生物信息学 DOI：10.13406/j.cnki.cyxb.003555

基于人工智能SGRN-Trans框架预测温胆汤中成分-靶点相互作用的研究

王艳菁[1]，李治琦[1]，魏冬青[1]，徐　威[2]，谭红胜[2]
（1.上海交通大学生命科学技术学院，上海　200240；2.上海交通大学医学院，上海　200025）

【摘　要】目的：以温胆汤为例，构建基于知识图谱和注意力机制的深度学习模型（SGRN-Trans）预测中医经典名方中药效成分与靶点的相互作用，评价其预测效果。**方法：**首次提出SGRN-Trans（Self-weighted Graph Relational Network-Transformer）预测模型，结合多生物数据源构建中医经典名方温胆汤知识图谱（Wendan Decoction Knowledge Graph，WDKG），利用图神经网络学习知识图谱中每个实体的低维嵌入表示，引入中药成分和靶点各自的结构特征，搭载基于注意力机制的Transformer模型进行药效成分-靶点相互作用的预测，结合分子对接及文献调研进行验证。**结果：**WDKG包含10个类型共14 292个实体，可用于深度学习模型的研究。SGRN-Trans预测模型与TransE、TransR、ComplEx、DistMult、ConvKB等其他知识图谱嵌入模型的性能相比，效果最优。将预测排序前20组的药效成分与靶点分别进行分子对接和可视化呈现，其中8组的结合能提示其药效成分与靶点有潜在的相互作用。以温胆汤中半夏的有效成分soya-cerebroside（大豆脑苷脂）与低密度脂蛋白受体（low density lipoprotein receptor，LDLR）相互作用为例，结合研究文献进行讨论，可能是温胆汤治疗动脉粥样硬化的机制之一。**结论：**本研究提出基于知识图谱和注意力机制的模型SGRN-Trans，可推广用于预测中医药经典名方复杂网络体系中成分与靶点的相互作用，为阐明经典名方的药效物质基础和作用机制提供新的工具。

【关键词】温胆汤；药物-靶点相互作用；知识图谱；图神经网络；注意力机制
【中图分类号】R221.2　　**【文献标志码】**A　　**【收稿日期】**2024-05-26

Prediction of component-target interactions in Wendan Decoction based on the artificial intelligence SGRN-Trans framework

Wang Yanjing[1]，*Li Zhiqi*[1]，*Wei Dongqing*[1]，*Xu Wei*[2]，*Tan Hongsheng*[2]
（*1.Shanghai Jiao Tong University School of Life Science and Technology*；
2.Shanghai Jiao Tong University School of Medicine）

【Abstract】*Objective*：To construct a deep learning model（SGRN-Trans）based on knowledge graph and attention mechanism for predicting the interaction between pharmacodynamic components and targets in classic traditional Chinese medicine（TCM）prescriptions with Wendan Decoction as an example，and to assess its predictive performance. ***Methods*：**The SGRN-Trans predictive model was proposed for the first time. Multiple biological data sources were used to construct the knowledge graph of Wendan Decoction（WDKG），and graph neural networks were used to learn the low-dimensional embedding representation of each entity in the knowledge graph. The respective structural features of TCM components and targets were introduced，and the Transformer model based on attention mechanism was used to predict the interaction between pharmacodynamic components and targets. Molecular docking and literature review were used for validation. ***Results*：**WDKG contained 10 types of entities，with 14292 entities in total，which could be used for the research on deep learning models. The SGRN-Trans predictive model showed the best performance compared with other knowledge graph embedding models such as TransE，TransR，ComplEx，DistMult，and ConvKB. Molecular docking and visualized presentation were performed for the top 20 groups of pharmacodynamic components and targets，among which 8 combinations suggested the potential interaction between pharmacodynamic components and targets. With the interaction between soya-cerebroside（an effective constituent of *Pinellia ternata* in Wendan Decoction）and low-density lipoprotein receptor as an example，the literature review showed that it might be one of the mechanisms for Wendan Decoction in the treatment of atherosclerosis. ***Conclusion*：**The SGRN-Trans model

作者介绍：王艳菁，Email：wangyanjing@sjtu.edu.cn，研究方向：人工智能药物设计。
通信作者：谭红胜，Email：tanhs@shsmu.edu.cn。
基金项目：中央高校基本科研业务费专项资金资助-上海交通大学医工交叉研究基金重点项目（编号：YG2021ZD02）；上海交通大学“交大之星（STAR）”资助项目（编号：20230101）。
优先出版：https://link.cnki.net/urlid/50.1046.R.20240819.0902.006（2024-08-19）

based on knowledge graph and attention mechanism proposed in this study can be widely used to predict the interaction between components and targets in the complex network system of classic TCM prescriptions, which provides a new tool for clarifying the pharmacodynamic material basis of classic TCM prescriptions and related mechanisms of action.

【Key words】Wendan Decoction; drug-target interaction; knowledge graph; graph neural network; attention mechanism

中医经典名方，泛指至今仍在中医临床被广泛应用、疗效确切、具有明显特色与优势的古代中医典籍所记载的方剂。近年来，我国出台了一系列方针政策及法律法规，积极鼓励研究和开发经典名方，运用现代科学技术揭示其科学内涵。国家中医药管理局办公室.国家药品监督管理局综合和规划财务司相继发布了《古代经典名方目录》第一批目录[1]和《古代经典名方关键信息表（7首方剂）》[2]，其中收载的经典名方温胆汤来源于古籍《备急千金要方》（唐代孙思邈），由六味中药组成（半夏、陈皮、枳实、竹茹、生姜、甘草）。具有理气化痰，和胃利胆的功效；主治胆胃不和，胆郁痰扰证。笔者基于温胆汤开展了历史沿革[3]、药材炮制[4]、基准物质[5]及疗效系统评价[6]等研究，提示温胆汤的临床应用广泛，对精神、心脑血管、消化及呼吸等多系统疾病的疗效确切，但是其药效物质基础及对应的靶点尚不明确。

由于中药复方通常可以看作含有多种化学成分的中药药材的混合，这些化学成分作为药物分子在药理上作用于不同的靶点进而调控多个人体内的生物机制[7]。尽管中医药的临床应用前景广阔，但其复杂的成分组成很难还原出其潜在的活性成分和生物学靶点之间的作用机制。大多数的中药药材包含大量的化学成分，其中只有一小部分在其疗效上发挥作用。通过实验来研究中药的药理机制需要识别其药物和靶点相互作用（Drug-Target Interaction，DTI），然而这个过程耗时且昂贵[8]。

