HypKG: Hypergraph-based Knowledge Graph Contextualization for Precision Healthcare

Yuzhang Xie, Xu Han, Ran Xu, Xiao Hu, Jiaying Lu, and Carl Yang

Emory University, Atlanta, GA, USA
{yuzhang.xie,andy.han,ran.xu,xiao.hu,jiaying.lu,j.carlyang}@emory.edu

Abstract. Knowledge graphs (KGs) are important products of the semantic web, which are widely used in various application domains. Healthcare is one of such domains where KGs are intensively used, due to the high requirement for knowledge accuracy and interconnected nature of healthcare data. However, KGs storing general factual information often lack the ability to account for important contexts of the knowledge such as the status of specific patients, which are crucial in precision healthcare. Meanwhile, electronic health records (EHRs) provide rich personal data, including various diagnoses and medications, which provide natural contexts for general KGs. In this paper, we propose HypKG, a framework that integrates patient information from EHRs into KGs to generate contextualized knowledge representations for accurate healthcare predictions. Using advanced entity-linking techniques, we connect relevant knowledge from general KGs with patient information from EHRs, and then utilize a hypergraph model to "contextualize" the knowledge with the patient information. Finally, we employ hypergraph transformers guided by downstream prediction tasks to jointly learn proper contextualized representations for both KGs and patients, fully leveraging existing knowledge in KGs and patient contexts in EHRs. In experiments using a large biomedical KG and two real-world EHR datasets, HypKG demonstrates significant improvements in healthcare prediction tasks across multiple evaluation metrics. Additionally, by integrating external contexts, HypKG can learn to adjust the representations of entities and relations in KG, potentially improving the quality and real-world utility of knowledge.

Keywords: Contextualized Knowledge graph · Hypergraph · Knowledge graph representation · Healthcare application of semantic webs.

1 Introduction

Knowledge graphs (KGs) are important products of the semantic web, organizing and linking entities such as people, places, and objects along with their relationships, in graph structures [26]. They provide a powerful method for representing factual knowledge about the real world [69][8], and have found broad applications across various domains. Healthcare is one of the domains where KGs are extensively utilized, driven by the need for highly accurate knowledge and

the inherently interconnected nature of healthcare data [1]. In the healthcare domain, KGs capture specialized knowledge about medical concepts, including drugs, diseases, genes, and symptoms [74]. In recent years, there has been significant growth in the development of large-scale healthcare KGs, such as UMLS [6], Concept5.5 [46], and iBKH [47]. These healthcare KGs provide rich factual knowledge, enhancing clinical research, decision-making, and healthcare delivery [68].

However, KGs storing general facts often fail to account for important contextual information about the knowledge, such as patients' statuses (e.g., demographics, medical conditions, lifestyles), which can lead to inaccuracies in the knowledge they provide. To ensure the accuracy and relevance of knowledge, it is essential to consider contexts from other data sources. In the healthcare domain, a natural consideration is to utilize KGs together with patient-specific contexts, which is crucial for precision healthcare. For instance, a patient's health status can significantly affect the relevance of drug-disease relationships retrieved from a KG. While a general KG may indicate that aspirin can be used for heart disease prevention [18], this recommendation may be inappropriate for patients with conditions such as gastrointestinal issues or bleeding disorders, where aspirin could cause adverse effects [14]. Another example involves the patient's medication context. A KG might provide information about commonly prescribed medications for depression, but these recommendations may not be suitable for patients taking other medications that could interact with antidepressants [56]. Without incorporating data about a patient's existing prescriptions, the KG might suggest treatments that could result in harmful drug interactions.

In the healthcare domain, a natural way to provide context for knowledge in KGs is to use patient information such as data from Electronic Health Records (EHRs). EHRs are widely recognized as valuable assets for comprehensive patient health analysis and informed decision-making [66]. They are usually constructed from a wide range of patients' digital medical information, including tabular data, clinical notes, medical images, genomics data, and other essential patient data types [2][63][48][62][58]. By consolidating this diverse data, EHRs facilitate a deep understanding of patients' health statuses, enabling more precise medical interventions. Moreover, EHR data often include rich contextual details, such as diagnosis, prescriptions, medication histories, admission and discharge times, and even demographic information like age and gender [37].

Motivated by the lack of context modeling in KGs for real-world applications like healthcare, and the rich contextual information found in EHRs, we propose HypKG, a novel framework to integrate patients' context information from EHR with extracted knowledge from KG, as exemplified in Figure 1. On one hand, the background knowledge extracted from KGs can serve as a valuable complement to the detailed patient data found in EHRs. This background knowledge can provide general medical information, such as common drug-disease interactions and treatment guidelines, which EHR alone may not explicitly capture. On the other hand, the rich context provided by EHRs, such as patient diagnoses, medication histories, lab results, and personal attributes, can inform the selection of
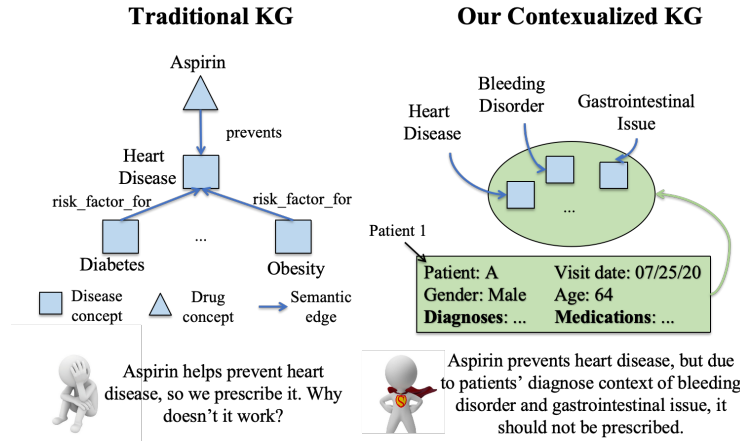
Fig. 1. A toy example of KG contextualization. *Left:* traditional KG. *Right:* our proposed contextualized KG.

the most relevant knowledge from large-scale KGs, tailoring it to the individual patient's specific needs. This integration of EHR and KG enriches the existing knowledge with patient-specific contexts, which could serve as a foundation for a wide range of downstream healthcare applications.

To combine KG knowledge with relevant context information, several challenges must be addressed. The first challenge is **how to connect knowledge and external contexts**. For example, there may be discrepancies in the entity naming conventions used in different sources [27]– a KG may refer to a disease as "Ellis-Van Creveld syndrome", while an EHR may refer to the same disease as "Chondroectodermal dysplasia". To address this challenge of biomedical entity linking, researchers have proposed many methods using crafted string-matching rules [15], constructed thesauri [4], and pre-trained language models [54]. In our approach, we leverage the prior knowledge and semantic understanding embedded in advanced *language models* to facilitate the linking process. Through this process, we robustly enable the KG to be linked with patient-specific context information from the EHR.

Another noteworthy challenge lies in **how to jointly model knowledge and contexts**. While researchers have proposed various models for EHR data, including feature engineering [72], deep neural networks (DNNs) [11], and graph neural networks (GNNs) [40][32], these methods cannot be directly applied to incorporate additional relational knowledge from KGs. In this work, we employ a more flexible and unified graph-based data structure of *hypergraphs* [65][75], where we model EHR attributes and KG entities with nodes and model patient contexts with hyperedges. Unlike standard graphs where edges connect two nodes, hypergraphs allow hyperedges to connect multiple nodes simultaneously, enabling us to directly model the patient-attribute (context-entity) relationships [24][29]. Additionally, the many-to-many connections in hypergraphs also allow

flexible capturing of entity-entity relationships within each patient context, as well as the context-context relationships for each entity.

