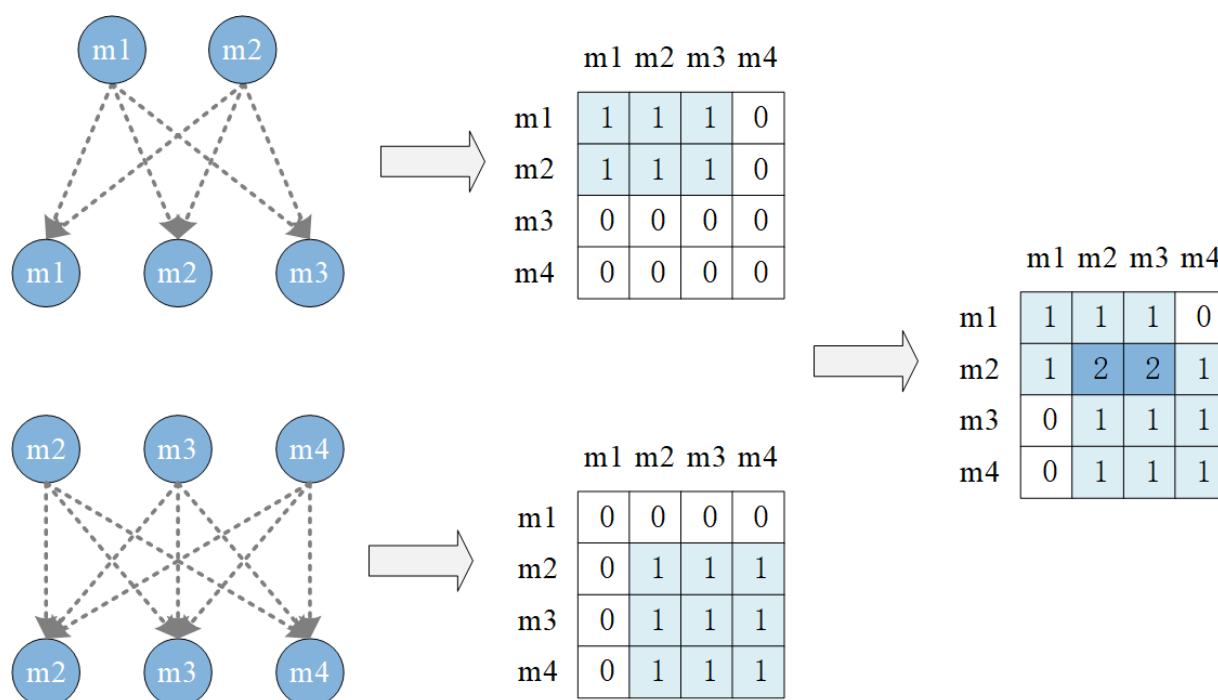
$V = [V^1, V^2, \dots, V^{T-1}]$, and heterogeneous medical entity graph G , the model recommends appropriate antihypertensive medications $\hat{y}_t \in \{0, 1\}^{|M|}$ to the patient.

Heterogeneous Collaborative Module

For a specific patient, we first constructed a medication homogeneous graph, which was explicitly designed to characterize relationships between medications and provide a foundation for subsequent cross-entity modeling. This graph comprised 2 components, both of which take medications as the sole node type.

One component is the medication collaboration graph $G_{mm}(t-1) = \{M(t-1), A^{mm}\}$, which serves to capture the patterns of combined medication use in clinical practice. Specifically, $M(t-1)$ denotes the collection of all prescribed medications from the patient's previous visits, and A^{mm} signifies the adjacency matrix representing collaborative interactions among medications. Each entry within this adjacency matrix indicates the initial collaboration weight between the respective medication nodes. The detailed procedure for generating this matrix is illustrated in Figure 2. Initially, A^{mm} is set as a zero matrix. If medication i and medication j co-occur during a specific visit, then $A[i, j] = 1$; if medication i and j co-occur across multiple visits, the corresponding value in A^{mm} is incremented, with higher values indicating stronger collaboration between the medications.

Figure 2. The medication-medication synergy matrix construction process.



Each medication node $m \in M(t-1)$ in the graph corresponds to its initial embedding vector, expressed as:

$$e_m = V_m^{t-1} E_m \quad (1)$$

Specifically, E_m refers to the embedding matrix associated with medication entities. The selector vector V_m^{t-1} specifically extracts the relevant embedding from E_m for the medication at the previous time step $t-1$.

Moreover, some medications may have harmful interactions (DDI) and should be avoided when used together. To address this, we integrated the medication collaboration graph G_{mm} with the medication safety graph G_{ddi} , enhancing the comprehensive representation of the medication nodes. Notably, the methodology used to construct the medication safety graph G_{ddi} closely mirrors that of G_{mm} . Specifically, we used A^{ddi} as the adjacency matrix, where $A^{ddi}[i, j] = 1$ signifies that the i -th and j -th medications exhibit a paired harmful medication interaction.

$$G_m = G_{mm} - \lambda G_{ddi} \quad (2)$$

Subsequently, we adopted a GAT-based graph neural approach to obtain embeddings for medication nodes within the medication graph G_m .

First, we utilized the attention mechanism to calculate the attention coefficients between a node and its neighbors, followed by normalization:

$$\alpha_{ij}^{(m)} = \frac{\exp(g_{ij}^{(m)})}{\sum_{k \in N_i^{(m)}} \exp(g_{ik}^{(m)})} \quad (3)$$

Here, $g_{ij}^{(m)} = \text{LeakyReLU}(\beta^{(m)T} (h_i^{(m)} \parallel h_j^{(m)}))$ denotes the output of the LeakyReLU activation function applied to the linear transformation of the concatenated feature vectors $h_i^{(m)}$ and $h_j^{(m)}$, parameterized by the learnable weight vector $\beta^{(m)}$. The notation $N_i^{(m)}$ represents the neighboring node set of node m_i in the graph G_m , while \parallel signifies the vector concatenation operation.

Second, we used the calculated attention coefficients to perform weighted aggregation of neighbor nodes, thereby obtaining the representation of node m_i :

$$e_i^{(m)} = \sigma \left(\sum_{j \in N_i^{(m)}} \gamma_{ij}^{(m)} h_j^{(m)} \right) \quad (4)$$

To overcome the limitations posed by single-view attention, we further introduced a multihead attention mechanism, using a linear layer to map node representations into multiple subspaces, then aggregate representations under each subspace. Formally, this operation is expressed as:

$$e_i^{(m)} = \parallel_{h=1}^H \sigma \left(\sum_{j \in N_i^{(m)}(h)} \gamma_{ij}^{(m)(h)} h_j^{(m)(h)} \right) \quad (5)$$

In this equation, H denotes the total number of attention heads, the superscript (h) indicates the current attention head index, and \parallel represents concatenating outputs from different heads. Through this strategy, distinct node-specific information is captured across various dimensions and integrated from multiple subspaces, thereby significantly improving the precision and robustness of the learned node representations.

During the training process, each medication entity is updated according to the aforementioned steps, resulting in an aggregated representation of the medication set $V_m = \{v_m^1, v_m^2, \dots, v_m^{t-1}\}$.

