

基于多方法的三角测量临床因果知识图谱构建研究

1. 引言

1.1 研究背景与动机

随着医疗信息化的快速发展和电子健康记录（EHR）的广泛应用，医疗数据呈现出规模庞大、结构复杂、关系多样的特点。如何从海量的医疗数据中挖掘出可靠的因果关系，构建结构化的医疗知识图谱，已成为医疗人工智能领域的重要研究方向。传统的医疗知识图谱主要依赖专家知识和文献挖掘，存在知识更新滞后、覆盖面有限、主观性强等问题。

因果推断作为统计学和机器学习的重要分支，为从观察数据中发现因果关系提供了理论基础和方法工具。然而，单一的因果发现方法往往存在算法偏差、假设限制等问题，难以保证结果的可靠性和稳定性。特别是在医疗领域，由于数据的复杂性、噪声的存在以及伦理约束等因素，传统的随机对照试验（RCT）往往难以实施，使得基于观察数据的因果推断显得尤为重要。

本研究的核心动机在于：（1）解决单一因果发现方法的局限性，通过多方融合提高因果关系识别的准确性；（2）建立系统性的证据评估框架，量化因果关系的可信度；（3）构建面向临床应用的结构化知识图谱，支持智能诊疗和决策辅助。

1.2 三角测量：多维度因果可信度评估的创新框架

在因果推断研究中，单一算法或证据来源往往存在偏倚或不确定性，导致推断结果在不同数据集或模型设定下表现不一致。为克服这一局限性并提升因果关系识别的稳健性与可信度，本研究使用了“方法三角测量”（的理论框架。该框架借鉴流行病学与社会科学研究中的“三角验证”思想，将不同类型的因果证据在统一的量化体系下进行综合评估，从而实现对每条候选因果边的多维可信度判定。

本研究构建的三角测量框架包含结构一致性（Sstruct）、参数拟合度（Sparam）、中介机制支持（Smed）和专家定向知识（Sexpert）四个维度。针对每条候选因果边 e 定义联合置信度函数如下：

$$Conf(e) = \alpha \cdot S_{struct}(e) + \beta \cdot S_{param}(e) + \gamma \cdot S_{med}(e) + \delta \cdot S_{expert}(e)$$

其中，参数 $\alpha, \beta, \gamma, \delta$ 分别为各维度的权重，满足 $\alpha + \beta + \gamma + \delta = 1$ 。该函数通过加权整合不同来源的证据，实现了因果可信度的综合量化。

结构一致性维度，该维度主要用于衡量因果边在多种结构学习算法中的稳定性。

研究选取了多种不同原理的因果发现算法（包括 PC 算法、贪婪等价搜索、爬山算法及树搜索等），计算边在不同算法结果中的出现频率与方向一致性，以此反映结构发现的跨算法稳健性。该维度权重最高，体现其在因果关系识别中的核心作用。参数拟合维度，该维度用于评估因果边对模型拟合优度的贡献。通过结合最大似然估计（MLE）、贝叶斯估计、期望最大化（EM）算法及结构方程模型（SEM）等多种参数学习方法，计算边级似然增益 $\Delta LL(e) = LL_{full} - LL_{drop}(e)$ ，中介机制维度，该维度从机制层面验证因果关系的合理性。基于贝叶斯中介分析方法，计算直接效应、间接效应及总效应，并检验中介路径的显著性与效应强度，从而揭示因果边在潜在生物学或临床机制中的作用位置。该维度的引入，使因果可信度的评估不仅依赖统计显著性，也考虑机制合理性。专家定向维度，该维度融合领域专家知识、临床经验及节点类型信息（如疾病、药物、检验等），对因果关系的生物学合理性进行定性判断。通过结合专家标注与大语言模型（LLM）的定向推理结果，对统计模型难以捕获的领域先验进行补充，确保评估结果兼具理论可解释性与实践相关性。