在过去的几十年里，越来越多针对DTI识别的计算方法被用于虚拟筛选，以缩小药物和候选蛋白质的搜索空间，减少成本并加快药物发现和开发的效率[9]。一般来说，DTI预测方法可以分为两类：基于分子对接的方法和基于机器学习的方法，前者将小分子放置在目标蛋白的活性位点，通过不断改变配体的构象，确定小分子（配体）与目标大分子相互作用的最佳构象，预测它们的结合模式和亲和力，这种方法会产生非常大的搜索空间和计算量[10]，对于未知蛋白质的靶点存在一定限制。后者根据对药物和靶点数据的处理过程可以分为三个子类：基于相似性的方法（Similarity-based），基于特征的方法（Feature-based）和基于网络的方法（Network-based）[11]。其中基于网络的方法近年来得到了快速的发展。它用网络图的形式表示药物、靶点以及疾病、通路等生物学实体以及它们之间的关系[12]。这些网络基于知识或使用各种数据资源进行计算和推理。除了使用仅包含DTI的单类型网络进行DTI预测的方法之外[13]，用异质信息（heterogeneous data）构建包含多类实体和关系的网络图并基于此进行DTI预测的方法受到了广泛的推广和使用[14]。

本研究以中医药经典名方温胆汤为实例开展研究，提出了一种基于异质信息的DTI人工智能预测框架（Self-weighted Graph Relational Network-Transformer，SGRN-Trans）。利用公开数据库以及开放的生物医药知识图谱，融合异质信息，构建了温胆汤知识图谱（Wendan decoction knowledge graph，WDKG），将DTI预测作为基于WDKG的链路预测任务；该预测框架利用图神经网络（Graph Neural Networks，GNN）学习知识图谱中每个实体的低维嵌入表示，同时引入药物和靶点各自的结构特征，并搭载基于注意力机制的Transformer模型进行DTI预测；通过不同知识图谱嵌入（Knowledge Graph Embedding，KGE）模型、不同分类模型和其他DTI预测模型的比较分析，SGRN-Trans在DTI预测任务上具有优异的性能；通过分子对接和文献案例分析获得了具有潜在相互作用的成分-靶点相互作用对。

1 研究方法

1.1 知识图谱数据构建与可视化

本研究以中草药数据库TCMSP（Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform，TCMSP）和生物医药知识图谱DRKG（Drug Repurposing Knowledge Graph，DRKG）为数据基础，构建以“Head-Relation-Tail”（头-关系-尾）形式的三元组组成的WDKG。首先，从TCMSP中获取温胆汤六味中药（Herb）所包含的药

物分子（compound）和各药物分子关联的靶点或基因，去重并滤除无法在PubChem（https://pubchem.ncbi.nlm.nih.gov/）数据库中查找到的药物分子和无法在DrugBank（https://go.drugbank.com/）数据库中查询得到或者非人类靶点的基因（Gene）。保留下来的药物和靶点之间的相互作用关系作为一类三元组（Compound:Gene），同时也是本研究的DTI数据集。将获得的基因实体在DRKG提取出包含这些基因实体在内的所有三元组。分析统计WDKG中不同类实体的数量、来源、占比。将WDKG中的实体关系通过Neo4j（https://neo4j.com/）图数据库进行存储和可视化。

1.2 基于关系图卷积神经网络（relational graph convolution network，RGCN）的知识图谱嵌入学习

RGCN是一种适用于拥有多种关系类型的图数据的图神经网络，其可以令不同类型的关系通过不同的信息率进行聚合。本研究使用RGCN来聚合知识图谱中的节点信息，以通过知识图谱的多关系类型的拓扑信息，学习出各实体以及各类关系的嵌入表示。

整个知识图谱是一个图结构的数据，本课题组用$G=(V,T,X)$来表示，其中V表示图中所有节点，也即知识图谱中所有实体的集合；T表示图中所有边，也即知识图谱中所有三元组的集合；$X\in\mathbb{R}^{N\times C}$是所有实体的嵌入表示组成的特征矩阵，$N$即实体的个数$|V|$，每个实体的嵌入长度为$C$。对于一个输入的图$G$中的节点$i$，RGCN按照如下公式聚合节点信息。

$$x_i^{(l+1)}=\sigma\left(\sum_{r\in R}\sum_{j\in N_i^r}\frac{1}{c_{i,r}}W_r^{(l)}x_j^{(l)}+W_0^{(l)}x_i^{(l)}\right)\qquad(1)$$

其中R是所有关系类型的集合，N_i^r表示以关系r与节点i相连的节点的集合，$x_i^{(l)}$表示经过第l层RGCN后节点i的隐层特征，$W_r^{(l)}\in\mathbb{R}^{C\times C}$则表示第$l$层RGCN中代表关系$r$的信息转化矩阵，$W_0^{(l)}$为第$l$层RGCN的自连接信息转化矩阵，$c_{i,r}$是一个归一化系数，其可以由训练习得也可以预先指定，这里令其为$c_{i,r}=|N_i^r|$。σ是激活函数，这里使用Leaky ReLU，以提高训练速度和模型性能。

在整个KGE部分，本课题组使用连续的L层RGCN作为编码器，其中每一层的输出作为下一层的输入。本课题组将第L层的输出$z^{(L)}$作为所有实体的嵌入表示，然后利用一个解码器，通过函数$s:\mathbb{R}^C\times R^{C\times C}\times\left(\mathbb{R}^C\right)^{\mathrm{T}}\rightarrow R$为三元组$(h,r,t)$打分。令头实体$h$的嵌入向量$e_h=x_h^{(L)}$，尾实体$t$的嵌入向量$e_t=x_t^{(L)}$。为了处理有向关系，并合理利用各层RGCN学习得到的关系特异的参数矩阵，本研究采用了一种自权重打分（self-weight scoring）的方法，使得打分函数中的R与RGCN中学习得到的关系参数矩阵相关联，具体形式如下。

$$R_r=\frac{1}{L}\sum_{l=1}^{L}W_r^{(l)}\qquad(2)$$

1.3 药物及靶点的特征处理

通过KGE得到的药物分子和靶点基因的嵌入表示仅包含了来自知识图谱的拓扑信息，并不具有药物分子和靶点基因各自独有的特征，为了使特征更加丰富，我们引入药物分子和靶点基因各自的结构特征与其嵌入表示相融合。药物分子的特征为由RDKit计算得到的摩根分子指纹（Morgan Fingerprints）[15]，靶点的特征为由PyBioMed[16]计算得到的CTD描述符[17]。