Similarly, to capture the unique clinical relationships of hypertension, such as the interactions between comorbidities, long-term monitoring data, and combination therapy dynamics, we defined 3 heterogeneous complete graphs: diagnosis-procedure graph $G_{dp}(t) = \{D(t), P(t), A^{dp}\}$, medication-diagnosis graph $G_{md}(t) = \{M(t-1), D(t), A^{md}\}$, and medication-procedure graph $G_{mp}(t) = \{M(t-1), P(t), A^{mp}\}$. This established an association model that is crucial for accurate hypertension medication recommendations but has low correlation with other medication categories. Following a similar learning approach to the medication graph G_m , we first initialized the diagnosis and procedure nodes in the 3 heterogeneous graphs.

$$e_d = V_d^t E_d, e_p = V_p^t E_p \quad (6)$$

Using the approach outlined in equations (3)-(5), we then learned from the diagnosis-procedure graph to obtain aggregated sets of diagnosis codes $V_d = \{v_d^1, v_d^2, \dots, v_d^t\}$ and procedure nodes $V_p = \{v_p^1, v_p^2, \dots, v_p^t\}$.

Finally, based on the aggregated medication nodes from the medication graph G_m and the learned diagnosis and procedure nodes from the diagnosis-procedure graph G_{dp} , we learned from the medication-diagnosis graph G_{md} and the medication-procedure graph G_{mp} , updating to obtain entity-level representations of the medication, diagnosis, and procedure sets:

$$\begin{aligned} V_D &= \{V_d^1, V_d^2, \dots, V_d^t\} \\ V_P &= \{V_p^1, V_p^2, \dots, V_p^t\} \\ V_M &= \{V_m^1, V_m^2, \dots, V_m^{t-1}\} \end{aligned} \quad (7)$$

It is noteworthy that this study used a multihead GAT mechanism to learn and update nodes within the graph. Although traditional GAT models consider intrinsic features and aggregate neighboring features, our proposed heterogeneous graph-based method places greater emphasis on enhancing the embeddings of medications, diagnoses, and procedures by explicitly modeling interactions among diverse

medical entities. Consequently, our model focuses more intensively on neighboring nodes to prevent excessive node information merging. To this end, we modified the standard GAT approach to explicitly focus on information derived from neighboring nodes within the heterogeneous graph learning process.

Temporal Selectivity Module

To effectively model the temporal evolution of patient health conditions, this paper introduces a bidirectional temporal selection mechanism using gated recurrent units (GRUs), with 3 key innovations that distinguish it from conventional applications.

First, we adopted a bidirectional temporal architecture to capture multiscale temporal dependencies. Specifically, we first used GRU^α to learn the diagnostic sequence, generating forward diagnostic selection coefficients that emphasize the impact of historical visits on current states (eg, past hypertensive crisis records influencing present medication titration):

$$g^1, g^2, \dots, g^t = GRU_d^\alpha(V_d^1, V_d^2, \dots, V_d^t) \quad (8)$$

$$\alpha^j = \tanh(W^\alpha g^j + b^\alpha), j = 1, \dots, t \quad (9)$$

Concurrently, a backward GRU^β was used in the reverse temporal order to learn the diagnostic sequence, generating backward selection coefficients at different time steps that highlight recent critical changes. For example, recent fluctuations in blood pressure require immediate therapeutic adjustment. This bidirectional design enhances computational stability while overcoming the limitations of unidirectional GRUs or traditional recurrent neural networks (RNNs), which often overlook either long-term or short-term temporal cues:

$$h^t, h^{t-1}, \dots, h^1 = GRU_d^\beta(V_d^t, V_d^{t-1}, \dots, V_d^1) \quad (10)$$

$$\beta^j = \tanhshrink(W^\beta h^j + b^\beta), j = t, \dots, 1 \quad (11)$$

Second, we proposed an adaptive selection coefficient integration strategy. Based on the generated bidirectional diagnostic selection coefficients, we can capture key visit information and entity information within the visit sequence, rather than relying on static aggregation methods. This allowed us to capture key visit information and entity interactions, thereby obtaining a diagnostic representation that integrates historical context with current needs:

$$d^t = \sum_{j=1}^t \alpha^j \beta^j \odot V_d^j \quad (12)$$

Third, leveraging GRU's inherent advantages, our design achieved computational efficiency without sacrificing performance. Compared with traditional RNNs, GRU effectively mitigates gradient vanishing issues; relative to LSTMs, its simplified gating mechanism reduces parameter complexity by avoiding redundant memory cells, resulting in faster training efficiency, a critical advantage for handling large-scale longitudinal EHR data in medication

recommendation tasks. After a series of similar processing steps, we obtained the patient's final procedural and medication representations.

$$p^t = \sum_{j=1}^t \alpha^j \beta^j \odot V_p^j, m^{t-1} = \sum_{j=1}^{t-1} \alpha^j \beta^j \odot V_m^j \quad (13)$$

Interaction Prediction Module

Based upon the outputs generated by the aforementioned modules, for the patient's t -th visit, we concatenated the medical entity sequence to make medication recommendations:

$$y^t = \sigma([d^t; p^t; m^{t-1}]) \quad (14)$$

In our approach, the medication recommendation task is formulated as a multilabel classification problem [25,26]. To address the complexities and potential imbalances in medical datasets, we used a comprehensive strategy during model training. We enhanced the model's generalization capabilities and mitigated overfitting through the use of regularization techniques, which limit the complexity of the parameters learned. Additionally, an early-stopping mechanism was implemented to curtail training based on validation set performance. To optimize parameters, we used the Adam optimization algorithm [27], which minimizes the binary cross-entropy loss function to promote efficient and stable model convergence.

$$L = - \sum_{t=1}^T \sum_{i=1}^{|M|} y_i^t \log \hat{y}_i^t + (1 - y_i^t) \log(1 - \hat{y}_i^t) \quad (15)$$

Ethical Considerations

This study made use of the standardized, publicly available MIMIC-III and MIMIC-IV datasets from the Massachusetts Institute of Technology [23,24] and was therefore deemed exempt from ethical approval requirements. Prior to their release, these datasets underwent comprehensive ethical review and privacy protection processes conducted by the data provider (Massachusetts Institute of Technology). These processes included deidentifying all personally identifiable information in patients' EHRs, such as names, hospital admission numbers, and dates of birth. Additionally, the datasets' usage license explicitly covers secondary analysis scenarios for academic research, eliminating the need for users to obtain additional ethical approval independently. Since this study did not involve independent collection of human subject data and solely relied on the aforementioned publicly available and compliant existing datasets for secondary analysis, we do not possess separate ethical approval documents to provide. The ethical approval statement is published on the official website of the data provider [23,24], and we can be contacted to obtain the original copy of the datasets' usage license.

Results

Dataset Description

Our experiments were conducted using the MIMIC-III v1.4 and MIMIC-IV v2.2 datasets provided by the Massachusetts Institute of Technology. The datasets comprise medical records collected from patients admitted to intensive care units, such as diagnoses, procedures, and medications. The diagnostic and procedural information uses the *International Classification of Diseases, Ninth Revision (ICD-9)* coding system. To study medication recommendations for patients with hypertension, this paper extracted relevant data from the aforementioned datasets based on the *ICD-9* hypertension codes under the guidance of clinical experts. Meanwhile, referring to previous research norms [4,21,22], patients who completed at least two visits were included, and medications with a frequency <2000 occurrences were excluded.