综上，证据三角测量框架通过多维度证据的系统性整合，实现了对因果可信度的定量化、可解释化与可调控化评估。其主要创新体现在以下四个方面：（1）证据互补性：整合结构、参数、机制及知识四类证据，减少单一方法带来的偏差；（2）量化可解释性：各维度具有明确的统计或领域含义，使最终置信度结果具 $\alpha, \beta, \gamma, \delta$ 进行灵活优化，实现框架的适应性扩展；（4）质量控制与缺失补偿机制：针对缺失或估算数据引入置信修正，确保综合评估的完整性与稳健性。

总体而言，本研究提出的“证据三角测量”框架实现了从结构识别、参数验证到机制解释的全流程因果可信度量化，构建了一种兼具理论严谨性与应用灵活性的多源证据融合方法，为复杂系统中的因果推断研究提供了一种新的方法论路径。

1.3 临床因果知识图谱：结构化知识表示的六元组模型

传统的医学知识图谱多基于实体 - 关系 (entity - relation) 三元组结构 (h, r, t) ，其优势在于语义表达简洁、结构清晰。然而，三元组模型通常仅能描述静态的相关性或语义关联，难以充分刻画因果关系的方向性、不确定性与多层次证据支撑，从而限制了其在因果推理与临床决策中的应用。为弥补这一不足，本研究在传统知识图谱的基础上提出了面向因果推断的六元组临床知识图谱模型，形式化定义如下：

$$KG = (V, E_{core}, R, W, \Theta, \Phi)$$

该模型在结构上对传统图谱进行了系统扩展，旨在实现因果结构的显式表达与多源证据的定量整合，具体组成及功能如下。

(1) V : 变量节点集合 (Variable Set)

节点集合 V 包含临床领域中的多类型医学实体，包括疾病 (Disease)、药物 (Drug) 与检验指标 (Laboratory Test) 等。每个节点均携带领域语义标注、实体类型、出现频率及网络拓扑特征 (如度中心性、介数中心性等)，以支持后续的语义推理与结构分析。

(2) E_{core} : 核心因果边集合 (Causal Edge Set)

核心边集合 E_{core} 由经“证据三角测量”框架 (Evidence Triangulation) 筛选出的高置信度因果关系构成。每条边均以有向形式 $v_i \rightarrow v_j$ 表示因果方向，并附带置信度、数据来源及边类型等属性。该集合构成因果知识图谱的核心骨架，确保了图谱结构的统计稳健性与逻辑一致性。

(3) R : 关系类型集合 (Relation Type Set)

关系类型集合 R 定义了图谱中因果关系的语义类别，包括疾病-药物关系 (disease-drug)、疾病-检验关系 (disease-test) 等。多层次关系类型的引入使图谱能够支持从宏观因果结构分析到微观机制解释的多粒度查询与推理。

(4) W : 边级权重属性 (Edge Weight Attributes)

权重属性 W 存储每条因果边的多维置信度评分，包括结构一致性、参数拟合度、中介支持度及专家定向权重四个子分量 ($S_{struct}, S_{param}, S_{med}, S_{expert}$)。此外，该属性还记录数据质量标识与证据来源可靠性，用于在下游分析中动态调整权重，实现对因果关系强度的定量表征。

(5) Θ : 参数属性 (Parameter Attributes)

参数属性 Θ 用于存储多方法参数学习的结果，包括条件概率表 (Conditional Probability Table, CPT)、边条件概率、似然增益值及其他统计估计量。该参数层使得图谱不仅可作为知识载体，还能支持基于贝叶斯推理或结构方程建模的概率计算，从而量化因果关系的不确定性与方向性。

(6) Φ : 效应属性 (Effect Attributes)

效应属性 Φ 承载中介分析与因果效应估计结果，包括间接效应 (Indirect Effect, IE)、总效应 (Total Effect, TE)、中介比例 (Mediation Ratio, MR)、高密度区间 (High-Density Interval, HDI) 以及效应显著性概率等。该部分为图谱提供了机制层面的解释能力，使其能够从结果层面验证因果链条的合理性与稳定性。

