

TAVI/R知识库问答效果评估

1. 项目简介

本项目旨在评估在TAVI/R医学领域中，引入专业知识库后，问答系统（RAGFlow + DeepSeek V3）相较于通用大语言模型（DeepSeek V3直接回答）在解答专业问题时的性能提升。

2. 评估设计

• 问题集与基准:

基于100余篇TAVI/R领域指南和共识意见，首先人工筛选核心文档，后由LLM辅助从中生成20个临床问题（见附录）。每个问题均配有从源文献提取的“原文片段”，此“原文片段”即为后续评估的基准。

• 核心评估指标与评分:

采用5分制，计划使用Gemini 2.5 Pro结合各问题的“原文片段”对以下主要指标进行评分：

- 答案相关性（与提问的匹配度）
- 内容一致性（与“原文片段”的符合度）
- 文献支撑可靠性（答案信息的准确性及依据可靠性）

• 对比分析方法:

除量化评分外，本报告还将辅以人工定性对比。选取典型案例，具体分析知识库增强系统与基础模型在回答内容、准确性及可靠性上的差异。

3. 核心结果

• 3.1 量化数据对比

为量化评估两个AI问答系统——“知识库系统 (KB System (RAG))”与“DeepSeek基础模型 (DeepSeek (Baseline))”——在处理TAVI/R领域专业问题时的性能，我们采用了统一的评估框架。针对预设的20个临床问题（已在附录中列出），我们主要考察了以下三个核心维度，所有评分均基于5分制。图表中的标签均使用英文以便清晰对应。

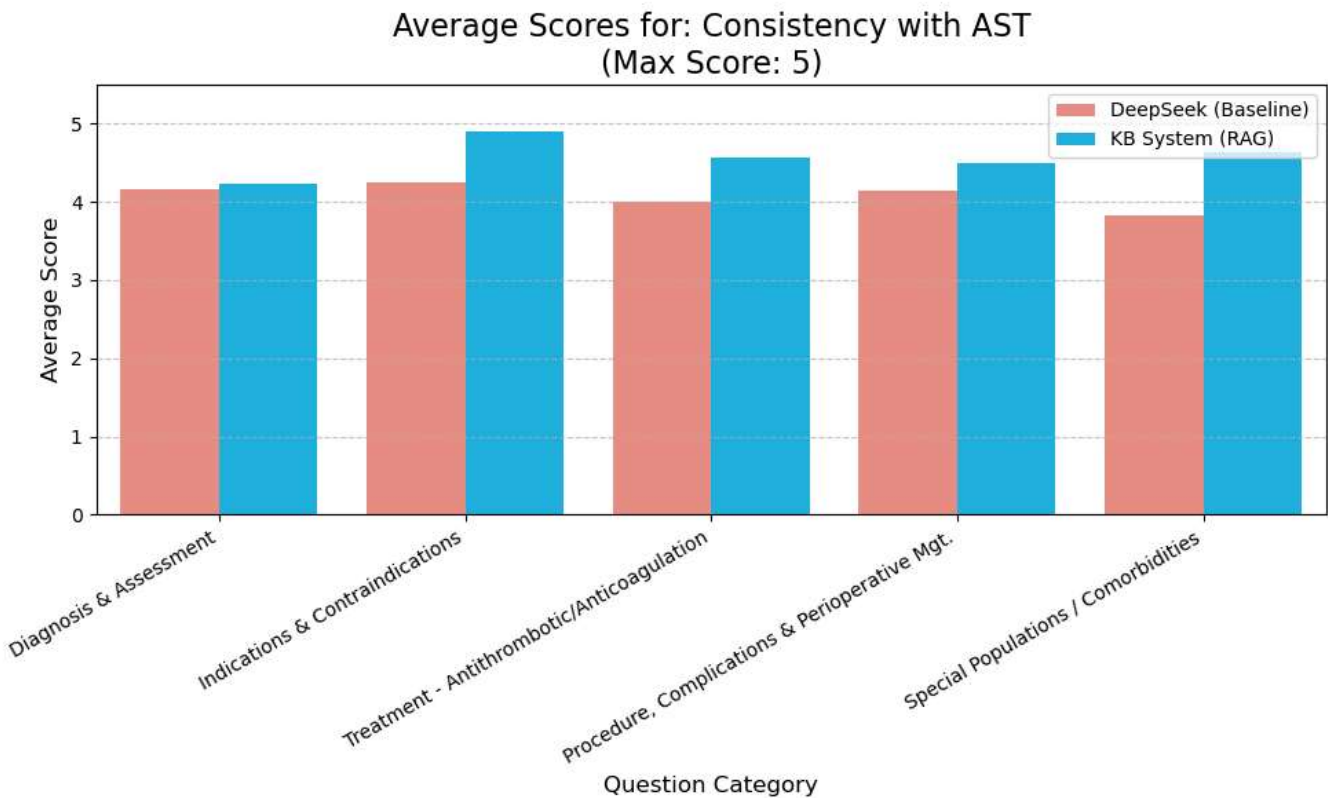
核心评估指标 (Key Evaluation Metrics):