During data analysis, we observed that the number of frequently used antihypertensive medications was relatively limited and the recommendations based on overly broad Anatomical Therapeutic Chemical (ATC) categories often

lacked the specificity needed for precise clinical decision-making. Given the systematic medication classification framework of the ATC classification system, this study overcame the limitation of previous studies that only focused on the ATC04 level for recommendation. It not only predicted the medication categories at the ATC04 level (such as C02A, anti-adrenergic medications with central effects) but also further refined to the specific medication types at the ATC05 level (such as C02AA, reserpine-like medications).

After data extraction, we carried out meticulous preprocessing. We normalized features to standardize the scale of different variables, which prevented certain features from dominating others during model training. Moreover, we selectively chose variables most relevant to hypertension medications. Finally, the preprocessed dataset was partitioned into training, validation, and testing subsets according to a ratio of $\frac{2}{3}:\frac{1}{6}:\frac{1}{6}$.

Detailed information about the dataset used in the experiments and examples from patients are shown in Table 1 and Table 2.

Table 1. Detailed information for the experimental datasets, by Anatomical Therapeutic Chemical (ATC)-level encoding.

Item	MIMIC-III ^a		MIMIC-IV	
	ATC04	ATC05	ATC04	ATC05
Patients, n	3115	3109	19,609	19,609
Visits, n	7308	7263	55,239	55,236
Diagnoses, n	1966	1965	2000	2000
Procedures, n	1145	1145	5488	5488
Medications, n	14	18	14	18
Number of visits, mean	2.3460	2.3361	2.8169	2.8169
Number of diagnoses, mean	10.8927	10.9262	9.3666	9.3666
Number of procedures, mean	4.0463	4.0609	2.5492	2.5492
Number of medications, mean	1.5508	1.4400	1.0804	1.1352

^aMIMIC: Medical Information Mart for Intensive Care.

Table 2. Samples from electronic health records.

Sub_ID	Hadm_ID	Diagnoses (Anatomical Therapeutic Chemical code)	Procedures	Medications
10001217	24597018	3240, 3484, 3485, 5180, 340, 04109, 3051, 4019, V168, V161	139, 331, 3897	HydrALAzine, LeVETiracetam, Vancomycin, Bisacodyl, Meropenem,...
10001217	27703517	3240, 3485, 340, 04102, 04184, 4019, 3051	139	HydrALAzine, Vancomycin, Meropenem, Bisacodyl, Lidocaine,...

Evaluation Metrics

To validate the effectiveness of CSRec, the evaluation metrics described in the following sections were used.

Jaccard Similarity Coefficient

A higher Jaccard coefficient reflects greater overlap between the predicted medication set and the actual medication set.

$$Jaccard = \frac{1}{T} \sum_{t=1}^T \frac{|y^t \cap \hat{y}^t|}{|y^t \cup \hat{y}^t|} \quad (16)$$

Area Under the Curve of Precision-Recall

A high area under the curve of precision-recall (PRAUC) indicates that the model recommends appropriate medications while keeping a low error rate.

$$\Delta \mathcal{R}call(i)^t = \mathcal{R}call(i)^t - \mathcal{R}call(i-1)^t \quad (17)$$

$$PRAUC = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{|M|} Precision(i)^t \Delta \mathcal{R}call(i)^t \quad (18)$$

F₁-Score

The elevated F_1 -score metric in medication recommendations signifies an optimal equilibrium between minimizing false negatives and maintaining classification specificity.

$$Precision^t = \frac{|y^t \cap \hat{y}^t|}{|\hat{y}^t|}, \quad Recall^t = \frac{|y^t \cap \hat{y}^t|}{|y^t|}$$

(19)

$$F1 = \frac{1}{T} \sum_{t=1}^T \frac{2 \times Precision^t \times Recall^t}{Precision^t + Recall^t}$$

(20)

DDI Rate

A lower DDI rate ensures that the recommended combination of medications is safer in clinical practice.

$$DDI = \frac{1}{T} \sum_{t=1}^T \frac{\sum_{i=1}^{|\hat{y}^t|} \sum_{j=i+1}^{|\hat{y}^t|} 1\{A_d[\hat{y}_i^t, \hat{y}_j^t]\}}{\sum_{i=1}^{|\hat{y}^t|} \sum_{j=i+1}^{|\hat{y}^t|} 1}$$

(21)

Comparative Models

The following baseline models were selected for comparison with the proposed CSRec:

- Logistic regression (LR) uses L2 regularization.
- Ensemble of classifier chains (ECC) [28] enhances predictions by connecting multiple classifiers, typically applied to multilabel classification scenarios for performance optimization.
- RETAIN [18] implements bidirectional temporal attention architecture designed for sequential clinical prediction tasks like treatment recommendation.
- Learn to Prescribe (LEAP) [29] uses RNNs to extract meaningful representations during current medical visits and generate medication sequences.

- Dual memory neural computer (DMNC) [20] incorporates dual-memory neural components to model asynchronous therapeutic pattern interactions.
- GAMENet [11] adopts graph-based memory architecture combining medication interaction knowledge and querying longitudinal EHR data for medication retrieval.
- MICRON [21] analyzes EHR temporal dynamics to adaptively optimize medication combinations upon symptom evolution.
- SafeDrug [16] evaluates molecular-level medication-patient compatibility to suggest safer therapeutic regimens.
- COGNet [22] implements medication copy or predict strategy integrating historical effective prescriptions into current recommendations.
- MoleRec [14] predicts medication mechanisms and interactions via molecular-patient relationships for personalized, safer recommendations.
- CausalMed [9] identifies medical entity causal relationships via causal discovery, accounting for dynamic health condition differences to generate causally linked recommendations.

Performance Comparison

In this section, we conducted an extensive comparative analysis between CSRec and the previously described baseline medication recommendation models to evaluate its effectiveness. Experimental results on the MIMIC-III v1.4 and MIMIC-IV 2.2 datasets are shown in Table 3 (“↑” indicates a preference for larger values, “↓” indicates a preference for smaller values).

Table 3. The comparison of experimental results on datasets of Anatomical Therapeutic Chemical (ATC)04 and ATC05 codes.

Model	Jaccard similarity coefficient ↑ ^a		PRAUC ↑ ^b		F ₁ -score ↑		DDI ↓ ^{c,d}	
	ATC04	ATC05	ATC04	ATC05	ATC04	ATC05	ATC04	ATC05
MIMIC-III ^e								
LR ^f	0.5203	0.4919	0.8117	0.7726	0.6347	0.5972	0.2951	0.3569
ECC ^g	0.5438	0.5056	0.8278	0.7846	0.6593	0.6127	0.3012	0.3979
RETAIN ^h	0.5540	0.5210	0.8143	0.7874	0.6783	0.6449	0.3921	0.5068
LEAP ⁱ	0.5228	0.5082	0.6346	0.6162	0.6405	0.6217	0.2627	0.3648
DMNC ^j	0.5342	0.5097	0.8225	0.7696	0.6463	0.6277	0.2365	0.2615
GAMENet ^k	0.5497	0.5169	0.8239	0.7861	0.6710	0.6308	0.2366	0.4199
MICRON	0.5352	0.5146	0.8285	0.7884	0.6423	0.6145	0.2357	0.2658
SafeDrug	0.5474	0.5063	0.8252	0.7764	0.6484	0.6146	0.2329 ^l	0.1975 ^l
COGNet	0.5325	0.5075	0.7972	0.7762	0.6448	0.6167	0.2902	0.3587
MoleRec	0.5501	0.5189	0.8261	0.7825	0.6716	0.6299	0.2709	0.3491
CausalMed	0.5626	0.5340	0.8226	0.7877	0.6741	0.6394	0.2340	0.2679
CSRec	0.5801 ^l	0.5534 ^l	0.8356 ^l	0.8123 ^l	0.6895 ^l	0.6602 ^l	0.2351	0.2663
MIMIC-IV								