该模型的创新之处在于将因果知识的多个维度进行了系统性的结构化表示：

- 节点集合 V ：包含临床变量（疾病、药物、检验指标），每个节点携带领域语义标注；
- 核心边集合 E_{core} ：通过证据三角测量筛选出的高置信度因果边，确保知识图谱的核心骨架具有较高的可信度；
- 关系类型集合 R ：定义了精细化的因果关系类型（直接因果、疾病-药物、中介路径等），支持不同粒度的因果查询；
- 权重属性 W ：存储每条边的四维证据评分，提供因果关系强度的量化依据；
- 参数属性 Θ ：整合多方法参数学习的结果（条件概率表、似然增益等），支持概率推理；
- 效应属性 Φ ：记录中介分析的效应估计（间接效应、总效应、HDI 区间等），支持机制解释。

这一六元组模型不仅保留了传统知识图谱的结构化表示优势，更重要的是将因果推断的不确定性、多方法证据和机制解释有机整合，形成了一个既具有统计

严谨性又具有临床可解释性的知识表示框架。

2. 研究内容

本研究构建了一套端到端的集成化流程，旨在实现从原始临床数据到结构化因果知识的自动化、高置信度转化。该流程以“因果知识发现—证据三角测量—知识图谱构建”为核心思想，通过六个紧密衔接的核心模块，系统性地解决了临床因果推断中的关键挑战。

2.1 数据预处理

为应对临床数据固有的高维、稀疏与噪声特性，本研究首先建立了一套严谨的数据治理流程。该流程起始于数据清洗与降维，通过系统性检测剔除无信息的全空变量，并实施一种基于临床变量前缀分类的特征选择策略，以降低计算复杂度的同时保留核心信息。随后，我们建立多维度质量评估体系，通过众数填充保障数据完整性，通过变量命名标准化实现一致性，并通过保留核心代表性样本来平衡计算效率与数据代表性。最终，所有经过治理的数据被统一转换为标准的 CSV 格式，为后续所有算法模块提供一个规范、高质量的统一数据接口。

2.2 基于多方法学的因果结构发现

为了克服单一因果发现算法在处理复杂真实数据时可能存在的局限性和偏见，本研究设计并实施了一个基于多方法学集成的因果结构发现框架。该框架通过整合多种不同原理的算法、建立多维度评估体系及应用智能筛选机制，旨在系统性地识别并提取高可靠性的因果关系。整体方法论主要包括以下三个核心步骤：

1. 基于多算法的候选因果关系生成

本研究的第一步是通过多种算法并行地对预处理后的医疗数据进行分析，以生成一个全面的候选因果关系池。为了确保方法学的多样性，我们选取了五种基于不同理论基础的因果发现算法：

PC 算法（基于约束）：该方法利用条件独立性检验来逐步剔除图中不存在的边，其理论基础为因果马尔可夫条件和忠实性假设。

爬山算法（基于评分）：作为一种贪心搜索策略，该算法通过迭代地添加、删除或反转边来优化整个网络结构的评分函数，以寻找最优模型。

贪婪等价搜索 (GES, 基于评分): 此算法在马尔可夫等价类的空间中进行搜索，通过前向和后向两个阶段贪心地增删边，能有效避免陷入局部最优解。

树搜索算法: 该算法旨在构建一个层次化的因果关系树状结构，通过识别根节点并按因果传导路径分层构建网络。

专家在循环算法: 此方法将数据驱动的发现过程与领域专家知识相结合，通过迭代反馈来优化和验证因果网络结构。

这五种算法的并行执行将产生一个庞大且存在冗余的初始因果边集合，该集合将作为后续评估与筛选的输入。

2. 多维度评分体系的构建与实施

为了对候选因果关系池中的每一条边进行量化评估，本研究建立了一个多维度的综合评分体系。该体系从以下四个正交的角度评估每条边的可靠性，并为每个维度分配了预设的权重：

频次评分 (Frequency Score, 权重 30%): 衡量一条因果边被多少种算法共同支持。其计算方式为支持该边的算法数量与总算法数量之比，用以评估该关系的共识度。

多样性评分 (Diversity Score, 权重 25%): 评估支持某条边的算法类型的多样性（如是否同时被基于约束和基于评分的算法所支持）。此评分旨在避免由某单一类型算法的系统性偏见所导致的高频次假象。