1.4 基于Transformer的药物-靶点相互作用预测

在按上述过程对Drug和Target的嵌入特征和结构特征进行处理后，我们构造了一个基于Transformer的分类器，该分类器主要由编码器、解码器以及全连接层组成。对于一对药物-靶点(D_i,T_j)，我们将该药物降维后的嵌入e_{D_i}以及变换后的特征p_{D_i}作为编码器的输入，将靶点降维后的嵌入e_{T_j}和变换后的特征p_{T_j}作为解码器的输入。编码器将对输入的特征进行学习，并将信息传递给解码器，解码器将其与输入解码器的特征进行融合，并最终由全连接层进行预测分类。每层编码器和解码器的结构如图2所示，每层编码器由Self-Attention层，Normalization层和Feed Forward层组成，而每层解码器相对于编码器多了一个Encoder-Decoder Attention层。

1.5 模型评估及比较

本研究所提出的DTI预测框架SGRN-Trans主要由KGE模型SGRN和分类器Transformer组成，我们将基于WDKG数据集，分步验证该框架各部分的优越性：首先比较不同特征（嵌入特征和结构特征）方案对模型预测性能的影响；其次，将SGRN与不同类型的经典主流KGE算法TransE，TransR，ComplEx，DistMult和ConvKB进行比较；接着将Transformer与：①在各场景下都取得优异表现的卷积神经网络（convolutional neural network，CNN），②在推荐系统中表现突出的NFM模型（neural factorization machines），③经典机器学习算法随机森林（random forest，RF）进行比较。最后，评估SGRN-Trans的整体性能，将其与其他各类型的最先进的关联预测算法进行比较，包括：DeepDTA[18]、RF和NFM、KGRN、KGE_NFM[19]。

在评估知识图谱嵌入的性能时，本课题组计算每个真实三元组的得分并将它们降序排列，得到每个三元组在其中的排名，并计算它们的平均倒数排名（mean reciprocal rank，MRR）以及排名前N的三元组所占的比例（Hits@N，N=1,3,10,50,100）。这些评估指标适用于评估知识图谱嵌入的整体性能，却并不适用于针对知识图谱中特定关系类型的下游任务。在DTI预测任务中，将其处理为一个二分类任务，为了评估其效果，引入了AUROC（area under the receiver operating characteristic curve，AUROC）和AUPR（area under precision-recall curve，AUPR）这2个在分类任务中应用广泛的评价指标，同时还引入了F1-score，其能够评估分类模型在不平衡数据集（在本研究中正样本和负样本的比例为1∶4）上的综合性能。

1.6 模型参数设置

模型使用Adam（Adaptive Moment Estimation）算法作为训练所有模型时的优化算法，在分别训练SGRN与Transformer时预先对以下超参数进行调整并最终设置见表1。

表 1　SGRN 和 Transformer 训练过程超参数设置

项目	学习率（learning rate）	一代训练（epoch）	批次大小（Batch size）
SGRN	0.001	300	10 000
Transformer	0.000 1	50	256

KGE 的嵌入向量的长度被设置为 200，主成分分析（Principal Component Analysis，PCA）降维后的主成分个数被设置为 120，Transformer 的 feed-forward 层的维度被设置为 512。为了防止过拟合，本课题组采用了早停机制（Early Stopping），即当测试集上的损失连续 7 个 epoch 没有得到降低时，则结束训练过程。本研究中所有训练过程都在 Ubuntu 16.04.7 LTS 上进行。CPU 芯片是 Intel（R）Xeon（R）CPU E5-2640 v3@2.60GHz，具有 32 个逻辑 CPU。GPU 的型号为 GeForce GTX 1070，具有 8G 内存。本研究中所有模型都使用 GPU 训练。使用的 Python 版本为 Python 3.8.3，Pytorch 的版本为 1.8.0。

2　结　果

2.1　SGRN-Trans 深度学习预测框架概述

药物与靶点之间的相互作用往往具有高度复杂性，涉及多种生物学特征和结构特征。Transformer 模型的多头自注意力机制能够同时关注输入序列中的不同部分，从而捕捉全局信息和局部依赖关系，并动态调整权重，使模型更好地理解药物和靶点之间的相互作用，适用于处理复杂的生物数据。本研究提出了一种基于知识图谱异质信息和注意力机制的药物-靶点相互作用预测框架：SGRN-Trans，由 3 部分组成：①基于 RGCN 图神经网络进行知识图谱嵌入学习以得到每个实体的低维嵌入表示；②利用 PCA 将药物和靶点的嵌入表示进行特征降维，引入药物和靶点各自的结构特征信息；③利用 Transformer 模型进行药物-靶点相互作用预测。整个预测框架的示意图见图 1。