Model	Jaccard similarity coefficient ↑ ^a		PRAUC ↑ ^b		F_1 -score ↑		DDI ↓ ^{c,d}	
	ATC04	ATC05	ATC04	ATC05	ATC04	ATC05	ATC04	ATC05
LR	0.4248	0.3518	0.7426	0.6873	0.5125	0.4315	0.2821	0.2734
ECC	0.4389	0.3779	0.7498	0.6923	0.5305	0.4624	0.3064	0.3047
RETAIN	0.5349	0.5045	0.7964	0.7641	0.6356	0.6008	0.3431	0.3197
LEAP	0.4186	0.3884	0.5047	0.4683	0.5326	0.4932	0.0104 ⁱ	0.0506 ⁱ
DMNC	0.5139	0.4787	0.7678	0.7277	0.5392	0.5094	0.2801	0.2792
GAMENet	0.5151	0.4738	0.7781	0.7326	0.6069	0.5662	0.2884	0.3050
MICRON	0.4297	0.3716	0.7543	0.7056	0.5151	0.4431	0.2866	0.2557
SafeDrug	0.5173	0.4767	0.7770	0.7313	0.6109	0.5648	0.3084	0.3255
COGNet	0.5087	0.4568	0.7448	0.7009	0.5948	0.5502	0.3261	0.3835
MoleRec	0.5275	0.4838	0.7624	0.7267	0.6262	0.5801	0.2733	0.2993
CausalMed	0.5440	0.5225	0.7812	0.7509	0.6412	0.6099	0.2677	0.2855
CSRec	0.5582 ^l	0.5307 ^l	0.8069 ^l	0.7655 ^l	0.6483 ^l	0.6195 ^l	0.2760	0.2931

^a ↑ indicates a preference for larger values.
^b PRAUC: area under the curve of precision-recall.
^c DDI: drug-drug interaction.
^d ↓ indicates a preference for smaller values.
^e MIMIC: Medical Information Mart for Intensive Care.
^f LR: logistic regression.
^g ECC: ensemble of classifier chains.
^h RETAIN: Reverse Time Attention model.
ⁱ LEAP: Learn to Prescribe.
^j DMNC: dual memory neural computer.
^k GAMENet: Graph-Augmented Memory Networks.
^l Optimal data.

After detailed analysis, we observed several key findings, as described in the following paragraphs.

The CSRec model proposed in this paper outperformed most comparative models across multiple evaluation metrics, fully demonstrating its significant effectiveness in hypertension medication recommendation.

LR and ECC performed poorly, primarily because these methods only focus on the patient’s current clinical condition, neglecting the influence of historical medical data on present treatment decisions. In contrast, the model presented in this paper, along with RETAIN, GAMENet, and other models integrating longitudinal medical histories, performed better, highlighting the importance of capturing historical medical information in hypertension medication recommendations.

In comparison with longitudinal models such as RETAIN, GAMENet, and COGNet, CSRec maintained superior performance, with a particularly notable enhancement in Jaccard scores. This superiority can be attributed to 2 core architectural advancements. The Temporal Selectivity Module enables accurate simulation of patient disease progression by assigning enhanced weights to clinically critical historical information, thereby ensuring that recommendations adhere to consistent clinical reasoning frameworks. Meanwhile, the Heterogeneous Collaborative Module facilitates effective capture of synergistic relationships among heterogeneous medical entities (diagnoses, procedures, medications), enabling the identification of medication combinations that are clinically coherent and reflective

of real-world coprescription patterns. Collectively, these innovations augment the degree of overlap between recommended and actually prescribed medications, as empirically validated by the elevated Jaccard index.

Although the CSRec model introduced in this study does not achieve the highest score on the DDI metric, it still ranks as suboptimal. This is mainly because the best performing LEAP model only considers the patient’s present health status, resulting in the recommendation of fewer medications and consequently yielding the lowest DDI scores.

Ablation Study

To systematically assess the contribution and validity of each module within CSRec, ablation experiments were conducted on MIMIC-IV by comparing CSRec with its variants:

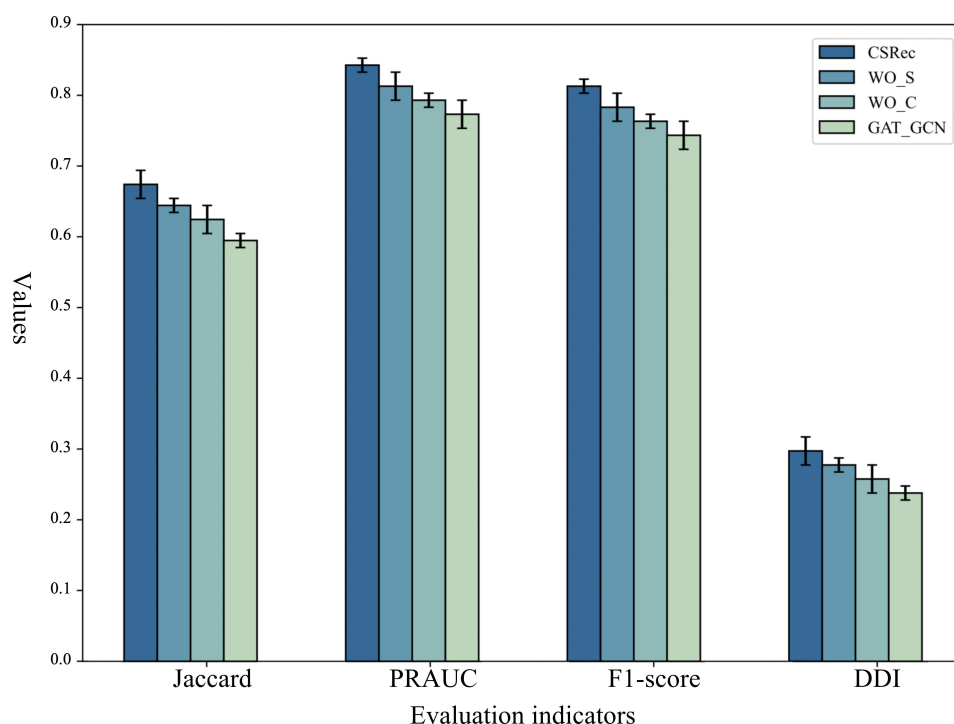
- WO_S removes the selection module based on the cyclic mechanism.
- WO_C removes the collaborative module based on the heterogeneous graph.
- GAT_GCN replaces the multihead GAT in the collaborative module with a graph convolutional network (GCN).