The third challenge is **how to learn a model that properly integrates knowledge and contexts**. The foundation of this process involves jointly generating entity representations (embeddings) from large-scale KGs and context (patient) representations from EHRs. For KGs alone, many KG embedding methods have been proposed [70], popular ones including TransE [7], ComplEx [50], and CompGCN [51]. These techniques compute compact embedding vectors that capture the entity relationships. We utilize the extensively studied KG embedding methods to capture knowledge in the original KG, and use them to initialize the node features in the hypergraph model. Then we incorporate patient context into the model by utilizing *hypergraph transformers* [10] to directly model the explicit entity-context relationships, and fully learn the entity-entity and context-context relationships from EHR data. We utilize *downstream prediction tasks* to guide the whole training process of hypergraph transformers. This approach enables the KG information to be combined with the rich, patient-specific context from EHRs. The final outcomes are patient-specific representations as well as contextualized KG representations that integrate the strengths of both data sources.

Overall, the core novelty of HypKG lies in its unified, hypergraph-based framework that integrates patient-specific contexts from EHRs directly into a general biomedical KG, rather than treating KG embeddings and EHR features separately or constructing isolated "personalized" KGs. Unlike prior related work that either (a) embeds KGs without patient-specific contextualization, (b) uses only one-hot or simple pooled KG vectors for patients, or (c) builds small per-user graphs, HypKG (1) leverages an LLM-augmented entity-linking pipeline to align EHR medical attributes with KG entities; (2) represents both KG concepts and patient visits as nodes and hyperedges in a single hypergraph, capturing high-order relations; and (3) employs hypergraph transformers supervised by downstream healthcare tasks to jointly learn contextualized embeddings.

In our experiments, we utilize the large-scale public KG, iBKH [47], as the foundational KG dataset. We enhance this KG by contextualizing it with patient-specific context information from two widely used EHR datasets: MIMIC-III [23] and PROMOTE [61]. The experimental results reveal that HypKG significantly improves KG representation, with an average relative performance gain of 12.15% on MIMIC-III and 9.66% on PROMOTE across multiple evaluation metrics. HypKG also outperforms the traditional baselines that directly model the high-dimensional sparse EHR attributes without using KGs. These results highlight the effectiveness of HypKG in integrating patient-specific context into KG representations, enabling more accurate and tailored predictions in healthcare applications. Additionally, case studies show that by integrating external contexts, HypKG can learn to adjust the representations of entities and relations in KG, potentially improving the quality and real-world utility of the KG.

2   Related Work

2.1   Biomedical entity linking.

Linking entities is a crucial part of building comprehensive healthcare knowledge graphs and is essential for entity predictions. Researchers have proposed conventional methods focusing on setting string-matching rules [15] and leveraging constructed thesauri [4], and machine learning-based methods [54][73][21][64] transforming biomedical concepts from raw text into embeddings which are then used to compute similarity scores via distance functions (*e.g.* cosine similarity), and rerank the candidate entities. Recently, some works have explored large language models (LLMs) for biomedical entity linking [60] due to their unprecedentedly rich prior knowledge as well as language capabilities [33][5][45][59]. Building on these advances, we incorporate entity linking techniques into a unified hypergraph framework for downstream predictive tasks.

2.2   KG Representations.

There are many methods aiming to embed entities and relations into vector spaces in a specific dimension. TransE tries to generate embeddings that minimize the distance between the head and the tail entities in a relation, which is one of the easiest methods to train [7]. The ComplEx embedding uses the Hermitian dot product to produce the embeddings and seperates the embeddings into imaginary and real number parts to handle a large variety of binary relations [50]. Composition-based Multi-Relational Graph Convolutional Network (CompGCN) extends the traditional Graph Convolutionary Network to include the relations between entities during the message-passing process. The embeddings are generated by aggregating the neighboring entities as well as the relations between entities [51]. These models effectively learn static embeddings over KGs, but none incorporate external patient contexts during embedding learning; our method advances KG representation by embedding both KG structure and patient-specific information within the same hypergraph.

2.3   Personalized KG for Healthcare.

Our contextualized KGs are distinct from personalized KGs in several ways. Unlike personalized KGs that focus on creating various small KGs to support the information needs of different users [22][67][71][49][9][17], through knowledge contextualization, we add external contexts into KG while still modeling it as a whole. This approach can achieve the following unique effects distinct from personalized KGs: (1) it enables global modeling of interactions between knowledge and contexts, allowing learning to happen across users; (2) it utilizes the contexts to adjust knowledge representations in the whole KG, improving the quality and utility of knowledge that can potentially benefit other related KG applications; (3) it leverages knowledge in the whole KG to improve the modeling of contexts and subsequently improve the modeling of specific users. Due to the often

different application scenarios of personalized KGs compared to our approach, as well as the often complicated and ad hoc processes of creating personalized KGs, we do not empirically compare with personalized KGs in this work and leave this direction as future work. Whereas personalized KGs isolate subgraphs per user, our unified hypergraph contextualization preserves KG coherence and enables collective learning across patients, demonstrating clear methodological and practical advantages.

3   Method

3.1   Preliminaries

Notations. All mathematical notations used in this paper are summarized in a comprehensive table in Section A of the Supplemental Materials.

Knowledge Graph. A knowledge graph is a multi-relational graph $\mathcal{KG} = (X, R, RT)$, where X denotes the set of entities (nodes), R denotes the set of relations (edges), and $RT \in X \times R \times X$ denotes the relational triples in the KG. In other words, knowledge in $\mathcal{KG}$ is stored as a collection of relational triples RT. For example, the relation triple *(Melatonin, Cause, Dry skin)* indicates the fact that melatonin can cause dry skin.

Knowledge Graph Representation. The objective of KG representation learning is to create the mapping functions $\mathcal{F}_X : X \mapsto \mathcal{Z}_X \in \mathbb{R}^{d_{KG}}$ and $\mathcal{F}_R : R \mapsto \mathcal{Z}_R \in \mathbb{R}^{d_{KG}}$ that transform the entities and relations in $\mathcal{KG}$ into a low-dimensional vector space. These representations, $\mathcal{Z}_X$ and $\mathcal{Z}_R$, are designed to capture and preserve the structural and attribute details of the original graph $\mathcal{KG}$, ensuring that the knowledge within the graph is retained.

Contextualizing KG Representation with EHR for Precision Healthcare. We study the KG contextualization problem in the context of patient EHR data. Let p represent a patient, and the associated EHR data can be represented as a sequence of time-stamped medical records: $\rho_p = \{r_1, r_2, \ldots, r_\tau\}$, where r_i represents the i-th record associated with the patient, and τ is the total number of events for the patient. Each medical record r_i can be described as a tuple: $r_i = (t_i, a_i, lt_i)$, where $t_i \in \mathbb{R}^+$ is the timestamp of the record (*i.e.* when the record is charted), a_i is the medical attribute (*e.g.* diagnosis such as heart failure, prescription such as desmopressin), and lt_i is the optional literal value of a_i (*e.g.* N/A for heart failure, 0.05mg for desmopressin). The medical attribute a_i of r_i can be linked to one entity X_k in $\mathcal{KG}$ using entity linking methods. Using the knowledge graph representation learning method $\mathcal{F}_X$, we can obtain the KG representations $(Z_{X_1}, \cdots, Z_{X_\tau})$ of a patient's related KG entities $(X_1, \cdots, X_\tau)$. Consequently, a patient's overall representation Z_p is defined as $Z_p = \mathcal{F}_\mathcal{Z}(Z_{X_1}, \cdots, Z_{X_\tau})$, where $\mathcal{F}_\mathcal{Z}$ is the contextualization function that contextualizes the general factual KG representation into patient-specific knowledge representation.

Hypergraph. A hypergraph $\mathcal{HG} = \{V, E\}$ consists of a finite set of vertices (nodes) V and hyperedges E, where each hyperedge $e \in E$ is a non-empty subset of the vertex set V. Unlike regular graphs where edges connect two vertices, hyperedges can connect any number of vertices (*e.g.* $e_1 = \{v_2, v_4, v_7\}$). The size of a hyperedge e is the number of vertices it connects, while the degree deg_v of a vertex v is the number of hyperedges it participates in. In our HypKG framework, we use hypergraphs to model patient contexts and knowledge, with hyperedges representing patient visits and vertices representing medical attributes.