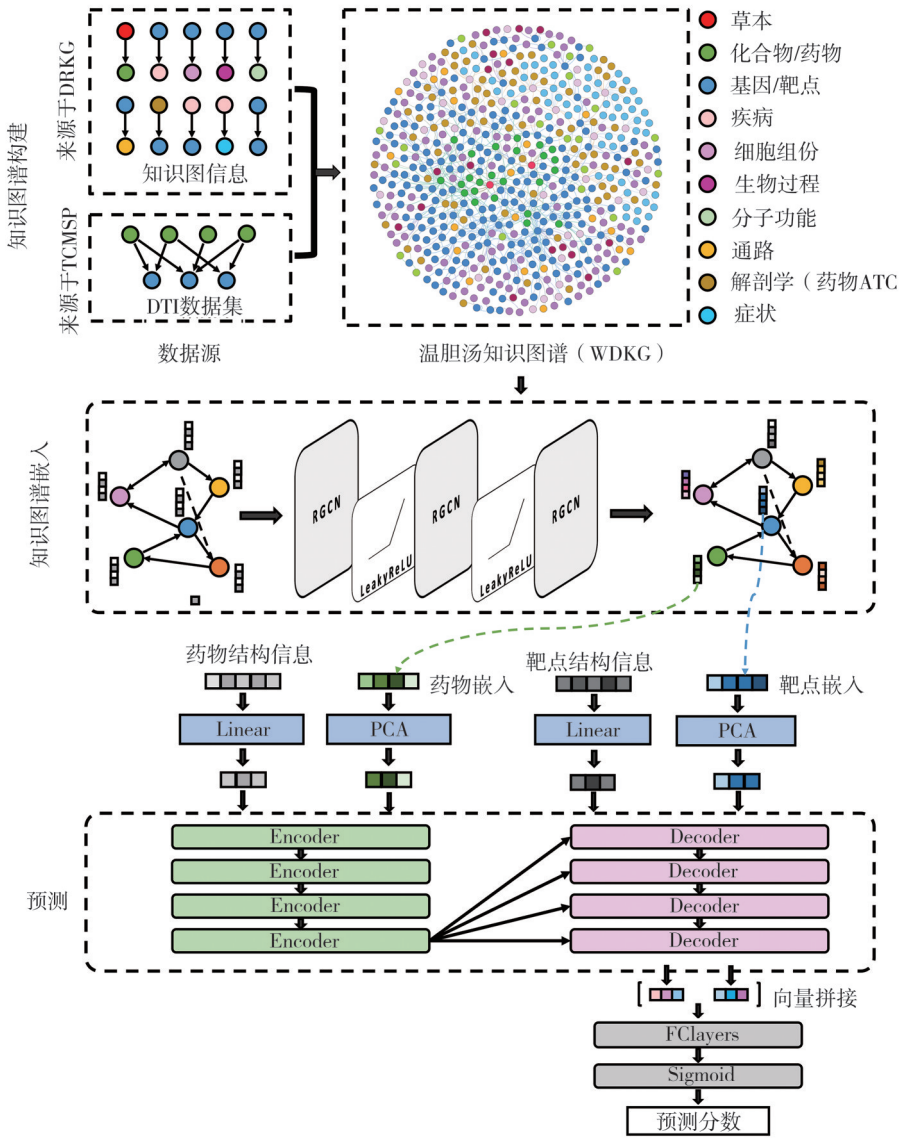


图 1　SGRN-Trans 框架示意图

2.2 WDKG的构建和可视化

本研究将从DRKG中提取得到的关系类型进行归纳整理和重新命名，最终WDKG共有169 165个三元组，包含26类关系，并涉及包含10个类型的共14 292个实体。表2显示了WDKG中不同类实体的数量及来源；表3展示了对收集数据进行关系归纳后的所有关系类型、各关系类型所要求的头实体和尾实体的类型以及各关系类型所占的百分比。图2展示了在Neo4j存储的知识图谱中都有哪些类型的节点以及它们之间是如何连接的。

表2 WDKG中各类实体的分布以及数据来源

实体（Entity Type）	TCMSP	DRKG
Herb	6	–
Compound/Drug	699	–
Gene/Target	685	–
Anatomy	–	190
Biological Process	–	6 957
Disease	–	1 935
Molecular Function	–	1 411
Cellular Component	–	608
Pathway	–	1 387
Symptom	–	414

2.3 模型性能评价分析

在SGRN-Trans中，将SGRN习得的嵌入特征以及药物和靶点自身的结构特征同时作为分类器的输入。分别只用“知识图谱嵌入特征（以Transformer-e表征）”及“药物和靶点的结构特征（以Transformer-s表征）”作为Transformer的输入，以比较这两部分特征对于预测的影响。表4显示了上述3种情况十折交叉验证下的AUROC和AUPR值，当使用SGRN-Trans时，在两个评价指标上的分值都是最高的，使其特征更加丰富，从而进一步提升模型预测的鲁棒性。

表5展示了不同知识图谱嵌入模型在本研究数据集上的性能比较结果。在MRR上，SGRN取得了最好的分值（0.0554），优于其他模型；在Hits@1，Hits@3以及Hits@10上，SGRN均为最高分且高于其他模型。分析表明，SGRN在本数据集上的整体性能最好，且可以为事实三元组赋予相对更高的分值。

图3展示了SGRN搭载不同分类算法时的评估结果，对于DTI预测任务，Transformer和NFM用于分类的性能更优秀，两者在AUROC和AUPR这两项指标上相当，但在可以体现模型整体的分类性能的F1-score上，SGRN-Trans优于SGRN+NFM。将SGRN-Trans与其他研究人员提出的关联

表3 WDKG中各类关系类型及比例分布

数据库来源	实体类型	Relation Name关系：名称	比例Ratio（%）
TCMSP	Herb：Compound	Contain：：Herb：Compound	0.45
	Compound：Gene	Interaction：：Compound：Gene	3.41
DRKG	Anatomy：Gene	Regulation：：Anatomy：Gene	16.71
	Disease：Anatomy	Localization：：Disease：Anatomy	1.13
	Disease：Gene	Association：：Disease：Gene	1.70
		Interaction：：Disease：Gene	0.01
		Regulation：：Disease：Gene	0.49
	Disease：Symptom	Presents：：Disease：Symptom	1.91
	Gene：Biological Process	Participate：：Gene：Biological Process	34.92
	Gene：Cellular Component	Participate：：Gene：Cellular Component	2.72
	Gene：Disease	ExpressionAssociation：：Gene：Disease	4.62
		MutationAssociation：：Gene：Disease	0.54
		PathologicalAssociatiton：：Gene：Disease	3.35
	Gene：Gene	Association：：Gene：Gene	1.50
		Binding：：Gene：Gene	1.48
		Colocalization：：Gene：Gene	0.05
		Covariation：：Gene：Gene	0.06
		Interaction：：Gene：Gene	3.93
		Other：：Gene：Gene	3.25
		PBCP：：Gene：Gene	0.56
		Reaction：：Gene：Gene	1.03
		Regulation：：Gene：Gene	4.77
		Same：：Gene：Gene	0.07
		Signaling：：Gene：Gene	0.15
	Gene：Molecular Function	Participation：：Gene：Molecular Function	4.62
	Gene：Pathway	Participation：：Gene：Pathway	6.60