The results displayed in Figure 3 indicate that the performance of CSRec declined when any module is removed or replaced, demonstrating that each component of CSRec is indispensable. The observed performance drops in evaluation metrics (Jaccard, PRAUC, F_1 -score, and DDI) can be attributed to the unique functional roles of each module: The collaborative module effectively modeled the

correlations between different medical events during each visit, and its removal (WO_C) led to incomplete capture of event associations, resulting in a decrease in recommendation accuracy as medical entity representations lost contextual relevance. The selective module globally modeled the patient's historical medical data, and its absence (WO_S) caused the model to fail at emphasizing critical time points in disease progression, leading to a decline in metrics

as the temporal continuity of hypertension development was disrupted. When replacing multihead GAT with GCN (GAT_GC_N), the loss of the attention mechanism resulted in insufficient differentiation of important medical entity relationships, reducing recommendation precision compared with the original model as convolutional operations cannot dynamically weight heterogeneous graph information.

Figure 3. The ablation study of different components. DDI: drug-drug interaction; GAT_GC_N: variant replacing the multihead graph attention network in the collaborative module with a graph convolutional network; Jaccard: Jaccard similarity coefficient; PRAUC: area under the curve of precision-recall; WO_C: variant removing the collaborative module based on the heterogeneous graph; WO_S: variant removing the selection module based on the cyclic mechanism.



Additionally, this study further explored the impact of different types of collaborative medical entity information (diagnosis-medication, procedure-medication, diagnosis-procedure, and medication-medication) on hypertension medication recommendations. The variants were designed as follows:

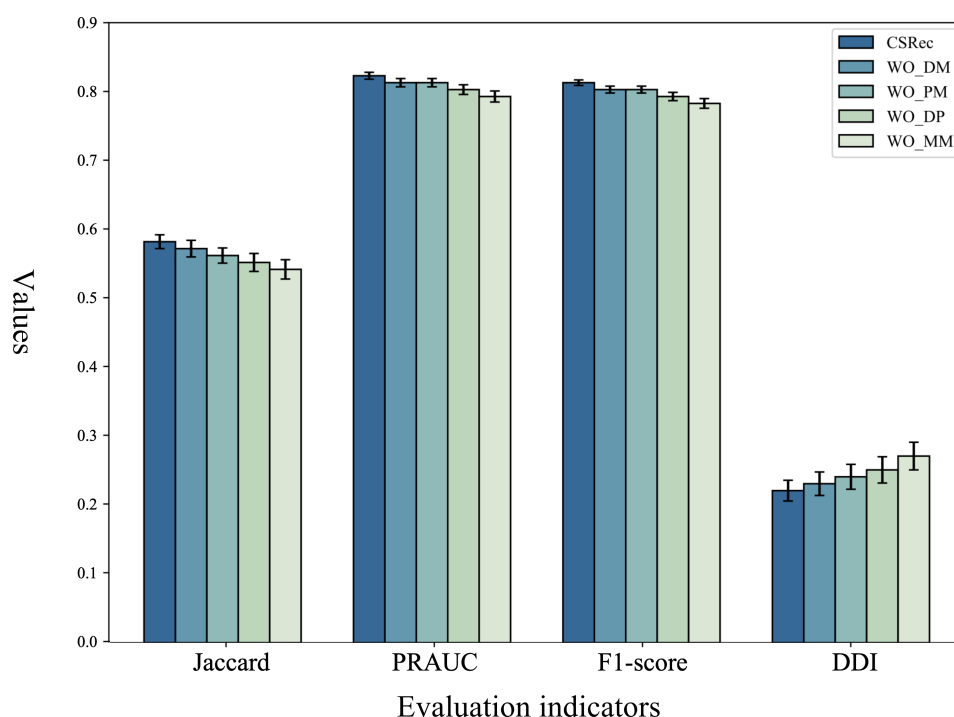
- WO_DM removes diagnosis-medication interaction information.
- WO_PM removes procedure-medication interaction information.
- WO_DP removes diagnosis-procedure interaction information.
- WO_MM removes medication co-occurrence and interaction information.

As illustrated in Figure 4, CSRec's performance decreased when any single type of collaborative medical entity information (WO_DM, WO_PM, WO_DP) was removed, confirming the strong correlation between such information and the medication recommendation task. The specific decline in metric degree corresponded to the clinical

relevance of each information type: Removing diagnosis-medication interactions (WO_DM) reduced metrics, as diagnostic information directly guides first-line hypertension medication selection. The WO_PM group's drop stemmed from lost procedure-related medication adjustments (eg, postinterventional anticoagulation needs). For WO_DP, the decline reflects disrupted diagnosis-procedure logical chains that inform therapeutic medication choices.

Notably, removing medication co-occurrence and interaction information (WO_MM) caused a significant decline in the DDI risk rate of the medication recommendation performance. This substantial change is explained by the frequent use of combination therapy in hypertension management. Without medication interaction data, the model cannot avoid contraindicated medication pairs, while the loss of co-occurrence patterns reduces the accuracy of synergistic medication recommendations. This underscores the essential role of medication co-occurrence and DDI information for ensuring hypertension treatment efficacy and safety.

Figure 4. The ablation study of different medical entities. DDI: drug-drug interaction; Jaccard: Jaccard similarity coefficient; PRAUC: area under the curve of precision-recall; WO_DM: variant removing diagnosis-medication interaction information; WO_DP: variant removing diagnosis-procedure interaction information; WO_MM: variant removing medication co-occurrence and interaction information; WO_PM: variant removing procedure-medication interaction information.



Parameter Sensitivity

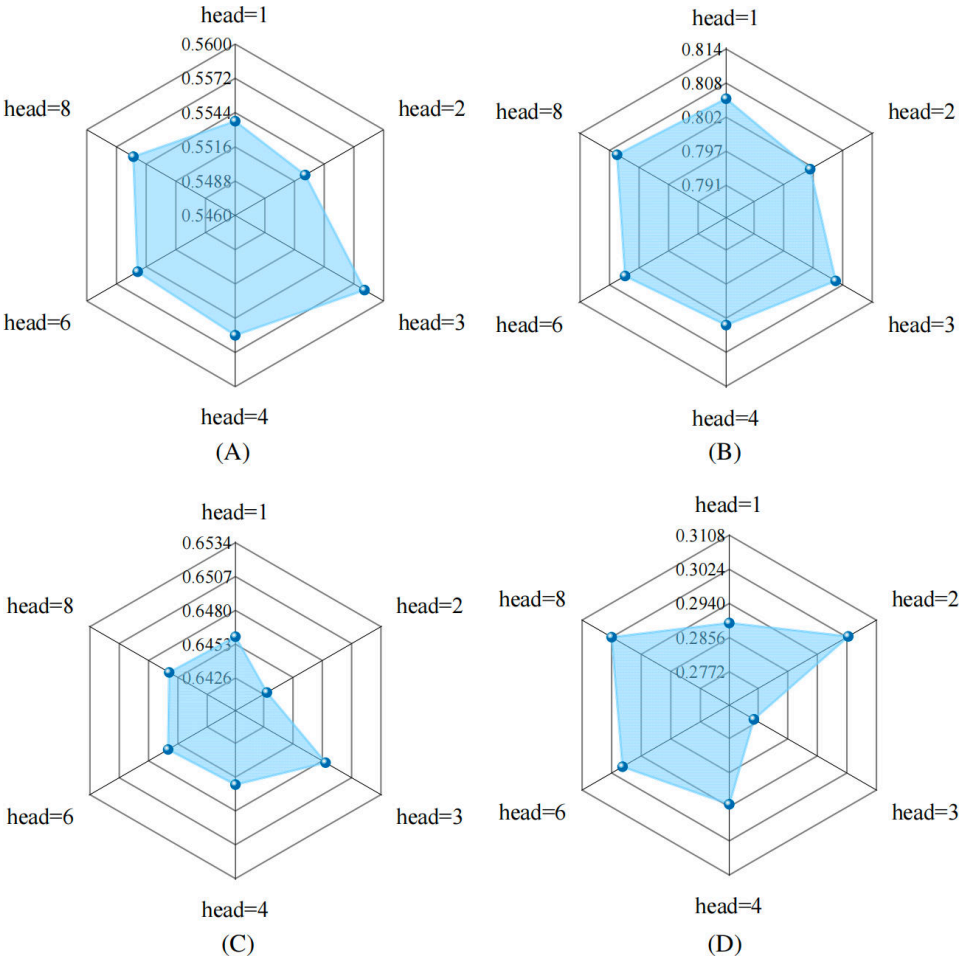
In order to investigate how varying the number of attention heads in the multihead graph attention mechanism within the heterogeneous collaborative module impacted model efficacy, multiple experiments were conducted on the MIMIC-IV dataset. The outcomes of these experiments, comparing different head counts, are presented in [Figure 5](#).