算法一致性评分 (Consistency Score, 权重 25%): 该评分引入了算法权重概念，允许为不同算法（可基于其理论完备性或过往表现）分配不同信誉度，并据此计算加权后的一致性得分。

网络拓扑评分 (Topological Score, 权重 20%): 基于图论，此评分通过综合评估一条边的两个端点在网络中的结构重要性来为边赋分。评估指标包括度中心性、介数中心性、接近中心性、PageRank 中心性和特征向量中心性。

每条候选因果边都将通过上述体系获得一个最终的综合集成评分，该分值反映了其整体的可靠性与重要性。

3. 基于数据驱动的智能筛选机制

获得所有候选边的综合评分后，最后一步是设计并应用一个智能筛选机制，以客观地区分高质量因果关系与低质量或虚假的关联。

筛选阈值的确定：为避免主观设定阈值，本研究将采用多种数据驱动技术来确定最佳分割点。这些技术包括肘部法则（Elbow Method），用于寻找评分分布曲线的“拐点”；聚类分析（如 K-means），将所有边根据评分聚类为“高质量”和“低质量”两组；以及异常检测算法，用于识别评分显著偏低的异常边。最终的筛选阈值将综合上述方法的结果来确定。

筛选与分级：所有综合评分高于该确定阈值的因果边将被保留，构成最终的高质量因果关系网络。低于该阈值的边则被视为证据不足而被过滤。此外，为便于结果的解释和使用，保留下来的边还将根据其具体分值被进一步划分为不同的质量等级（如白金级、黄金级等）。

2.3 参数学习与模型拟合度量化

在确定候选因果结构后，本研究的核心任务转向对因果关系的强度进行量化，并评估整个网络模型对观测数据的解释能力。为实现这一目标，我们构建了一个全面的参数学习与评估框架，该框架不仅旨在精确估计参数，还致力于评估这些估计的稳定性和各条因果边对模型的具体贡献。

为此，我们采用了四种具有互补优势的参数学习方法，以确保参数估计的鲁棒性和全面性：

最大似然估计 (Maximum Likelihood Estimation, MLE)：作为一种经典的频率派方法，MLE 旨在找到使观测数据出现概率最大化的参数值。它因其计算效率高和在大样本下具有良好的渐近特性而被广泛应用。

贝叶斯估计 (Bayesian Estimation)：该方法将参数视为随机变量，通过结合数据的似然和参数的先验分布来计算其后验分布。这种方法能够自然地量化参数的不确定性，并且在样本量较小时，通过引入先验知识能有效避免过拟合。

期望最大化 (Expectation-Maximization, EM) 算法：EM 算法是一种在数据缺失或存在未观测潜变量情况下的参数估计利器。它通过迭代执行“期望 (E) 步”（估计缺失数据）和“最大化 (M) 步”（基于完整数据重新估计参数），直至收敛，从而得到参数的局部最优解。

结构方程模型 (Structural Equation Models, SEM)：SEM 是一个强大的框架，用于检验关于观测变量和潜变量之间预设的因果关系假设。它通过分析变量间的协方差结构来估计参数，并能提供整个模型的拟合优度指数，从而可以直接量化因果效应。

对于每种方法，我们都生成了详尽的条件概率表 (Conditional Probability Tables, CPTs)，用以精确描述网络中每个节点状态如何依赖于其父节点状态的概率分布。此外，为

了对参数估计的稳定性进行量化评估，我们引入了 Bootstrap 重采样技术。该技术通过从原始数据集中有放回地多次抽样，构建多个重采样数据集，并在每个数据集上重新进行参数估计，最终形成参数的经验分布，从而评估其稳定性和置信区间。

为客观评估因果网络参数学习的质量与可靠性，尤其是在采用多方法学框架时，必须建立一套标准化的量化评估体系。本研究提出一个双维度评估框架，该框架不依赖于单一的拟合指标，而是从“边际贡献度”和“跨方法一致性”两个正交的视角出发，深入剖析网络中每一条因果关系的强度与可信度。该体系主要由边际似然增益分析和参数稳定性评估两大核心方法构成。