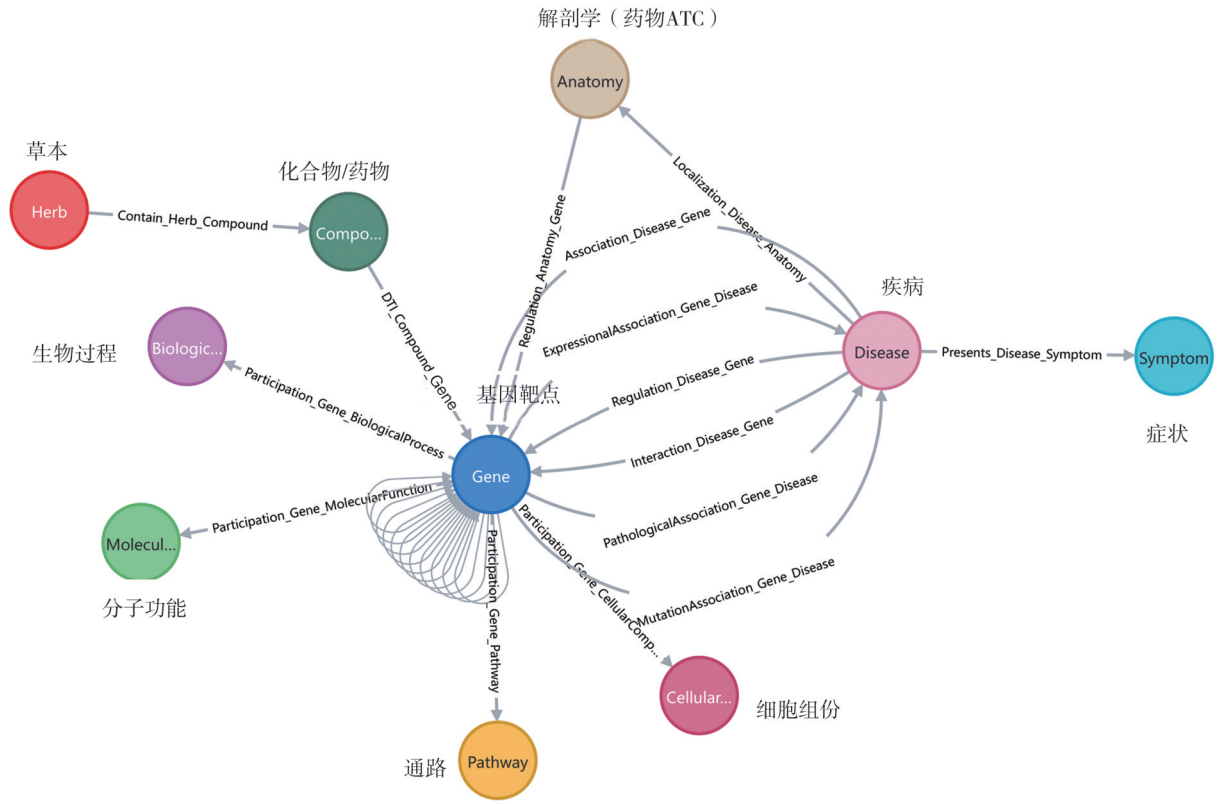


图2　Neo4j 中 WDKG 包含的节点类型以及节点之间的关系

表 4　不同特征方案对 DTI 预测性能的影响

Transformer 输入	AUROC	AUPR
SGRN-Trans	0.977 4	0.910 5
Transformer-e	0.976 6	0.909 9
Transformer-s	0.976 3	0.908 6

表 5　SGRN 与其他各类知识图谱嵌入模型的性能比较

类别	算法模型	MRR	Hits@N				
			N=1	N=3	N=10	N=50	N=100
Translational Distance	TransE	0.040 6	0.004 9	0.026 0	0.097 3	0.368 5	0.539 2
	TransR	0.046 8	0.011 6	0.035 3	0.101 7	0.345 0	0.514 1
Semantic Matching	ComplEx	0.030 0	0.006 1	0.019 5	0.062 4	0.239 1	0.370 4
	DistMult	0.026 2	0.006 4	0.017 3	0.050 6	0.200 3	0.338 3
Neural Network	ConvKB	0.009 5	0.001 2	0.006 1	0.014 4	0.064 9	0.124 2
Proposed	SGRN	0.055 4	0.017 5	0.045 6	0.113 4	0.364 6	0.531 9

预测算法进行对比（图 4），从箱型图中可以看出，SGRN-Trans（AUROC=0.9785，AUPR=0.913 5，F1-score=0.856 1）在各项指标上的表现都是最好的，在本研究任务中相对其他先进算法具有一定的优势。

2.4　对接结果分析

为了探究 SGRN-Trans 的实际预测效果，我们用训练好的 SGRN-Trans 模型对所有药物-靶点对进行预测打分，并剔除训练集中出现过的组合，对首次发现未经报道的 8 组药物-靶点对进行分子对接以进一步验证。从 PubChem 数据库中得到药物分子的 3D 结构，并从 PDB 数据库（https://www.rcsb.org/）中选取靶点蛋白的 3D 结构，使用 AutoDock 进行分子对接。分子对接结合能结果见表 6，对接构象见图 5。8 组分子对接的平均结合能为-8.445 kcal/mol，其中 6 组的结合能均在-7 kcal/mol 以下，更有 3 组结合能在-10 kcal/mol 以下，表明这些药物分子和靶点蛋白之间可能存在潜在的相互作用。