A setting of 3 attention heads (head=3) achieved optimal performance. Consequently, this configuration (head=3) was adopted consistently throughout the remaining experiments

in this study. Moreover, the experimental results illustrated that variations in the number of attention heads led to minimal performance differences, highlighting the stability and reliability of the proposed model.

Furthermore, to investigate the impact of the number of included hypertension medications on model performance, we conducted a sensitivity analysis by varying the minimum occurrence threshold for medication inclusion from 500 to 3000. The number of medications ranged from 26 (≥ 500) to 9 (≥ 3000).

Figure 5. The effect of different numbers of heads in the multihead graph attention network (GAT), as measured using (A) the Jaccard similarity coefficient, (B) area under the curve of precision-recall (PRAUC), (C) F_1 -score, and (D) drug-drug interaction (DDI) rate.



As shown in Table 4, performance across all 4 metrics—Jaccard, PRAUC, F_1 -score, and DDI—remained remarkably stable, with only marginal fluctuations. Notably, the best performance was observed at the 2000 threshold, which is used in our paper. This finding supported our design choice and indicated that our framework remained robust even under different medication inclusion criteria, alleviating concerns

regarding limited sample size. In addition, setting a lower threshold introduced lower-frequency medications that may lack sufficient clinical representation, potentially compromising the statistical reliability of training. Thus, the current configuration struck a reasonable trade-off between medication variety and prediction stability.

Table 4. Sensitivity analysis of the medication frequency threshold.

Minimum medication occurrence	Medications, n	Jaccard similarity coefficient \uparrow^a	PRAUC \uparrow^b	F_1 -score \uparrow	DDI $\downarrow^{c,d}$
≥ 500	26	0.5531	0.7983	0.6428	0.2847
≥ 1000	20	0.5554	0.8027	0.6456	0.2802
≥ 1500	17	0.5576	0.8049	0.6471	0.2781
≥ 2000	14	0.5582 ^e	0.8069 ^e	0.6483 ^e	0.2760 ^e
≥ 2500	12	0.5568	0.8051	0.6475	0.2766
≥ 3000	9	0.5560	0.8050	0.6452	0.2764

^a \uparrow indicates a preference for larger values.

^b PRAUC: area under the curve of precision-recall.

^c DDI: drug-drug interaction.

^d \downarrow indicates a preference for smaller values.

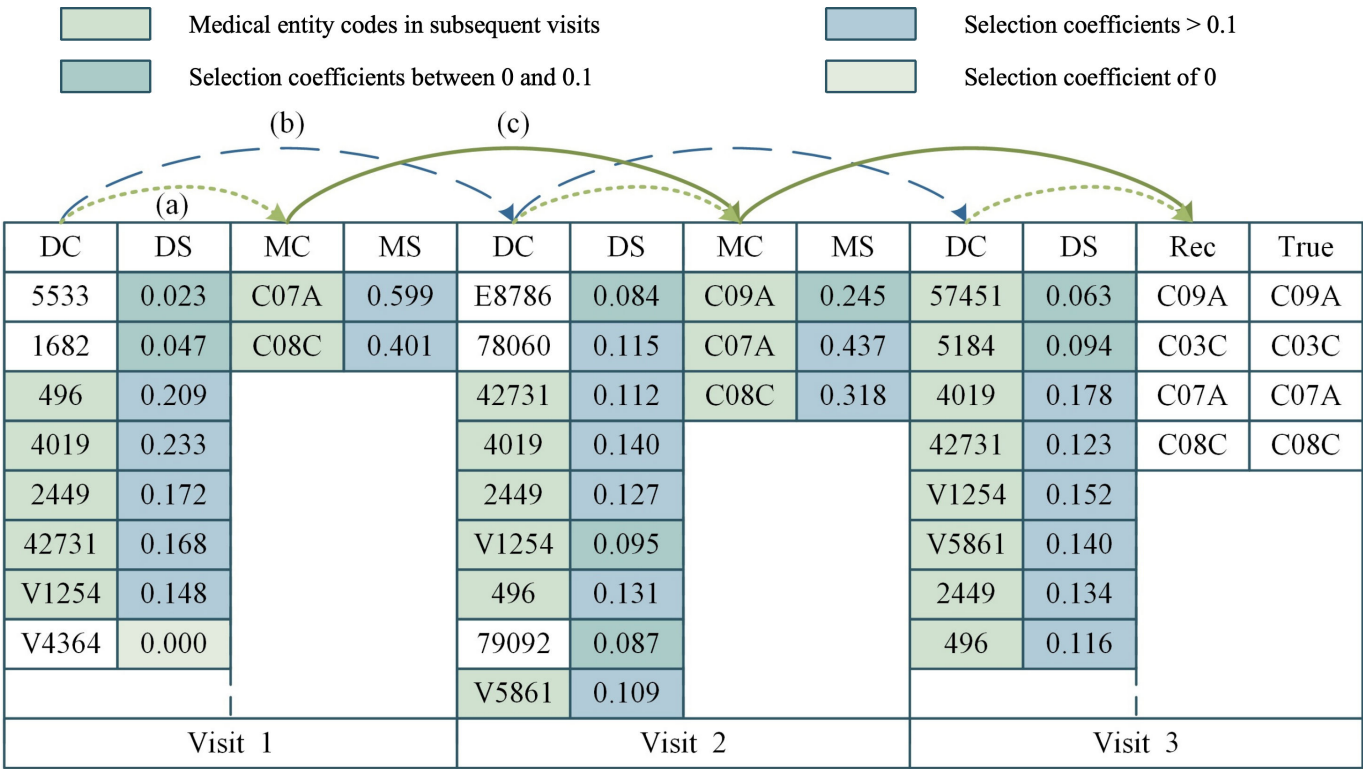
^e Optimal data.