3.2 HypKG Framework

HypKG is a framework designed to contextualize KG representation with patient-specific context for precision healthcare, and the pipeline is described in Figure 2. First, HypKG connects KG entities with relevant context information from EHR by linking medical entities between them (described in Section 3.3). Then, HypKG jointly represents KG knowledge and contextual information from EHR in a hypergraph structure, capturing key relationships between patients, features, and other patients (described in Section 3.4). Finally, the node and hyperedge embeddings in the hypergraph structure are learned and optimized for downstream precision healthcare tasks (described in Section 3.5). This framework requires the availability of contextual data such as EHRs, reliable mappings between patients and their clinical features, and a robust embedding or entity linking model to initialize node representations effectively. These components ensure that context is accurately integrated into the KG and can be leveraged for improved predictive performance.
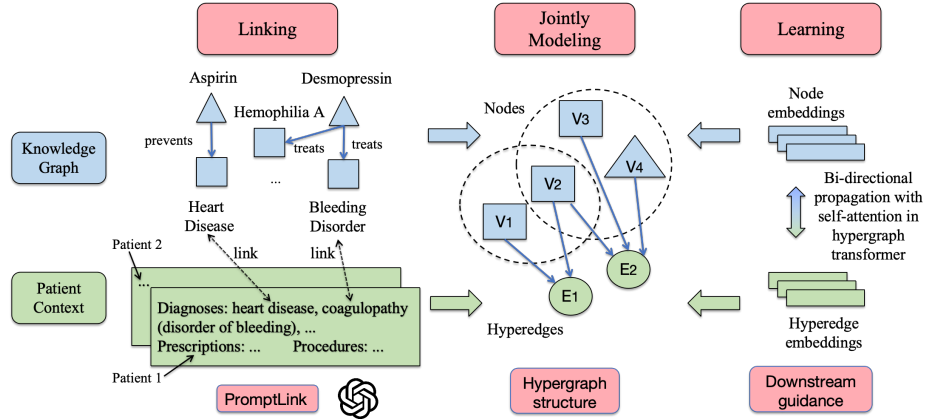

Fig. 2. Overview of HypKG framework. *Left:* Linking knowledge and contexts. *Middle:* Jointly modeling knowledge (e.g., node V_1) and contexts (e.g., hyperedge E_1, which includes nodes V_1 and V_2; hyperedge E_2, which includes nodes V_2, V_3, and V_4). *Right:* Learning to integrate knowledge and contexts.

3.3   Linking Knowledge and External Contexts

To connect KG-extracted knowledge with relevant patient context, we leverage the advanced LLM-based entity-linking technique - PromptLink [60], to accurately link patient-specific context from EHR with KG entities. The process begins with preprocessing the contextual medical attribute a in EHR ρ_p (*e.g.* diagnosis names, medication names, procedure names) by lowercasing and removing punctuation, ensuring uniformity across both EHR and KG entities. We then leverage a pre-trained language model (SAPBERT [31]) to generate embeddings $Z_{lm,a} \in \mathbb{R}^{d_{lm}}$ for EHR entity a and $Z_{lm,X_k} \in \mathbb{R}^{d_{lm}}$ for KG entity X_k. Next, we compute the cosine similarity $SIM \in [0,1]$ between the embeddings $Z_{lm,a}$ and Z_{lm,X_k}, identifying the top-LC candidates (where $LC = 10$) with the highest similarity scores as potential matches for further analysis. At the final stage, we employ a LLM (GPT-4), to establish the entity linking by analyzing the semantic meanings of the candidates. Using both the embeddings and the biomedical knowledge stored in the LLM, we generate final predictions for linking LK (between KG entities X_k and EHR medical attribute a). This multi-step approach, combining deep embeddings and LLM-based reasoning, ensures accurate entity linking, thereby enabling the integration of patient-specific data into the KG for contextualized healthcare applications. The process can be formalized as:

$$LK(a, X_k) = \text{PromptLink}(a, X_k). \tag{1}$$

3.4   Jointly Modeling Knowledge and Contexts

After connecting KG knowledge and contextual information from EHR, HypKG jointly represents them in a hypergraph structure. In this hypergraph model, medical attributes such as diagnosis and medications are treated as nodes V. A patient's individual visit ρ_p (*i.e.* single encounter) is represented as a hyperedge e, which connects the relevant nodes corresponding to the medical attributes from that visit. The collection of these hyperedges forms the hyperedge set E within the hypergraph. Unlike standard graphs that link only two nodes, hypergraphs allow hyperedges to connect multiple nodes at once, making it easier to model the relationship within the patient context. This hypergraph-based approach effectively represents the rich contextual information in EHR data and captures high-order interactions between patients and medical features, providing a more structured and comprehensive representation.

3.5   Learning to Integrate Knowledge and Contexts

HypKG learns to integrate knowledge and contexts. To facilitate this process, we initialize node embeddings within the hypergraph to incorporate KG information and patient context. For each node v, associated with an EHR attribute a and connected to a KG entity X_k, the initial node embedding $\mathcal{Z}_v^0 \in \mathbb{R}^{d_{KG}}$ is derived from the corresponding KG representation.

We design the learning process inspired by hypergraph transformers [30][10][65]. The structures consist of L layers for message passing. In each layer, the embeddings are updated as follows: hyperedge embeddings $\mathcal{Z}_e^l$ of layer l are obtained by aggregating node information from layer $l-1$ within each hyperedge e using aggregation function $f_{V \to E}$, and then the node embeddings $\mathcal{Z}_v^l$ are updated by aggregating information from hyperedges connected to the node v using aggregation function $f_{E \to V}$. These message-passing processes are iteratively applied across layers:

$$\mathcal{Z}_e^l = f_{V \to E}(\mathcal{Z}_v^{l-1}), \quad \mathcal{Z}_v^l = f_{E \to V}(\mathcal{Z}_e^l) \tag{2}$$

Additionally, in the first layer, the hyperedge embedding is aggregated using the initial node embedding $\mathcal{Z}_v^0$.

To implement the message-passing functions $f_{V \to E}$ and $f_{E \to V}$, we apply multi-head attention. Here we use notation S to represent a set, which can be a hyperedge or a node. Specifically, if S represents a hyperedge, the set S consists of its connected nodes, and then the embedding Z_S is updated using the embeddings of its connected nodes from the previous layer. Conversely, if S represents a node, the set S consists of its connected hyperedges, and its embedding Z_S is updated using the embeddings of its connected hyperedges from the previous layer. Given an embedding matrix $\mathcal{Z}_S \in \mathbb{R}^{|S| \times d_{hi}}$, the calculation of output $\mathcal{Z}'_S$ is defined as:

$$\mathcal{Z}'_S = \|_{i=1}^H \text{Attention}_i(Z_S), \tag{3}$$

$$\text{Attention}_i(Z_S) = \text{softmax}\left(\frac{W_{Q,i}^l (Z_S . W_{K,i}^l)^\top}{\sqrt{d_{hi}/H}} \right) Z_S . W_{V,i}^l, \tag{4}$$

where $W_{Q,i}^l \in \mathbb{R}^{1 \times \lfloor d_{hi}/H \rfloor}$, $W_{K,i}^l \in \mathbb{R}^{d_{hi} \times \lfloor d_{hi}/H \rfloor}$, $W_{V,i}^l \in \mathbb{R}^{d_{hi} \times \lfloor d_{hi}/H \rfloor}$ are learnable weight matrices in layer l, H is the number of attention heads, and d_{hi} is the hidden layer dimension.