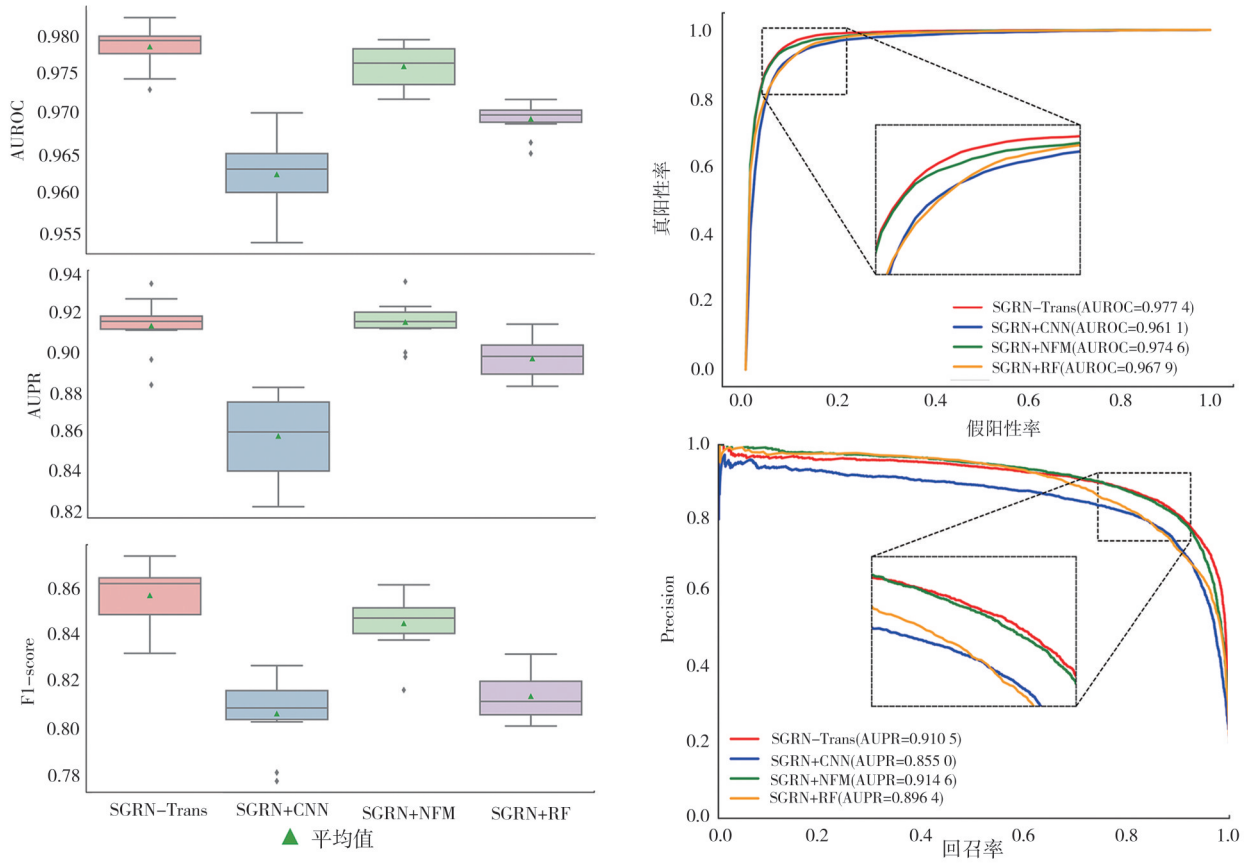


图3　将SGRN搭载不同的分类算法与SGRN-Trans比较

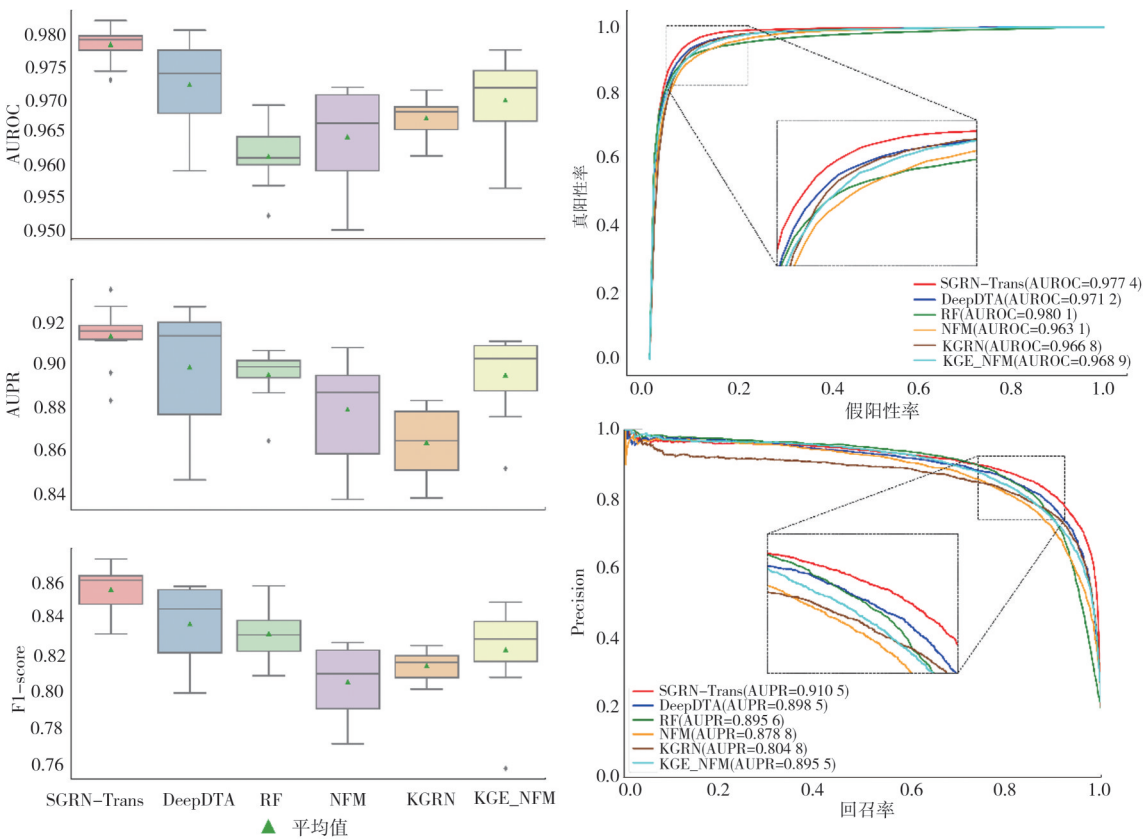


图4　SGRN-Trans与其他先进的DTI预测算法比较

表6　未报道过可能作用的8组分子对接相关信息

成分名称	靶标名称	PDB-ID	结合能(kcal/mol)	网格中心(xyz)	网格大小(xyz)
Soya-cerebroside	LDLR	2FCW	-10.50	13.7,15.2,29.3	40,60,80
Kanzonol E	HNF1A	1IC8[35]	-11.45	45.9,8.3,22.6	118,60,110
alpha-Inosine	FKBP1A	1BL4[36]	-7.63	10.0,15.7,2.6	70,60,70
Kanzonol E	AHCY	1LI4	-8.18	58,-4.8,104.4	120,70,100
Docosanoic acid	LDLR	2FCW	-5.60	13.7,15.2,29.3	40,60,80
Kanzonol E	CHRNA2	AF-Q15822-F1[37]	-7.38	-15.4,-14.9,11.1	100,70,100
Kanzonol E	CPT2	AF-P23786-F1	-11.06	-3.5,-1.7,0.9	100,70,100
(S)-(-)-Citronellic acid	ERBB2	3PP0[38]	-5.76	17.6,15.2,27.4	100,100,100