1. 边际似然增益 (Marginal Likelihood Gain)

传统的模型评估方法侧重于网络的整体拟合优度，但这无法揭示特定因果关系对模型解释能力的具体贡献。为了实现更细粒度的分析，本研究引入边际似然增益 (Marginal Likelihood Gain) 分析，其核心目标是独立地量化每一条因果“边”对整个模型数据拟合能力的提升程度。定义与计算方法：

$$Gain_{edge} = \log L(\text{模型含该边}) - \log L(\text{模型除该边})$$

其中， $L(\cdot)$ 代表模型在给定观测数据下的似然函数值。计算过程需要对网络中的每一条边进行迭代：首先计算包含所有边的完整模型的总对数似然度，然后依次暂时移除每一条边，并重新计算模型的对数似然度，最后求取二者之差。

该方法将模型的整体性能分解到每一个独立的结构组件（即“边”）上。一个较高的正向增益值表明，该因果关系是解释数据变异性的关键因素，其存在显著提升了模型的拟合能力。反之，一个接近于零或为负的增益值则暗示该边可能是冗余的或错误的，对模型的贡献微乎其微甚至产生负面影响。通过这种方式，研究者能够识别出网络中的核心因果路径，并对结构学习阶段发现的边进行量化筛选。

2. 参数稳定性 (Parameter Stability)

任何单一的参数估计算法都受其底层统计假设的限制，可能产生系统性偏差。为了克服这一局限性并确保结论的鲁棒性，本研究设计了参数稳定性 (Parameter Stability) 评估方法。该方法的核心思想是：如果一条因果关系是真实且稳固的，那么使用不同理论基础的参数估计算法（如最大似然估计、贝叶斯估计等）对其进行量化时，得到的结果应该具有高度的一致性。

参数稳定性通过衡量不同方法对同一条边所估计参数值的离散程度来定义。具体而言，我们采用变异系数 (Coefficient of Variation, CV) 作为核心度量，因为它是一个无量纲的相对差异指标。为了使评估结果更具可解释性，我们将其转换为一个标准化的稳定性分数，其计算公式如下：

$$Stability = 1 - \frac{CV}{CV_{\max}}$$

其中， CV 是基于多种方法（例如，MLE、Bayesian、EM、SEM）所得参数估计值计算出的变异系数。 CV_{\max} 是一个归一化因子，代表在特定数据类型和分布下理论上可能出现的最大变异系数，以确保稳定性分数落在的区间内。

参数稳定性提供了一个超越单一算法性能的元评估 (meta-evaluation) 视角。一个接近 1 的稳定性分数意味着该因果关系的量化结果在不同统计学派的模型中得到了交叉验证，其结论非常可靠。相反，一个较低的分数则警示研究者，该关系的量化结果存在显著的方法依赖性，其背后可能隐藏着数据稀疏、模型假设不匹配或存在未被观察到的混杂因素等问题。这种评估机制强制研究者审视那些看似存在但并不稳固的因果关系，从而有效提升了最终网络模型的可信度。

2.4 基于贝叶斯推断的因果机制解析

在因果推断的研究中，仅仅确定变量间的因果关系是不够的，更深层次的科学问题在于揭示这种关系背后的作用机制。因果中介分析 (Causal Mediation Analysis) 为此提供了一个强大的理论框架，其核心目标是将一个变量 (X) 对另一个变量 (Y) 的总效应 (Total Effect) 分解为两个核心部分：

直接效应 (Direct Effect, DE)：变量 X 在不通过中介变量 (M) 的情况下，对 Y 产生的直接影响。

间接效应 (Indirect Effect, IE)：变量 X 通过影响中介变量 M ，再由 M 对 Y 产生影响的传导路径。这部分效应揭示了因果链条中的“黑箱”，是理解机制的关键。