It is crucial to capture three key relationships within the patient context: (1) Patient-Attribute Relations – linking patients to their respective attributes; (2) Attribute-Attribute Relations – illustrating how attributes influence each other within a patient; and (3) Patient-Patient Relations – identifying similarities between patients with comparable medical histories or conditions. Our HypKG, which utilizes a hypergraph structure with patient-level hyperedges and attribute-level nodes, facilitates the capture of Patient-Attribute Relations through node-to-hyperedge and hyperedge-to-node propagation, as described in Eq. 2. Additionally, the multi-head attention mechanism, defined in Eqs. 3 and 4, enables the model to measure Attribute-Attribute Relations via attention between nodes within a hyperedge, and Patient-Patient Relations through attention between hyperedges connected to a common node.

To guide learning, we define downstream tasks ($\mathcal{TA}$) and divide the hypergraph data into training, validation, and test sets for robust evaluation. During training, KG information is contextualized using patient data, capturing intricate relationships like patient-attribute, attribute-attribute, and patient-patient

relations. The resulting final representation, $\mathcal{Z}_{final}$, incorporates both patient-specific context and KG-derived knowledge, and can then be utilized in various healthcare applications. Since we leverage supervised classification tasks for guidance, the final representation $\mathcal{Z}_{final}$ is passed to a two-layer fully-connected layer (with size $d_{final}=48$) and the loss function is the binary cross-entropy.

4   Experiments

4.1   Experimental Settings

KG Dataset. We use a large-scale public knowledge graph, iBKH [47], as the primary KG dataset. iBKH integrates data from various biomedical knowledge bases, offering a comprehensive resource with 2,384,501 entities across 11 categories, including drugs, diseases, symptoms, genes, and pathways. Moreover, with over 48 million relation triples, iBKH facilitates deeper insights into complex biological interactions. More details about how we utilize the KG dataset to generate embeddings are described in Section D of Supplemental Materials.

EHR Datasets. We contextualize the knowledge graph by integrating patient-specific data from two EHR datasets: MIMIC-III [23] and PROMOTE [61]. Overview of the two EHR datasets are described in Table 1. MIMIC-III, a public dataset, includes over 53,000 patient records from critical care units at Beth Israel Deaconess Medical Center between 2001 and 2012. From MIMIC-III, we follow the settings of Xu et al. [65], to keep 12,353 patient visits (*i.e.* encounters) with 7,423 key medical attributes, including 846 diagnoses, 4,525 prescriptions, and 2,032 procedures. For the MIMIC-III dataset, we perform phenotyping prediction, framed as a multi-label classification task. This involves predicting the presence of 25 acute health conditions (e.g., chronic kidney disease) during a patient's future visits based on their current ICU stay records. PROMOTE is a private dataset containing records of 7,780 stroke patients treated from 2012 to 2021. For this dataset, we extract 2,595 medical attributes, including 1,480 ICD-10 diagnosis codes and 1,115 prescribed medications recorded up to the patients' discharge after their index stroke. For the PROMOTE dataset, we use the occurrence of post-stroke cognitive impairment (PSCI) as the label and define it as a binary classification task. More details about the two EHR datasets are described in Section C of Supplemental Materials.

Table 1. Overview of the EHR datasets. Dx denotes diagnosis, Rx denotes prescriptions, and Px denotes procedures.

EHR Dataset	Patient Encounter Count	List of Features	Task Count
MIMIC-III	12,353	Dx, Rx, Px	25
PROMOTE	7,780	Dx, Rx	1

Evaluation Metrics. In our experiments, we evaluate model performance using key metrics for classification tasks: Accuracy, Area Under the Receiver Operating Characteristic Curve (AUROC), Area Under the Precision-Recall Curve (AUCPR), and the Macro-F1 score. For the binary classification performance on PROMOTE, we follow the default way to compute these metrics. For the multilabel classification performance on MIMIC-III, all metrics are macro-averaged to ensure that performance is evaluated equally across all labels. The model's final performance is determined based on the highest AUROC score achieved on the validation set, as AUROC is a comprehensive metric for evaluating the classification model's effectiveness, and the corresponding test set performance from that epoch is then reported. To ensure robustness, we repeat the experiments five times and report the average performance across these runs. These metrics provide a thorough assessment of the model's predictive capabilities.

Baseline Methods. In our experiments, we compare the downstream risk prediction performance of the following embedding approaches, all of which are patient representations for downstream prediction tasks:

– Binary Embedding: A traditional patient risk factor encoding approach [52][36][57] that represents patients using one-hot binary embeddings for each medical attribute. This results in a high-dimensional binary embedding with over one thousand dimensions (7423 for MIMIC-III and 2595 for PROMOTE), where each dimension indicates the presence or absence of a specific medical attribute during a patient's visit.

– KG Embedding: This method leverages KG information by mean-pooling KG representations based on the linked patient's medical attributes during a visit. While we use PromptLink [60] to perform the entity linking between EHR attributes and KG entities, our approach extends beyond PromptLink by generating and aggregating the KG embeddings into patient-level representations. The resulting 128-dimensional vector captures the patient's relevant KG context in a compact form, offering a lower-dimensional alternative to traditional high-dimensional binary feature embeddings.

– HypKG Embedding: This 128-D embedding is generated using our proposed method HypKG, which integrates patient-specific contextual information with the KG information.

Each of these embeddings is evaluated using several machine learning methods: Logistic Regression (LR), Support Vector Machine (SVM), Multi-layer Perceptron (MLP), Random Forest (RF), and XGBoost (XGB). Additionally, for the HypKG embedding, we also assess the performance during the training process using the patient context integration structure - hypergraph transformer (Hypergraph), since the contextualization process itself also uses the same downstream prediction task. These models allow us to thoroughly assess the effectiveness of the different embeddings in capturing the relevant features for healthcare tasks. The experiments for downstream prediction (Section 4.2) are designed to demonstrate that contextualizing KG embeddings with patient data yields substantially better risk-prediction performance. Moreover, our ablation studies (Section 4.3)

show that HypKG remains robust under variations in hypergraph construction, entity-linking strategies, and other design choices.

Implementation Details. Implementation details of the experiments are described in Section D of Supplemental Materials. These details include the KG embedding generation method, hypergraph modeling hyperparameters, and baseline ML models' hyperparameter configurations.

4.2   Performance on Downstream Prediction

Table 2. Performance comparison of different embeddings on the downstream risk prediction task. For each evaluation metric, the best performance is highlighted in bold, and the second-best result is underlined.