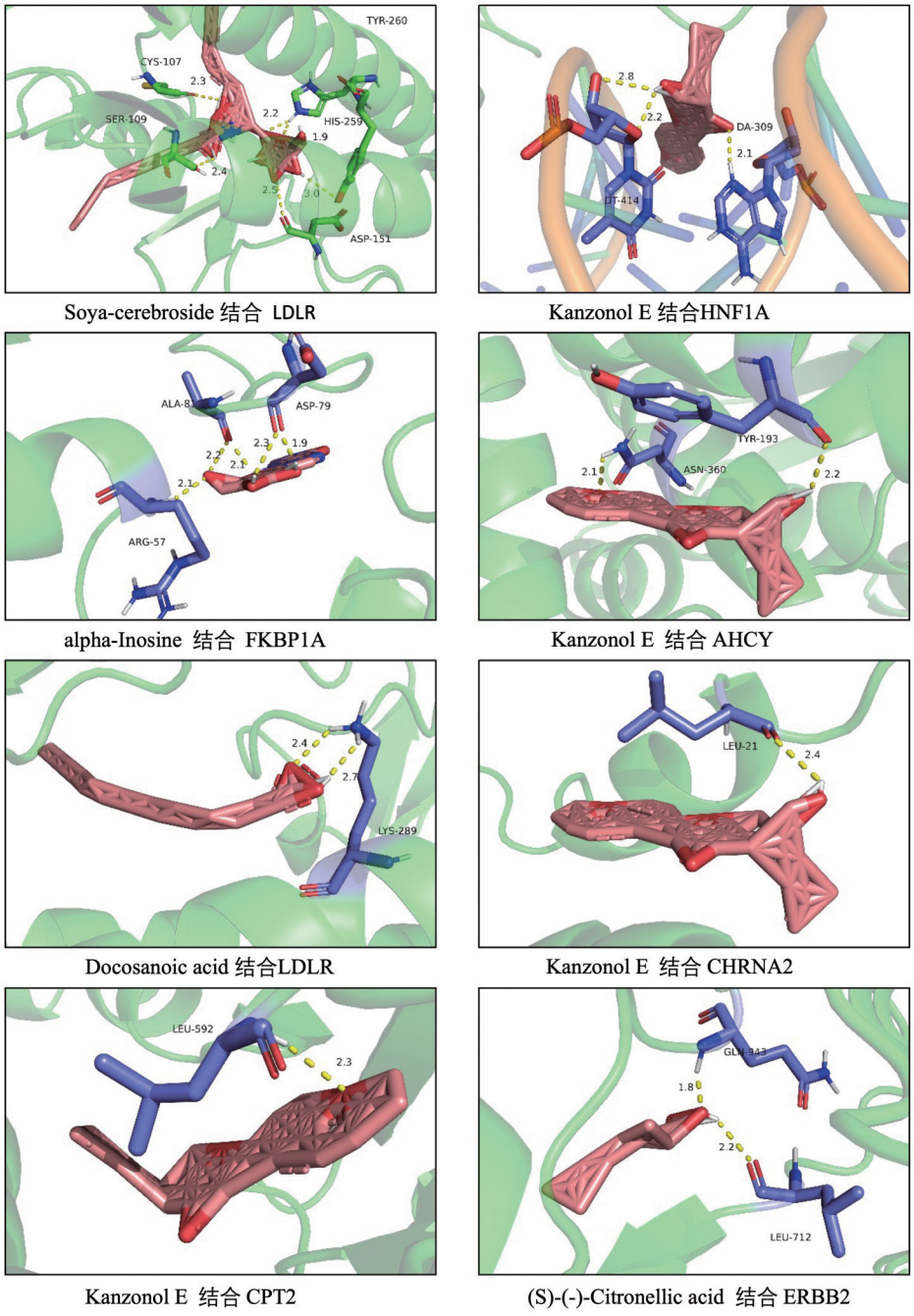


图5　前20组药物-靶点中未知的8组的分子对接可视化结果

3 讨 论

在现代中医证候临床研究中，血脂代谢异常是中医痰证的物质基础之一[20]，也是动脉粥样硬化性病变的重要因素。温胆汤作为化痰的经典代表方剂，在治疗动脉粥样硬化性病变显示出良好疗效[21-22]。研究表明，温胆汤能够有效降低血甘油三酯（triglyceride，TG）、总胆固醇（total cholesterol，TC）和低密度脂蛋白胆固醇（low-density lipoprotein cholesterol，LDL-C）水平，有效抑制 ox-LDL 诱导的巨噬细胞泡沫化[23]，改善动脉粥样硬化性心脏病患者临床症状及预后[24]。然而，其作用机制尚未明确。

低密度脂蛋白受体（low density lipoprotein receptor，LDLR）的主要作用是通过受体内吞作用降解低密度脂蛋白（LDL）[25]。通过抑制 LDLR 的合成，减少 LDL 摄取，可以防止胆固醇过度积聚。因此，LDLR 在维持体内 LDL 水平和胆固醇稳态中起到至关重要的作用。当血浆中 LDL 和自由基的含量增多时，就会导致 oxLDL（oxidized LDL）增多，从而增加动脉粥样硬化的风险[25]。此外，动脉粥样硬化的另一个特征是动脉壁中富含脂质的巨噬细胞泡沫细胞的积累[26]。Soya-cerebroside（大豆脑苷脂）是温胆汤中半夏的有效成分，能抑制 FAK（Focal adhesion kinase）-Src 信号通路的效应，但其作用的机制尚不明确[27]。FAK-Src 信号通路是 oxLDL 影响巨噬细胞迁移的机制中的重要部分，而 oxLDL 由 LDL 与血浆中自由基氧化而得，血浆中 LDL 的水平会影响 FAK-Src 信号通路的表达[28]。

在本研究中，模型预测和分子对接结果显示 soya-cerebroside 和 LDLR 之间存在相互作用，结合上述相关的文献调研，推测可能存在以下机制：soya-cerebroside 与 LDLR 相互作用，通过调控 LDLR 的蛋白功能，调节 LDL 的水平，进而抑制 FAK-Src 信号通路，抑制 oxLDL 诱导的细胞迁移，达到缓解动脉粥样硬化的效果。soya-cerebroside 具有作为治疗动脉粥样硬化等血管疾病药物的潜力。

中医经典名方中活性成分的作用靶点及机制研究，对揭示其生物学基础和中药药效的现代科学内涵至关重要。本研究首次将人工智能深度学习技术应用于中药经典名方温胆汤，提出了基于知识图谱和 Transformer 的 DTI 预测框架 SGRN-Trans。研究利用多生物数据源构建温胆汤知识图谱（WDKG），通过 GNN 学习知识图谱中每个实体的低维嵌入表示，同时引入药物和靶点的结构特征，并基于注意力机制的 Transformer 模型进行 DTI 预测。实验结果表明，与其他先进的 DTI 预测模型对比，SGRN-Trans 具有一定的性能优势。对接实验和文献研究表明，温胆汤中半夏的有效成分 soya-cerebroside 和 LDLR 存在潜在的相互作用，对温胆汤治疗动脉粥样硬化的科学内涵提供重要的支持。本研究为温胆汤的药效成分及作用机制提供了新的研究思路，未来，仍需在其他优质 DTI 数据集上对 SGRN-Trans 进行应用，以评估该模型的泛化性。进一步通过科学、规范的动物及细胞生物学实验进行验证。

参 考 文 献

[1] 国家中医药管理局. 古代经典名方目录（第一批）[EB/OL].（2018-04-13）[2024-05-26]. http://www.natcm.gov.cn/kejisi/zhengcewenjian/2018-04-16/7107.html.

State Administration of Traditional Chinese Medicine. Catalogue of Ancient Classic Prescriptions（First Batch）[EB/OL].（2018-04-13）[2024-05-26]. http://www.natcm.gov.cn/kejisi/zhengcewenjian/2018-04-16/7107.html.

[2] 国家中医药管理局办公室. 国家药品监督管理局综合和规划财务司关于发布《古代经典名方关键信息考证原则》《古代经典名方关键信息表（7 首方剂）》的通知[EB/OL].（2020-11-11）[2024-05-26]. https://www.nmpa.gov.cn/xxgk/fgwj/gzwj/gzwjyp/20201111091109170.html.