Case Study

To visualize and validate our core innovations, which included heterogeneous collaborative interaction modeling and the temporal selectivity mechanism, we randomly

selected a patient from the test set of the MIMIC-IV dataset for a case study. Taking diagnosis and medication entities as examples, the entire learning process is shown in [Figure 6](#).

Figure 6. Case study. DC: diagnosis code; DS: diagnosis selection coefficient; MC: medication code; MS: medication selection coefficient; Rec: recommended medication; True: prescribed medication.



Initially, in the collaborative process (a), the model’s collaborative interaction mechanism was showcased: In the visit-specific level, the model dynamically enhanced the integration of diagnostic information into medication recommendation reasoning by assigning context-aware interaction weights to the interplays between diagnostic codes (DC column) and medication codes (MC column). For the first visit of the illustrative patient, the DC column encompasses comorbidity-related codes, including chronic obstructive pulmonary disease (*ICD-9*: 496), hypothyroidism (*ICD-9*: 2449), atrial fibrillation (*ICD-9*: 42731), and a documented history of cardiovascular disease (*ICD-9*: V1254), while the MC column includes hypertension-targeted agents such as selective β 1-receptor blockers (ATC-04: C07A) and dihydropyridine calcium channel blockers (ATC-04: C08C). This aligns with our innovation of capturing disease-specific collaborative patterns, evident in the consistently higher selection coefficients for hypertension-related diagnostic codes in each DC column, which prioritizes the disease context critical for hypertension treatment.

Furthermore, to demonstrate our temporal selectivity mechanism, the case study highlights how the model captures longitudinal EHR dependencies. Given hypertension’s chronic nature, recommended medications depend

on both current (third visit) diagnoses (DC column) and historical data. As shown in step (b), diagnostic codes from previous visits received higher selection coefficients, while step (c) reveals that historical medication codes were similarly prioritized. This reflects our model’s ability to dynamically weight temporal information, addressing the evolving nature of hypertension and representing an advantage over static models.

Discussion

This paper presents the CSRec model for hypertension medication recommendation, which is based on the collaboration and selection of heterogeneous medical entities. By effectively capturing relationships among diverse medical entities and combining the temporal evolution characteristics of clinical entities, the model provides more precise and effective medication guidance for hypertension treatment. Future research will focus on 3 main areas: (1) in-depth exploration of the multilabel imbalance problem existing in the current method, (2) further investigation of the model’s performance in addressing the cold-start problem, and (3) exploration of recommendations regarding medication dosages and medication types.

Acknowledgments

This research is supported by the Industrial Support Program of Colleges and Universities in Gansu Province (number 2025CYZC-010); National Natural Science Foundation of China (number 62163033); Talent Innovation and Entrepreneurship Project of Lanzhou, China (number 2021-RC-49); Natural Science Foundation of Gansu Province, China (numbers 22JR5RA145, 21JR7RA781, and 21JR7RA116); and Major Research Project Incubation Program of Northwest Normal University, China (number NWNLU-LKZD2021-06). The funding agencies did not participate in the study design, data gathering and analysis, decision-making regarding publication, nor manuscript preparation.

Data Availability

The code for the model is available on GitHub [30].

Conflicts of Interest

None declared.

References

1. Hypertension Branch of Chinese Geriatrics Society, Beijing Hypertension Association, National Clinical Research Center of the Geriatric Diseases, Hua Q, Fan L, Wang ZW, Li J. 2023 guideline for the management of hypertension in the elderly population in China. *J Geriatr Cardiol*. Jun 28, 2024;21(6):589-630. [doi: [10.26599/1671-5411.2024.06.001](https://doi.org/10.26599/1671-5411.2024.06.001)] [Medline: [38973827](https://pubmed.ncbi.nlm.nih.gov/38973827/)]
2. The Writing Committee of the Report on Cardiovascular Health and Diseases in China. Interpretation of report on cardiovascular health and diseases in China 2022. *Chin J Cardiovasc Med*. 2023;28(4):297-312. [doi: [10.12114/j.issn.1007-9572.2023.0408](https://doi.org/10.12114/j.issn.1007-9572.2023.0408)]
3. Wu H, Xie HW. Research on hypertension diagnosis and treatment system based on ontology and CBR. *Comput Appl Softw*. 2013;30(12):155-159. [doi: [10.3969/j.issn.1000-386x.2013.12.041](https://doi.org/10.3969/j.issn.1000-386x.2013.12.041)]
4. Li X, Liang SP, Hou YL, Ma TF. StratMed: relevance stratification between biomedical entities for sparsity on medication recommendation. *Knowl Based Syst*. Jan 2024;284:111239. [doi: [10.1016/j.knosys.2023.111239](https://doi.org/10.1016/j.knosys.2023.111239)]
5. Zhang HJ, Yang X, Bai L, Liang JY. Enhancing drug recommendations via heterogeneous graph representation learning in EHR networks. *IEEE Trans Knowl Data Eng*. 2024;36(7):3024-3035. [doi: [10.1109/TKDE.2023.3329025](https://doi.org/10.1109/TKDE.2023.3329025)]
6. Choi E, Xiao C, Stewart WF, Sun J. MiME: multilevel medical embedding of electronic health records for predictive healthcare. Presented at: 32nd International Conference on Neural Information Processing Systems; Dec 3-8, 2018; Montréal, Canada. URL: <https://dl.acm.org/doi/10.5555/3327345.3327366> [Accessed 2025-09-25]
7. Choi E, Xu Z, Li Y, et al. Learning the graphical structure of electronic health records with graph convolutional transformer. *AAAI*. 2020;34(1):606-613. [doi: [10.1609/aaai.v34i01.5400](https://doi.org/10.1609/aaai.v34i01.5400)]
8. Hosseini A, Chen T, Wu W, et al. HeteroMed: heterogeneous information network for medical diagnosis. Presented at: 27th ACM International Conference on Information and Knowledge Management; Oct 22-26, 2018; Turin, Italy. [doi: [10.1145/3269206.3271805](https://doi.org/10.1145/3269206.3271805)]
9. Li X, Liang S, Lei Y, et al. CausalMed: causality-based personalized medication recommendation centered on patient health state. Presented at: CIKM '24: The 33rd ACM International Conference on Information and Knowledge Management; Oct 21-25, 2024; Boise, ID, USA. [doi: [10.1145/3627673.3679542](https://doi.org/10.1145/3627673.3679542)]
10. Poulain R, Beheshti R. Graph transformers on EHRs: better representation improves downstream performance. Presented at: The Twelfth International Conference on Learning Representations; May 7-11, 2024; Vienna, Austria. URL: <https://openreview.net/forum?id=pe0Vdv7rsL> [Accessed 2025-10-12]
11. Shang JY, Xiao C, Ma TF, Li HY, Sun JM. GAMENet: graph augmented memory networks for recommending medication combination. *Proc AAAI Conf Artif Intell*. 2019;33(1):1126-1133. [doi: [10.1609/aaai.v33i01.33011126](https://doi.org/10.1609/aaai.v33i01.33011126)]
12. Shang JY, Ma TF, Xiao C, Sun JM. Pre-training of graph augmented transformers for medication recommendation. Presented at: Twenty-Eighth International Joint Conference on Artificial Intelligence; Aug 10-16, 2019; Macao, China. [doi: [10.24963/ijcai.2019/825](https://doi.org/10.24963/ijcai.2019/825)]
13. Liang X, Yang JY, Lu GM, Zhang D. CompNet: competitive neural network for palmprint recognition using learnable gabor kernel. *IEEE Signal Process Lett*. 2021;28:1739-1743. [doi: [10.1109/LSP.2021.3103475](https://doi.org/10.1109/LSP.2021.3103475)]
14. Yang N, Zeng K, Wu Q, Yan J. MoleRec: combinatorial drug recommendation with substructure-aware molecular representation learning. Presented at: WWW '23: The ACM Web Conference 2023; Apr 30 to May 4, 2023; Austin, TX. [doi: [10.1145/3543507.3583872](https://doi.org/10.1145/3543507.3583872)]
15. Li XW, Zhang YJ, Li XB, Wei H, Lu MY. DGCL: distance-wise and graph contrastive learning for medication recommendation. *J Biomed Inform*. Mar 2023;139:104301. [doi: [10.1016/j.jbi.2023.104301](https://doi.org/10.1016/j.jbi.2023.104301)]