Embedding	Method	MIMIC-III				PROMOTE			
		Accuracy(%)	AUROC(%)	AUCPR(%)	Macro-F1(%)	Accuracy(%)	AUROC(%)	AUCPR(%)	Macro-F1(%)
Binary Embedding	LR	78.58±0.18	81.23±0.14	68.22±0.32	45.79±0.44	74.08±0.69	60.56±1.96	30.67±1.91	53.84±1.37
	SVM	80.41±0.10	83.21±0.23	71.29±0.37	41.77±0.04	77.42±0.69	63.13±1.20	33.48±1.61	44.28±0.26
	MLP	76.05±0.11	78.09±0.22	63.74±0.25	44.21±0.22	69.30±0.58	57.04±1.13	28.22±1.92	53.46±0.71
	RF	77.17±0.25	81.22±0.11	67.45±0.40	23.60±0.21	77.40±0.11	64.73±1.33	32.77±1.55	43.80±1.12
	XGB	79.62±0.20	82.72±0.15	70.83±0.21	49.06±0.34	77.11±0.91	65.01±1.70	33.30±2.42	49.84±1.17
KG Embedding	LR	77.72±0.22	80.66±0.09	66.20±0.36	35.80±0.34	77.40±0.82	61.62±1.70	29.79±2.41	44.08±0.34
	SVM	78.38±0.22	80.84±0.19	67.49±0.39	35.40±0.30	77.49±0.83	54.97±1.39	24.49±0.94	43.78±0.26
	MLP	72.08±0.21	73.12±0.21	53.79±0.54	42.47±0.39	68.77±1.11	57.64±1.55	24.05±6.48	53.39±1.81
	RF	76.57±0.17	78.98±0.16	64.36±0.56	25.13±0.95	77.59±0.83	56.49±1.69	25.79±1.53	43.75±0.26
	XGB	75.35±0.13	77.27±0.19	61.77±0.51	38.65±0.49	77.48±0.87	59.04±1.31	28.95±1.63	46.21±0.84
HypKG Embedding	LR	80.32±0.29	83.98±0.45	72.84±0.91	46.28±0.54	77.62±0.74	66.97±0.51	35.65±1.56	47.58±0.95
	SVM	80.51±0.16	83.07±0.35	71.59±0.62	42.83±0.24	**77.70±0.90**	58.07±1.22	28.49±1.93	44.06±0.54
	MLP	74.61±0.18	77.10±0.30	62.15±4.47	46.91±0.62	69.17±1.32	58.72±1.59	32.19±1.61	**54.60±1.52**
	RF	80.54±0.23	84.01±0.22	72.80±0.47	41.72±0.29	74.26±7.92	63.74±1.15	32.19±1.92	46.98±4.74
	XGB	78.86±0.21	81.89±0.61	69.25±0.63	**47.12±0.63**	77.02±0.52	64.22±1.28	32.65±2.85	49.43±0.82
	Hypergraph	**80.86±0.17**	**84.26±0.17**	**73.12±0.37**	45.30±0.49	77.06±0.62	**67.37±1.02**	**35.74±1.67**	53.28±1.54

Table 2 summarizes the downstream risk prediction results on the two datasets to show the impact of contextualization. In this table, our proposed contextualized HypKG Embedding consistently outperforms other baseline methods. Across various evaluation metrics on both datasets, the HypKG Embedding consistently achieves the best and second-best performance. This performance gain demonstrates the effectiveness of integrating KG information with patient context to generate richer, more representative embeddings. Notably, the hypergraph model within the HypKG framework achieves the strongest results across most metrics. This finding suggests that the HypKG model alone is powerful enough to provide robust representations and the hypergraph model has a strong modeling ability, eliminating the need to extract embeddings and apply them to external machine learning models.

The advantages of contextualization within HypKG are further emphasized when compared with the KG Embedding. Across all evaluation metrics and on both datasets, the HypKG Embedding consistently achieves better results, demonstrating the effectiveness of incorporating contextual information and KG information. Moreover, we observe an average relative improvement of 12.15% on MIMIC-III and 9.66% on PROMOTE by comparing all methods utilizing

HypKG Embedding and those using KG Embedding. This improvement highlights the significant impact of contextualization, allowing for a deeper understanding of patient information and stronger performance in downstream tasks.

In comparison to HypKG, the other two methods, Binary Embedding and KG Embedding, show comparable performance in certain models, such as Binary Embedding with SVM. However, both embeddings suffer from significant limitations. Binary Embedding, a traditional approach for generating patient representations in healthcare tasks, faces the curse of dimensionality, especially in medical datasets where the number of attributes can exceed 1,000. This leads to an exponential increase in computational complexity. For instance, when using SVM with Binary Embedding on the MIMIC-III dataset, the computation time exceeded 10 hours—a time cost that is over 20 times longer than the time required by HypKG for the same task with SVM. This inefficiency highlights the unsustainable nature of Binary Embedding for large-scale applications. On the other hand, KG Embedding reduces the representation size and also incorporates patient information with KG information. However, the method used for combining them (simple mean-pooling)lacks the sophistication required to capture the intricate relationships between the knowledge graph and patient contexts. As a result, the performance of KG Embedding lags behind that of both Binary Embedding and HypKG Embedding. The main limitation here is that the KG information is not sufficiently contextualized, leading to a less effective representation. This is where HypKG excels, as it effectively integrates KG and contextual information, modeling intricate relations between them and producing better global representations efficiently for downstream applications.

4.3   Ablation Studies

HypKG consists of three key components: linking, joint modeling, and learning, leveraging EHR data to contextualize KG representations. To validate its design, we conduct comprehensive ablation studies. These studies evaluate the impact of entity linking (Supplemental Materials Section E), KG embedding generation (Supplemental Materials Section F), joint modeling of KG and EHR (Supplemental Materials Section G), hypergraph design (Supplemental Materials Section H), and learning hyperparameters (Supplemental Materials Section I). These ablation studies confirm the effectiveness of HypKG's implementation.

Entity Linking Methods. We evaluate the impact of different entity linking methods on HypKG, comparing PromptLink with BM25, BioBERT, and shuffled embeddings. Results show that PromptLink consistently achieves the best performance. Meanwhile, HypKG remains robust with only minor performance drops when using simpler or noisier linking methods. This highlights both the effectiveness of PromptLink and the resilience of HypKG to linking errors in real-world settings. Details of the experiments are described in Section E of Supplemental Materials.

KG Embedding Generation Methods. We evaluate the impact of different KG embedding methods on HypKG's performance. Details of the experiments are described in Section F of Supplemental Materials. Results show that our default setting (ComplEx + Large) achieves the best overall performance, while alternative methods like TransE and CompGCN perform slightly worse. Overall, HypKG remains robust across different embedding configurations, highlighting its adaptability and efficiency in practical applications.

Jointly Modeling EHR and KG. To assess the importance of jointly modeling EHR and KG information, we compare HypKG against two ablation baselines: KG Only and EHR Only, as demonstrated in Section G of Supplemental Materials. Results across both datasets show that HypKG significantly outperforms the baselines, demonstrating the complementary value of KG embeddings and EHR-derived hypergraph structures. This confirms the effectiveness of our joint modeling design in enhancing contextualization and predictive accuracy.

Hypergraph Design. We evaluate HypKG's hypergraph design by comparing it with state-of-the-art hypergraph models and alternative message-passing strategies. Details and results are described in Section H of Supplemental Materials. In experiments, HypKG consistently outperforms baselines across key metrics, demonstrating its effectiveness in integrating structural and semantic information from both EHRs and KGs. These results validate our contextualization approach and highlight HypKG's superior scalability and domain-specific design for precision healthcare.

Hyperparameter Studies. We conduct a comprehensive hyperparameter study to assess the impact of key settings on HypKG's performance, as described in Section I of Supplemental Materials. Results on both MIMIC-III and PRO-MOTE datasets show that our selected configurations achieve strong and stable outcomes, with learning rate, hypergraph transformer depth, hidden size, and attention heads all influencing performance. Among these, the learning rate plays the most critical role in balancing model optimization.

4.4 Case Studies on Knowledge Representation

To demonstrate that the representations generated by HypKG are effectively contextualized, we compare the knowledge embeddings before and after training. For each KG medical entity pair (Entity A, Entity B), we analyze them from three key perspectives: (1) Similarity: We compute the cosine similarity between the knowledge representations of each entity pair both before and after training. By comparing these values, we observe an increase in similarity after the application of HypKG for most pairs (more details analyzed in Section K of Supplemental Materials), indicating that the contextual relationships between entities are better captured. From this analysis, we identify twenty top entity

pairs (including pairs of diagnosis and prescription) that show a great increase in representation similarity within both the MIMIC-III and PROMOTE datasets for a case study. Since the cosine similarity values are within the range $[-1, 1]$, the increases in similarity for these listed pairs are significant as they approach 1. (2) Occurrence and Prevalence: We further investigate the frequency with which each entity pair co-occurs in the same patient context. Additionally, we determine the prevalence of each entity in these contexts by dividing the number of co-occurrences by the total number of occurrences for each entity individually (i.e., Prevalence A and Prevalence B). This allows us to understand how common it is for both entities to appear together and how strongly their occurrence is correlated within a patient's data. (3) Relation Analysis: To provide further insights into the relationships between the entities, we consult with clinicians for qualitative assessments and cite relevant resources.