Notice of the Office of the State Administration of Traditional Chinese Medicine and the Comprehensive and Planning Finance Department of the State Medical Products Administration on Issuing the "Key Information Research Principles for Ancient Classic Famous Prescriptions" and "Key Information Table for Ancient Classic Famous Prescriptions（7 Prescriptions）" [EB/OL].（2020-11-11）[2024-05-26]. https://www.nmpa.gov.cn/xxgk/fgwj/gzwj/gzwjyp/20201111091109170.html.

[3] 王 婧，董丽萍，时欹未，等. 温胆汤历史沿革研究进展[J]. 世界中医药，2020，15（14）：2193-2198.

Wang J，Dong LP，Shi XW，et al. Study on the progress of historical evolution of Wendan Decoction[J]. World Chin Med，2020，15（14）：2193-2198.

[4] 董丽萍，赵家莹，詹 梁，等. 经典名方温胆汤中枳实模拟古法的炮制工艺与质量控制研究[J]. 上海中医药杂志，2019，53（10）：90-96.

Dong LP，Zhao JY，Zhan L，et al. Technology and quality control for processing Fructus Aurantii Immaturus simulating ancient method in classic recipe Wendan Decoction[J]. Shanghai J Tradit Chin Med，2019，53

(10):90-96.

[5] 杨思雨,董丽萍,谭红胜,等.温胆汤基准样品量值传递研究[J].中成药,2023,45(6):1788-1794.

Yang SY, Dong LP, Tan HS, et al. Quality value transmitting of Wendan Decoction reference samples[J]. Chin Tradit Pat Med, 2023, 45(6): 1788-1794.

[6] Yang YT, Chen R, Li CX, et al. A synthetic external control study comparing the clinical efficacy of Wendan Decoction and 19 antidepressants[J]. Int J Neuropsychopharmacol, 2023, 26(10): 739-746.

[7] Wang N, Li P, Hu XC, et al. Herb target prediction based on representation learning of symptom related heterogeneous network[J]. Comput Struct Biotechnol J, 2019, 17: 282-290.

[8] Xu ZG. Modernization: one step at a time[J]. Nature, 2011, 480(7378): 90-92.

[9] Bagherian M, Sabeti E, Wang K, et al. Machine learning approaches and databases for prediction of drug-target interaction: a survey paper[J]. Brief Bioinform, 2021, 22(1): 247-269.

[10] Ru XQ, Ye XC, Sakurai T, et al. Current status and future prospects of drug-target interaction prediction[J]. Brief Funct Genomics, 2021, 20(5): 312-322.

[11] Ding H, Takigawa I, Mamitsuka H, et al. Similarity-based machine learning methods for predicting drug-target interactions: a brief review[J]. Brief Bioinform, 2014, 15(5): 734-747.

[12] Wan FP, Hong LX, Xiao A, et al. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions[J]. Bioinformatics, 2019, 35(1): 104-111.

[13] Cheng FX, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference[J]. PLoS Comput Biol, 2012, 8(5): e1002503.

[14] Chen X, Liu MX, Yan GY. Drug-target interaction prediction by random walk on the heterogeneous network[J]. Mol Biosyst, 2012, 8(7): 1970-1978.

[15] Morgan HL. The generation of a unique machine description for chemical structures-a technique developed at chemical abstracts service[J]. J Chem Doc, 1965, 5: 107-113.

[16] Dong J, Yao ZJ, Zhang L, et al. PyBioMed: a python library for various molecular representations of chemicals, proteins and DNAs and their interactions[J]. J Cheminform, 2018, 10(1): 16.

[17] Dubchak I, Muchnik I, Holbrook SR, et al. Prediction of protein folding class using global description of amino acid sequence[J]. Proc Natl Acad Sci U S A, 1995, 92(19): 8700-8704.

[18] Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug-target binding affinity prediction[J]. Bioinformatics, 2018, 34(17): i821-i829.

[19] Ye Q, Hsieh CY, Yang ZY, et al. A unified drug-target interaction prediction framework based on knowledge graph and recommendation system[J]. Nat Commun, 2021, 12(1): 6775.

[20] 陈玄晶,徐丹苹,陈小光,等.冠心病中医痰证与客观化指标关系研究概述[J].中华中医药杂志,2017,32(7):3089-3092.

Chen XJ, Xu DP, Chen XG, et al. Overview of the relationship between phlegm syndrome and objective indexes in coronary heart disease[J]. China J Tradit Chin Med Pharm, 2017, 32(7): 3089-3092.

[21] Xu JH, Huang YM, Ling W, et al. Wen Dan Decoction for hemorrhagic stroke and ischemic stroke[J]. Complement Ther Med, 2015, 23(2): 298-308.

[22] 何皓颋.加味温胆汤对气虚痰瘀型冠状动脉粥样硬化性心脏病多支病变不完全血运重建患者心绞痛的影响[J].中医临床研究,2019,11(10):73-75.

He HT. Effects of the Wendan Decoction on angina of incomplete revascularization patients with coronary heart disease of the Qixu Tanyu type[J]. Clin J Chin Med, 2019, 11(10): 73-75.

[23] 陈玄晶,钟碧莹,吴焕林,等.温胆汤抑制氧化低密度脂蛋白诱导的巨噬细胞泡沫化[J].中国病理生理杂志,2020,36(11):1952-1959.

Chen XJ, Zhong BY, Wu HL, et al. Inhibitory effect of Wendan Decoction on formation of foam cells induced by ox-LDL[J]. Chin J Pathophysiol, 2020, 36(11): 1952-1959.

[24] 陈玄晶,陈小光,徐丹苹,等.温胆汤调节血脂异常系统评价[J].中国中医药信息杂志,2017,24(8):93-98.

Chen XJ, Chen XG, Xu DP, et al. Systematic assessment on effects of Wendan Decoction on dyslipidemia[J]. Chin J Inf Tradit Chin Med, 2017, 24(8): 93-98.

[25] Jeon H, Blacklow SC. Structure and physiologic function of the low-density lipoprotein receptor[J]. Annu Rev Biochem, 2005, 74: 535-562.

[26] Chen YL, Kennedy DJ, Ramakrishnan DP, et al. Oxidized LDL-bound CD36 recruits an Na^+/K^+-ATPase-Lyn complex in macrophages that promotes atherosclerosis[J]. Sci Signal, 2015, 8(393): ra91.

[27] Lee HP, Wang SW, Wu YC, et al. Soya-cerebroside inhibits VEGF-facilitated angiogenesis in endothelial progenitor cells[J]. Food Agric Immunol, 2020, 31: 193-204.

[28] Park YM, Febbraio M, Silverstein RL. CD36 modulates migration of mouse and human macrophages in response to oxidized LDL and may contribute to macrophage trapping in the arterial intima[J]. J Clin Invest, 2009, 119(1): 136-145.

（责任编辑：周一青）