本研究采用贝叶斯推断框架进行中介分析，旨在利用其在不确定性量化、小样本处理和融合先验知识方面的独特优势，系统性地发现并量化复杂的因果传导路径。本研究的中介分析建立在线性结构方程模型的基础之上。对于一个经典的三变量中介模型 ($X \rightarrow M \rightarrow Y$)，其数学表达如下：

$$M = \alpha + \alpha \cdot X + \varepsilon_M$$

$$Y = \beta + c' \cdot X + b \cdot M + \varepsilon_Y$$

模型参数定义：

路径 a: 自变量 X 对中介变量 M 的效应系数。

路径 b: 在控制了 X 的影响后，中介变量 M 对因变量 Y 的效应系数。

路径 c': 在控制了 M 的影响后，X 对 Y 的直接效应系数。

效应分解：

间接效应 (IE): 通过路径 a 和 b 的乘积来量化，即 $IE = a \times b$ 。它代表了通过中介变量 M 传导的因果效应的大小和方向。

总效应 (TE): 直接效应与间接效应之和，即 $TE = c' + (a \times b)$ 。它反映了 X 对 Y 的总体影响。

选择贝叶斯方法进行中介分析，主要基于以下几点方法学优势：

完整的不确定性量化: 贝叶斯推断不产生单一的点估计值，而是为每个参数 (a, b, c') 以及由它们计算出的间接效应 (IE) 生成一个完整的后验概率分布。这使得我们能够全面地评估参数估计的不确定性。

直观的显著性检验: 我们采用**最高密度区间 (Highest Density Interval, HDI) **来判断效应的显著性。通常，我们会计算 95% HDI。如果间接效应 IE 的 95% HDI 不包含零点，我们则认为该中介效应是显著的。这种方式比传统的 p 值更易于解释且在统计上更为稳健。

先验知识的融合: 贝叶斯框架允许通过设定先验分布 (Prior Distribution)，将领域专家的知识或以往的研究结论整合到模型中。这在处理样本量有限的医疗数据时尤其有价值，有助于获得更稳定和合理的估计结果。

为确保分析的系统性和可复现性，本研究构建了以下标准化的分析流程：

1. **中介路径的发现与筛选:** 基于已构建的因果网络，首先采用算法系统性地枚举所有可能的 $X \rightarrow M \rightarrow Y$ 形式的三元组路径。随后，邀请领域专家根据医学知识对这些潜在路径进行筛选，剔除不具备生物学或临床合理性的路径。
2. **模型构建与后验推理:** 为每一条通过筛选的路径构建贝叶斯中介模型。通过设定无信息或弱信息的先验分布来减少主观偏见，然后利用马尔可夫链蒙特卡洛 (MCMC) 算法（如 Gibbs 采样）对模型参数的后验分布进行抽样，从而完成贝叶斯推理过程。

3. 效应量化与中介比例计算：基于 MCMC 采样得到的后验分布，计算直接效应（DE）、间接效应（IE）和总效应（TE）的后验均值、标准差和 95% HDI。同时，计算中介比例（Proportion Mediated, PM），其定义为 $PM = |IE| / |TE|$ ，用于评估间接效应在总效应中所占的比重。

2.5 四维证据评估体系

在复杂的生物医学系统中，任何单一的因果发现算法或统计模型都存在固有的假设和局限性。为了克服这一挑战，本研究构建了一个三角测量（Evidence Triangulation）框架。其核心原则是：一个真正稳健的因果关系，应当能够经受住来自多个不同维度、采用不同方法的证据的交叉验证。

本框架通过系统性地整合来自因果发现流程中不同阶段的证据，旨在从众多候选因果关系中，筛选出那些最可靠、最具解释力的核心因果边。

1. 四维证据评估体系

我们为每条候选边构建了四个维度的评分，分别代表了不同层面的证据支持强度：

结构一致性评分（S_struct(e)）：此维度评估不同因果发现算法对该条边存在的“共识”程度。它整合了三个子指标：该边被算法识别的频次、识别算法的多样性（如是否同时被基于约束和基于评分的方法支持）以及不同算法对其方向判断的一致性。该评分反映了因果结构在算法层面的稳健性。