Table 3. Case studies on knowledge representation. "Increase" denotes the increase in similarity between entity pairs after applying HypKG. "Occurrence" indicates the number of times the entity pair co-occurs in the same patient context. "Prevalence A" and "Prevalence B" represent the ratio of co-occurrences to the total occurrences for each entity individually. "References" denote existing literature that discusses the relationship between the entity pairs.

Dataset and Pair	ID	Entity A	Entity B	Increase	Occurrence	Prevalence A	Prevalence B	References
MIMIC-III	I	Benign neonatal seizures	Pneumonia	1.08	25	67.57%	40.98%	[43]
Diagnosis	II	Benign neonatal seizures	Congenital hemolytic anemia	1.01	30	81.08%	34.88%	[16]
	III	Upper extremity deep vein thrombosis	Limb injury	1.01	4	57.14%	100.00%	[44]
	IV	Congenital hemolytic anemia	Pneumonia	1.00	38	44.19%	62.30%	[28]
	V	Closed fracture of bones, unspecified	Pneumothorax	0.99	43	36.44%	78.18%	[3]
MIMIC-III	VI	Starter PN D5	Heparin sodium	1.21	7	53.85%	58.33%	[39]
Prescription	VII	Insulin	Acetaminophen	1.13	5566	75.34%	69.38%	[35]
	VIII	Drug withdrawal syndrome neonatal	Indomethacin sodium	1.12	9	36.00%	81.82%	[34]
	IX	Syringe (Neonatal)	Indomethacin sodium	1.11	11	44.00%	84.62%	[53]
	X	Urethral stricture	Pulmonary tuberculosis	1.11	19	79.17%	63.33%	[38]
PROMOTE	XI	Respiratory system disease	Medical device complication	1.09	1043	66.01%	68.21%	[42]
Diagnosis	XII	Human immunodeficiency virus infectious disease	HIV infection	1.07	95	75.40%	84.07%	[19]
	XIII	Symptoms of nervous & musculoskeletal systems	Long term (current) drug therapy	1.06	2990	72.89%	55.12%	[12]
	XIV	Essential hypertension	Chronic fatigue syndrome	1.02	2966	44.37%	89.61%	[20]
	XV	Essential hypertension	Cerebrovascular disease	0.99	4282	64.06%	87.57%	[55]
PROMOTE	XVI	Isopropyl alcohol	Lancets	0.92	95	98.96%	50.00%	[25]
Prescription	XVII	Blood glucose	Lancets	0.87	181	88.73%	95.26%	[25]
	XVIII	Emtricitabine & Tenofovir disoproxil fumarate	Ritonavir	0.87	12	75.00%	66.67%	[41]
	XIX	Mycophenolate mofetil	Tacrolimus	0.83	59	71.08%	74.68%	[13]
	XX	Blood glucose	Isopropyl alcohol	0.76	86	54.78%	89.58%	[25]

The case study results are described in Table 3. Overall, HypKG enhances the similarity between entity pairs that are similar in patient contexts, even when their relationships are not explicitly captured by the KG, showing the power of our contextualization. For instance, in Case I, "Benign neonatal seizures" and "Pneumonia" are two distinct diagnostic conditions that, according to the KG, have no direct causal link. However, in the EHR context, these dignoses are associated with the medical attributes "Convulsions in newborn" and "Pneumonia", which often co-occur in newborns. Pneumonia, for example, may indirectly contribute to neonatal seizures by causing physiological stress that makes seizures more likely in affected infants [43]. While the KG might not capture this indirect relationship, the contextual co-occurrence in patient data suggests a meaningful association. Another example is seen in Case VII, which involves the prescription entity pair of Insulin and Acetaminophen. These two drugs do not have a direct pharmacological interaction, as insulin is primarily used to manage blood sugar

levels in diabetic patients, while acetaminophen is typically prescribed for pain or fever relief. However, they are linked to their corresponding medical attributes within the EHR context and frequently co-occur. This can be explained by the phenomenon that patients with diabetes often require both medications—insulin for managing their condition and acetaminophen to alleviate common symptoms like pain or fever [35]. Although the KG does not reflect a direct connection between these drugs, their frequent co-prescription in clinical practice highlights a contextual relationship, which HypKG is able to capture. In summary, this case study analysis demonstrates that HypKG contextualizes the KG knowledge and generates better representations.

5   Conclusion

In this study, we propose HypKG, a novel framework that addresses the critical need for contextualized knowledge representations for precision healthcare by integrating context-specific information from EHRs with general knowledge from KGs. In this framework, we first use entity-linking techniques to connect general KG knowledge with patient contexts from EHRs, and then apply a hypergraph model to contextualize the knowledge via hyperedges. Finally, we employ hypergraph transformers guided by downstream prediction tasks to jointly learn contextualized representations for both KG and patients, leveraging the combined knowledge. In experiments, HypKG shows significant improvements in healthcare prediction tasks across multiple evaluation metrics. Additionally, by integrating external contexts, HypKG can learn to adjust the representations of entities and relations in KG, potentially improving the quality and real-world utility of knowledge. The application of HypKG allows for the effective contextualization of knowledge and advances precision healthcare, capturing the intricate relationships between patient data and general knowledge in a way that traditional KGs fail to do.

Supplemental Material Statement

The source code and dataset download instructions of this paper can be found at https://github.com/constantjxyz/HypKG. Additional materials—including the notation table, algorithm descriptions, dataset details, implementation specifics, and ablation studies—are provided in the Supplementary Materials of the extended version of the paper on arXiv.

References

1. Abu-Salih, B., Al-Qurishi, M., Alweshah, M., Al-Smadi, M., Alfayez, R., Saadeh, H.: Healthcare knowledge graph construction: A systematic review of the state-of-the-art, open issues, and opportunities. Journal of Big Data **10**(1), 81 (2023)
2. Abul-Husn, N.S., Kenny, E.E.: Personalized medicine and the power of electronic health records. Cell **177**(1), 58–69 (2019)
3. Amer, K.M., Congiusta, D.V., Suri, P., Choudhry, A., Otero, K., Adams, M.: Clavicle fractures: Associated trauma and morbidity. Journal of clinical orthopaedics and trauma **13**, 53–56 (2021)
4. Aronson, A.R., Lang, F.M.: An overview of metamap: historical perspective and recent advances. Journal of the American Medical Informatics Association **17**(3), 229–236 (2010)
5. Bhasuran, B., Jin, Q., Xie, Y., Yang, C., Hanna, K., Costa, J., Shavor, C., Han, W., Lu, Z., He, Z.: Preliminary analysis of the impact of lab results on large language model generated differential diagnoses. npj Digital Medicine **8**(1), 166 (2025)
6. Bodenreider, O.: The unified medical language system (umls): integrating biomedical terminology. Nucleic acids research **32**(suppl_1), D267–D270 (2004)
7. Bordes, A., Weston, J., Collobert, R., Bengio, Y.: Learning structured embeddings of knowledge bases. In: Proceedings of the AAAI conference on artificial intelligence. vol. 25, pp. 301–306 (2011)
8. Cao, J., Fang, J., Meng, Z., Liang, S.: Knowledge graph embedding: A survey from the perspective of representation spaces. ACM Computing Surveys **56**(6), 1–42 (2024)
9. Carvalho, R.M., Oliveira, D., Pesquita, C.: Knowledge graph embeddings for icu readmission prediction. BMC Medical Informatics and Decision Making **23**(1), 12 (2023)
10. Chien, E., Pan, C., Peng, J., Milenkovic, O.: You are allset: A multiset function framework for hypergraph neural networks. arXiv preprint arXiv:2106.13264 (2021)
11. Choi, Y., Chiu, C.Y.I., Sontag, D.: Learning low-dimensional representations of medical concepts. AMIA Summits on Translational Science Proceedings **2016**, 41 (2016)
12. Clauw, D.J.: Diagnosing and treating chronic musculoskeletal pain based on the underlying mechanism (s). Best Practice & Research Clinical Rheumatology **29**(1), 6–19 (2015)
13. Dalal, P., Shah, G., Chhabra, D., Gallon, L.: Role of tacrolimus combination therapy with mycophenolate mofetil in the prevention of organ rejection in kidney transplant patients. International journal of nephrology and renovascular disease pp. 107–115 (2010)
14. De Berardis, G., Lucisano, G., D'Ettorre, A., Pellegrini, F., Lepore, V., Tognoni, G., Nicolucci, A.: Association of aspirin use with major bleeding in patients with and without diabetes. Jama **307**(21), 2286–2294 (2012)
15. D'Souza, J., Ng, V.: Sieve-based entity linking for the biomedical domain. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). pp. 297–302 (2015)
16. Gallagher, P.G.: Anemia in the pediatric patient. Blood, The Journal of the American Society of Hematology **140**(6), 571–593 (2022)
17. Gao, Y., Li, R., Caskey, J., Dligach, D., Miller, T., Churpek, M.M., Afshar, M.: Leveraging a medical knowledge graph into large language models for diagnosis prediction. arXiv preprint arXiv:2308.14321 (2023)