16. Yang C, Xiao C, Ma F, Glass L, Sun J. SafeDrug: dual molecular graph encoders for recommending effective and safe drug combinations. Presented at: Thirtieth International Joint Conference on Artificial Intelligence; Aug 19-27, 2021; Montreal, Canada. [doi: [10.24963/ijcai.2021/514](https://doi.org/10.24963/ijcai.2021/514)]
17. Liu SC, Wang XL, Xiang Y, Xu H, Wang H, Tang BZ. Multi-channel fusion LSTM for medical event prediction using EHRs. J Biomed Inform. Mar 2022;127:104011. [doi: [10.1016/j.jbi.2022.104011](https://doi.org/10.1016/j.jbi.2022.104011)]
18. Choi E, Bahadori MT, Kulas JA, Schuetz A, Stewart WF, Sun JM. RETAIN: an interpretable predictive model for healthcare using reverse time attention mechanism. Presented at: 30th International Conference on Neural Information Processing Systems (NIPS'16); Dec 5-10, 2016; Barcelona, Spain. URL: <https://dl.acm.org/doi/10.5555/3157382.3157490> [Accessed 2025-09-25] [doi: [10.5555/3157382.3157490](https://doi.org/10.5555/3157382.3157490)]
19. An Y, Zhang L, You M, Tian XQ, Jin B, Wei XP. MeSIN: multilevel selective and interactive network for medication recommendation. Knowl Based Syst. Dec 2021;233:107534. [doi: [10.1016/j.knosys.2021.107534](https://doi.org/10.1016/j.knosys.2021.107534)]
20. Le H, Tran T, Venkatesh S. Dual memory neural computer for asynchronous two-view sequential learning. Presented at: KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 19-23, 2018; London, United Kingdom. [doi: [10.1145/3219819.3219981](https://doi.org/10.1145/3219819.3219981)]
21. Yang CQ, Xiao C, Glass L, Sun JM. Change matters: medication change prediction with recurrent residual networks. Presented at: Thirtieth International Joint Conference on Artificial Intelligence; Aug 19-27, 2021; Montreal, Canada. [doi: [10.24963/ijcai.2021/513](https://doi.org/10.24963/ijcai.2021/513)]
22. Wu R, Qiu ZP, Jiang JC, Qi GL, Wu X. Conditional generation net for medication recommendation. Presented at: WWW '22: The ACM Web Conference 2022; Apr 25-29, 2022; Virtual Event, Lyon, France. [doi: [10.1145/3485447.3511936](https://doi.org/10.1145/3485447.3511936)]
23. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible electronic health record dataset. Sci Data. 2016;3:160035. [doi: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35)]
24. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. Jan 3, 2023;10(1):1. [doi: [10.1038/s41597-022-01899-x](https://doi.org/10.1038/s41597-022-01899-x)] [Medline: [36596836](https://pubmed.ncbi.nlm.nih.gov/36596836/)]
25. Zhao D, Shi YL, Cheng L, Li H, Zhang LG, Guo HM. Time interval uncertainty-aware and text-enhanced based disease prediction. J Biomed Inform. Mar 2023;139:104239. [doi: [10.1016/j.jbi.2022.104239](https://doi.org/10.1016/j.jbi.2022.104239)] [Medline: [36356933](https://pubmed.ncbi.nlm.nih.gov/36356933/)]
26. Nguyen MV, Nguyen DT, Trinh QH, Le B. ALGNet: attention light graph memory network for medical recommendation system. Presented at: SOICT 2023: The 12th International Symposium on Information and Communication Technology; Dec 7-8, 2023; Ho Chi Minh, Vietnam. [doi: [10.1145/3628797.3628983](https://doi.org/10.1145/3628797.3628983)]
27. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv. Preprint posted online on Dec 22, 2014. [doi: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980)]
28. Read J, Pfahringer B, Holmes G, Frank E. Classifier chains for multi-label classification. Mach Learn. 2011;85:333-359. [doi: [10.1007/978-3-642-04174-7_17](https://doi.org/10.1007/978-3-642-04174-7_17)]
29. Zhang YT, Chen R, Tang J, Stewart W, Sun JM. LEAP: learning to prescribe effective and safe treatment combinations for multimorbidity. Presented at: KDD '17: The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2017; Halifax, Canada. [doi: [10.1145/3097983.3098109](https://doi.org/10.1145/3097983.3098109)]
30. Zk0814/csrec. GitHub. URL: <https://github.com/zk0814/CSRec> [Accessed 2025-09-18]

Abbreviations

ATC: Anatomical Therapeutic Chemical
CompNet: competitive neural network
DDI: drug-drug interaction
DMNC: dual memory neural computer
ECC: ensemble of classifier chains
EHR : electronic health record
G-BERT: graph-augmented bidirectional encoder representations from transformers
GAMENet: Graph-Augmented Memory Networks
GAT: graph attention network
GCN: graph convolutional network
GCT: graph convolutional transformer
GRU: gated recurrent unit
GT-BEHT: graph transformers of bidirectional encoder representations from transformers on EHRs
HeteroMed: Heterogeneous Information Network for Medical Diagnosis
ICD-9: *International Classification of Diseases, Ninth Revision*
LEAP: Learn to Prescribe
LR: logistic regression
LSTM: long short-term memory

MiME: Multilevel Medical Embedding
MIMIC: Medical Information Mart for Intensive Care
PRAUC: area under the curve of precision-recall
RETAIN: Reverse Time Attention model
RNN: recurrent neural network

Edited by Ling Luo, Qiao Jin; peer-reviewed by Hongbin Lu, Qingqing Zhu, Shunpan Liang; submitted 19.Mar.2025; final revised version received 27.Aug.2025; accepted 28.Aug.2025; published 25.Nov.2025

Please cite as:

Zhang K, Zhang Z, Liang Y, Wang W, Wang X

Hypertension Medication Recommendation via Synergistic and Selective Modeling of Heterogeneous Medical Entities: Development and Evaluation Study of a New Model

JMIR Med Inform 2025;13:e74170

URL: <https://medinform.jmir.org/2025/1/e74170>

doi: [10.2196/74170](https://doi.org/10.2196/74170)

© Ke Zhang, Zhichang Zhang, Yali Liang, Wei Wang, Xia Wang. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 25.Nov.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.