参数拟合评分（S_param(e)）：此维度从统计学角度评估因果关系模型的拟合优度。我们采用多种参数学习方法（如最大似然估计、贝叶斯估计、EM 算法等）对包含该边的局部结构进行拟合，并基于似然增益等指标来量化该边的加入对模型解释力的提升程度。该评分反映了因果关系在数据层面的统计支持。

中介支持评分（S_med(e)）：此维度从机制角度评估该边在更长的因果链条中的作用。通过贝叶斯中介分析，我们考察该边作为中介路径 ($X \rightarrow e \rightarrow Y$) 一部分时的间接效应强度和显著性概率。一条边如果在多条中介路径中均表现出显著的效应，则其作为因果机制关键环节的可能性更高。评分还会根据节点类型（如“疾病→药物”）进行权重调整，以符合临床逻辑。

专家定向评分（S_expert(e)）：此维度整合了领域知识的先验信息。它一方面来源于

支持人机交互的专家算法的直接输出，另一方面来源于根据医学常识预设的领域知识库（例如，“疾病→药物”的因果方向通常比反向更合理）。该评分确保了筛选出的因果关系符合公认的医学理论。

2. 联合置信度计算与质量控制

为了得到一个统一的评估标准，我们将上述四维评分通过加权求和的方式融合成一个联合置信度函数

$$\text{Conf}(e) = \alpha \cdot S_{\text{struct}}(e) + \beta \cdot S_{\text{param}}(e) + \gamma \cdot S_{\text{med}}(e) + \delta \cdot S_{\text{expert}}(e)$$

其中，权重系数 ($\alpha, \beta, \gamma, \delta$) 经过设定以反映各维度证据的重要性（例如，结构发现作为基础，其权重 α 被赋予较高值）。

同时，为了控制评估过程中的数据质量，我们还设计了一个并行的数据质量评分，用于评估用于计算参数拟合、中介支持等维度评分的底层数据的完整性和可靠性。

3. 核心因果边的双重筛选机制

最后，我们采用一套严格的双重筛选标准来识别核心因果边：

置信度阈值筛选：候选边的联合置信度 $\text{Conf}(e)$ 必须达到预设的阈值 (≥ 0.6)。数

据质量阈值筛选：支持该边评估的数据质量评分也必须达到相应的标准 (≥ 0.6)。

只有同时满足这两个条件的候选边，才被最终认定为“核心因果边”。这一过程确保了所筛选出的每一条因果关系不仅在多源证据上具有高度一致性，而且其评估过程本身也基于高质量、可靠的数据。通过这种方式，我们为后续的因果网络构建和临床应用提供了坚实的、经过多重验证的因果关系基础。

2.6 面向临床应用的因果知识图谱构建

本研究旨在构建一个超越传统关联网络的增强型因果知识图谱（Enhanced Causal Knowledge Graph）。其核心理念是：一条高质量的因果知识不应依赖于单一来源的证据，而必须通过一个从发现、筛选、验证到属性增强的系统性流程进行构建。为此，我们设计了一个多层次证据融合的框架，确保图谱中的每一条因果关系都具备结构稳定性、统计显著性、机制可解释性和领域知识一致性。为实现对复杂因果知识的完整表示，我们将医疗因果知识图谱定义为一个六元组 KG：

$$KG=(V, E_{core}, R, W, \Theta, \Phi)$$

该模型不仅包含节点（V）、核心因果边（E_{core}）和关系类型（R），还创新性地整合了边级权重属性（W）、参数属性（Θ）以及效应属性（Φ），实现了因果知识及其证据链的完整表征。

V (Vertices): 节点的集合，代表医疗实体，如疾病、药物、检验指标。

E_{core} (Core Edges): 核心因果边的集合，代表经过多阶段验证的高质量因果关系。

R (Relations): 关系类型的集合，对边的语义进行分类。

W (Weights): 权重属性集合，量化每条边的可信度、强度和质量等级。

Θ (Parameters): 参数属性集合，存储描述因果关系的统计参数，如条件概率表（CPT）和模型拟合指标。

Φ (Effects): 效应属性集合，量化因果效应的细节，如直接效应、间接效应和中介类型。