18. Gasparyan, A.Y., Watson, T., Lip, G.Y.: The role of aspirin in cardiovascular prevention: implications of aspirin resistance. Journal of the American College of Cardiology **51**(19), 1829–1843 (2008)
19. German Advisory Committee Blood (Arbeitskreis Blut), S.A.o.P.T.b.B.: Human immunodeficiency virus (hiv). Transfusion medicine and hemotherapy **43**(3), 203–222 (2016)
20. Hulens, M., Rasschaert, R., Vansant, G., Stalmans, I., Bruyninckx, F., Dankaerts, W.: The link between idiopathic intracranial hypertension, fibromyalgia, and chronic fatigue syndrome: exploration of a shared pathophysiology. Journal of pain research pp. 3129–3140 (2018)
21. Ji, Z., Wei, Q., Xu, H.: Bert-based ranking for biomedical entity normalization. AMIA Summits on Translational Science Proceedings **2020**, 269 (2020)
22. Jiang, P., Xiao, C., Cross, A.R., Sun, J.: Graphcare: Enhancing healthcare predictions with personalized knowledge graphs. In: The Twelfth International Conference on Learning Representations (2024)
23. Johnson, A.E., Pollard, T.J., Shen, L., Lehman, L.w.H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Anthony Celi, L., Mark, R.G.: Mimic-iii, a freely accessible critical care database. Scientific data **3**(1), 1–9 (2016)
24. Johnson, R., Li, M.M., Noori, A., Queen, O., Zitnik, M.: Graph artificial intelligence in medicine. Annual Review of Biomedical Data Science **7** (2024)
25. Jońca, M., Krótki, F., Tomasik, P.: The effect of disinfecting procedure on the glucose concentration using a personal glucose meter. Primary Care Diabetes **15**(5), 848–852 (2021)
26. Kejriwal, M.: Knowledge graphs: A practical review of the research landscape. Information **13**(4), 161 (2022)
27. Kohane, I.S., Aronow, B.J., Avillach, P., Beaulieu-Jones, B.K., Bellazzi, R., Bradford, R.L., Brat, G.A., Cannataro, M., Cimino, J.J., García-Barrio, N., et al.: What every reader should know about studies using electronic health record data but may be afraid to ask. Journal of medical Internet research **23**(3), e22219 (2021)
28. Kurugol, Z., Onen, S.S., Koturoglu, G.: Severe hemolytic anemia associated with mild pneumonia caused by mycoplasma pneumonia. Case reports in medicine **2012**(1), 649850 (2012)
29. Lee, G., Bu, F., Eliassi-Rad, T., Shin, K.: A survey on hypergraph mining: Patterns, tools, and generators. arXiv preprint arXiv:2401.08878 (2024)
30. Lee, J., Lee, Y., Kim, J., Kosiorek, A., Choi, S., Teh, Y.W.: Set transformer: A framework for attention-based permutation-invariant neural networks. In: International conference on machine learning. pp. 3744–3753. PMLR (2019)
31. Liu, F., Shareghi, E., Meng, Z., Basaldella, M., Collier, N.: Self-alignment pretraining for biomedical entity representations. arXiv preprint arXiv:2010.11784 (2020)
32. Liu, Z., Wang, X., Wang, B., Huang, Z., Yang, C., Jin, W.: Graph odes and beyond: A comprehensive survey on integrating differential equations with graph neural networks. arXiv preprint arXiv:2503.23167 (2025)
33. Lu, J., Shen, J., Xiong, B., Ma, W., Staab, S., Yang, C.: Hiprompt: Few-shot biomedical knowledge fusion via hierarchy-oriented prompting. In: 46th International ACM SIGIR Conference on Research and Development in Information Retrieval - Short Paper (2023)
34. Lucas, S.: The pharmacology of indomethacin. Headache: The Journal of Head and Face Pain **56**(2), 436–446 (2016)
35. Maahs, D.M., DeSalvo, D., Pyle, L., Ly, T., Messer, L., Clinton, P., Westfall, E., Wadwa, R.P., Buckingham, B.: Effect of acetaminophen on cgm glucose in an outpatient setting. Diabetes Care **38**(10), e158–e159 (2015)

36. Manikandan, R., Kuwelkar, S., Sivakumar, R.: An hybrid technique for optimized clustering of ehr using binary particle swarm and constrained optimization for better performance in prediction of cardiovascular diseases. Measurement: Sensors **25**, 100577 (2023)
37. Murali, L., Gopakumar, G., Viswanathan, D.M., Nedungadi, P.: Towards electronic health record-based medical knowledge graph construction, completion, and applications: A literature study. Journal of biomedical informatics **143**, 104403 (2023)
38. Naeem, M., Zulfiqar, M., Siddiqui, M.A., Shetty, A.S., Haq, A., Varela, C., Siegel, C., Menias, C.O.: Imaging manifestations of genitourinary tuberculosis. Radiographics **41**(4), 1123–1143 (2021)
39. Novick, K., Cober, M.P.: Evaluation of inpatient starter parenteral nutrition use in the neonatal intensive care unit. The Journal of Pediatric Pharmacology and Therapeutics **27**(6), 524–528 (2022)
40. Ochoa, J.G.D., Mustafa, F.E.: Graph neural network modelling as a potentially effective method for predicting and analyzing procedures based on patients' diagnoses. Artificial Intelligence in Medicine **131**, 102359 (2022)
41. Reynes, J., Trinh, R., Pulido, F., Soto-Malave, R., Gathe, J., Qaqish, R., Tian, M., Fredrick, L., Podsadecki, T., Norton, M., et al.: Lopinavir/ritonavir combined with raltegravir or tenofovir/emtricitabine in antiretroviral-naive subjects: 96-week results of the progress study. AIDS research and human retroviruses **29**(2), 256–265 (2013)
42. Samore, M.H., Evans, R.S., Lassen, A., Gould, P., Lloyd, J., Gardner, R.M., Abouzelof, R., Taylor, C., Woodbury, D.A., Willy, M., et al.: Surveillance of medical device–related hazards and adverse events in hospitalized patients. Jama **291**(3), 325–334 (2004)
43. Schüssler, S.C., Paul, A., Niederreiter, U., Deiters, L., Fahlbusch, F.B., Morhart, P., Trollmann, R.: Seizures in preterm infants with germinal-matrix-intraventricular hemorrhage (gm-ivh): a retrospective monocentric study on predictors and neurodevelopmental outcome. European Journal of Paediatric Neurology (2025)
44. Singh, O., Juneja, D.: Upper extremity deep vein thrombosis: an intensivist's perspective. World Journal of Critical Care Medicine **12**(3), 130 (2023)
45. Singhal, K., Azizi, S., Tu, T., Mahdavi, S.S., Wei, J., Chung, H.W., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., et al.: Large language models encode clinical knowledge. Nature **620**(7972), 172–180 (2023)
46. Speer, R., Chin, J., Havasi, C.: Conceptnet 5.5: An open multilingual graph of general knowledge. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
47. Su, C., Hou, Y., Zhou, M., Rajendran, S., Maasch, J.R., Abedi, Z., Zhang, H., Bai, Z., Cuturrufo, A., Guo, W., et al.: Biomedical discovery through the integrative biomedical knowledge hub (ibkh). Iscience **26**(4) (2023)
48. Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S., Wang, G.: Data processing and text mining technologies on electronic medical records: a review. Journal of healthcare engineering **2018** (2018)
49. Theodoropoulos, C., Mulligan, N., Stappenbeck, T., Bettencourt-Silva, J.: Representation learning for person or entity-centric knowledge graphs: An application in healthcare. In: Proceedings of the 12th Knowledge Capture Conference 2023. pp. 225–233 (2023)
50. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., Bouchard, G.: Complex embeddings for simple link prediction. In: International conference on machine learning. pp. 2071–2080. PMLR (2016)

51. Vashishth, S., Sanyal, S., Nitin, V., Talukdar, P.: Composition-based multi-relational graph convolutional networks. In: International Conference on Learning Representations
52. Vasiljeva, I., Arandjelović, O.: Diagnosis prediction from electronic health records using the binary diagnosis history vector representation. Journal of Computational Biology **24**(8), 767–786 (2017)
53. Walker, S.E., Gray, S., Schmidt, B.: Stability of reconstituted indomethacin sodium trihydrate in original vials and polypropylene syringes. American journal of health-system pharmacy **55**(2), 154–158 (1998)
54. Wang, B., Xie, Q., Pei, J., Chen, Z., Tiwari, P., Li, Z., Fu, J.: Pre-trained language models in biomedical domain: A systematic survey. ACM Computing Surveys **56**(3), 1–52 (2023)
55. Webb, A.J., Werring, D.J.: New insights into cerebrovascular pathophysiology and hypertension. Stroke **53**(4), 1054–1064 (2022)
56. Wolff, J., Reißner, P., Hefner, G., Normann, C., Kaier, K., Binder, H., Hiemke, C., Toto, S., Domschke, K., Marschollek, M., et al.: Pharmacotherapy, drug-drug interactions and potentially inappropriate medication in depressive disorders. PLoS One **16**(7), e0255192 (2021)
57. Wu, J., Roy, J., Stewart, W.F.: Prediction modeling using ehr data: challenges, strategies, and a comparison of machine learning approaches. Medical care **48**(6), S106–S113 (2010)
58. Wu, Y., Nahab, F., Ge, Y., Xie, Y., Aboul-Nour, H., Yang, C., Hu, X.: Prediction of post-stroke af in esus patients is enhanced by combining expert-derived predictors and embedding of full diagnostic codes using pre-trained hypergraph neural networks. In: STROKE. vol. 56. LIPPINCOTT WILLIAMS & WILKINS TWO COMMERCE SQ, 2001 MARKET ST, PHILADELPHIA . . . (2025)
59. Xie, Y., Cui, H., Zhang, Z., Lu, J., Shu, K., Nahab, F., Hu, X., Yang, C.: Kerap: A knowledge-enhanced reasoning approach for accurate zero-shot diagnosis prediction using multi-agent llms. arXiv preprint arXiv:2507.02773 (2025)
60. Xie, Y., Lu, J., Ho, J., Nahab, F., Hu, X., Yang, C.: Promptlink: Leveraging large language models for cross-source biomedical concept linking. In: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 2589–2593 (2024)
61. Xie, Y., Nahab, F., Ge, Y., Wu, Y., Saurman, J., Yang, C., Hu, X.: Abstract wp175: Predicting post-stroke cognitive impairment (psci) using multiple machine learning approaches. Stroke **56**(Suppl_1), AWP175–AWP175 (2025)
62. Xie, Y., Niu, G., Da, Q., Dai, W., Yang, Y.: Survival prediction for gastric cancer via multimodal learning of whole slide images and gene expression. In: 2022 IEEE international conference on bioinformatics and biomedicine (BIBM). pp. 1311–1316. IEEE (2022)
63. Xie, Y., Sang, Q., Da, Q., Niu, G., Deng, S., Feng, H., Chen, Y., Li, Y.Y., Liu, B., Yang, Y., et al.: Improving diagnosis and outcome prediction of gastric cancer via multimodal learning using whole slide pathological images and gene expression. Artificial Intelligence in Medicine **152**, 102871 (2024)
64. Xu, D., Zhang, Z., Bethard, S.: A generate-and-rank framework with semantic type regularization for biomedical concept normalization. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 8452–8464 (2020)
65. Xu, R., Ali, M.K., Ho, J.C., Yang, C.: Hypergraph transformers for ehr-based clinical predictions. AMIA Summits on Translational Science Proceedings **2023**, 582 (2023)

66. Xu, R., Yu, Y., Zhang, C., Ali, M.K., Ho, J.C., Yang, C.: Counterfactual and factual reasoning over hypergraphs for interpretable clinical predictions on ehr. In: Machine Learning for Health. pp. 259–278. PMLR (2022)

67. Xu, Y., Chu, X., Yang, K., Wang, Z., Zou, P., Ding, H., Zhao, J., Wang, Y., Xie, B.: Seqcare: Sequential training with external medical knowledge graph for diagnosis prediction in healthcare data. In: Proceedings of the ACM Web Conference 2023. pp. 2819–2830 (2023)

68. Yang, C., Cui, H., Lu, J., Wang, S., Xu, R., Ma, W., Yu, Y., Yu, S., Kan, X., Ling, C., Fu, T., Zhao, L., Ho, J., Wang, F.: A review on knowledge graphs for healthcare: Resources, applications, and promises. arXiv preprint arXiv:2306.04802 (2023)

69. Yang, H., Lin, Z., Zhang, M.: Rethinking knowledge graph evaluation under the open-world assumption. Advances in Neural Information Processing Systems **35**, 8374–8385 (2022)

70. Yang, Y., Huang, C., Xia, L., Li, C.: Knowledge graph contrastive learning for recommendation. In: Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval. pp. 1434–1443 (2022)

71. Ye, M., Cui, S., Wang, Y., Luo, J., Xiao, C., Ma, F.: Medpath: Augmenting health risk prediction via medical knowledge paths. In: Proceedings of the Web Conference 2021. pp. 1397–1409 (2021)

72. Yu, Y., Zuo, S., Jiang, H., Ren, W., Zhao, T., Zhang, C.: Fine-tuning pre-trained language model with weak supervision: A contrastive-regularized self-training approach. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 1063–1077 (2021)

73. Zhang, S., Cheng, H., Vashishth, S., Wong, C., Xiao, J., Liu, X., Naumann, T., Gao, J., Poon, H.: Knowledge-rich self-supervision for biomedical entity linking. In: Findings of the Association for Computational Linguistics: EMNLP 2022. pp. 868–880 (2022)

74. Zhang, Y., Sheng, M., Zhou, R., Wang, Y., Han, G., Zhang, H., Xing, C., Dong, J.: Hkgb: an inclusive, extensible, intelligent, semi-auto-constructed knowledge graph framework for healthcare with clinicians' expertise incorporated. Information Processing & Management **57**(6), 102324 (2020)

75. Zhang, Z., Cui, H., Xu, R., Xie, Y., Ho, J.C., Yang, C.: Tacco: Task-guided co-clustering of clinical concepts and patient visits for disease subtyping based on ehr data. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. pp. 6324–6334 (2024)