- 与原文片段的一致性 (Consistency with AST):** 衡量系统回答内容与预设“原文片段”的符合程度。
- 答案相关性 (Answer Relevance):** 评估系统回答与所提问题的匹配程度。
- 文献支撑可靠性 (Literature Support Reliability):** 考量答案信息的准确性、是否有臆断以及依据的可靠性（对于知识库系统，尤其体现为对其检索上下文的忠实度；对于基础模型，则主要参照其信息与“原文片段”所代表的事实的一致性）。

接下来，我们将分别展示各个指标下的具体得分对比结果。

1. 与原文片段的一致性 (Consistency with AST) 得分对比

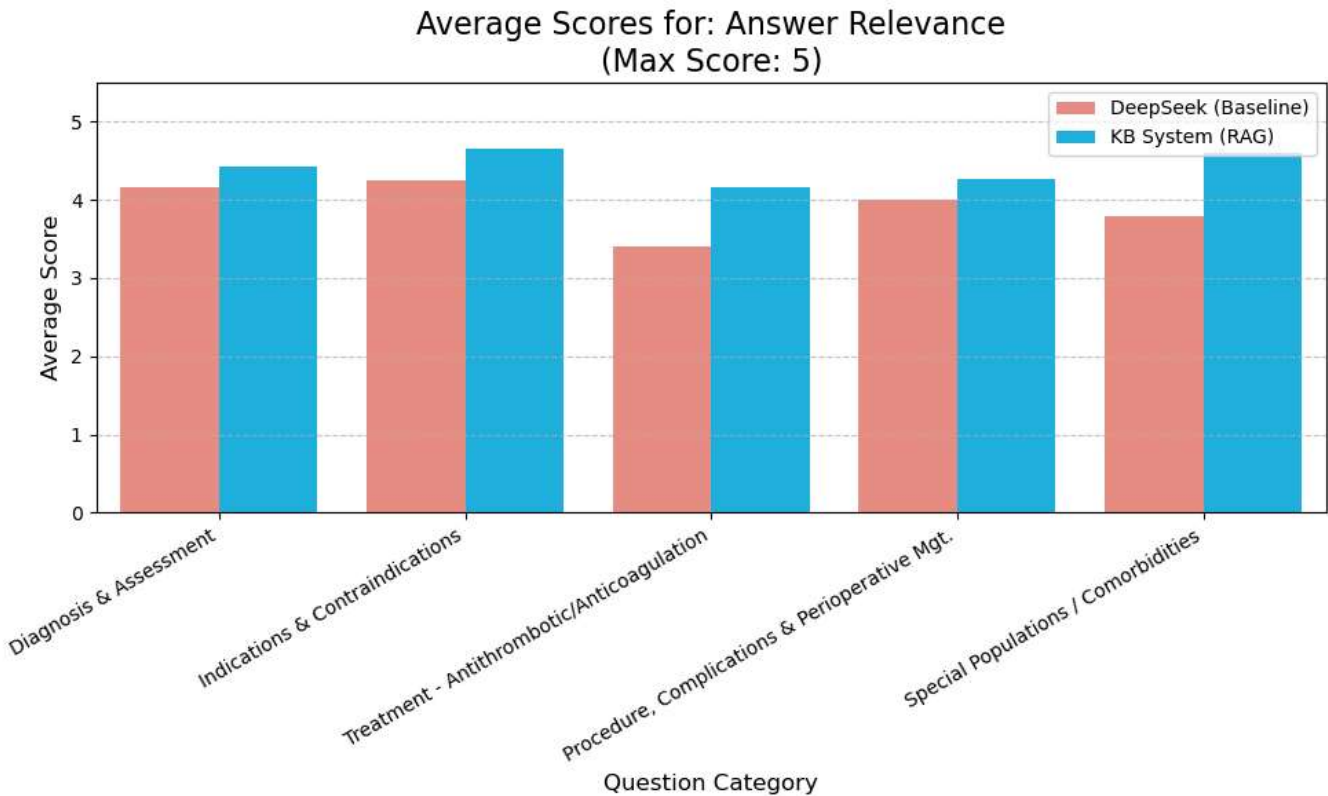
下图展示了两个系统在“与原文片段的一致性” (Consistency with AST) 指标上，于五个不同问题类别的平均得分对比情况。



简要分析: 结果显示，知识库系统（KB System (RAG)）在所有问题类别的内容一致性平均得分上均优于DeepSeek基础模型，尤其在“适应证与禁忌证 (Indications & Contraindications)”及“特殊人群/共病管理 (Special Populations / Comorbidities)”类别中展现出更明显的优势。

2. 答案相关性 (Answer Relevance) 得分对比

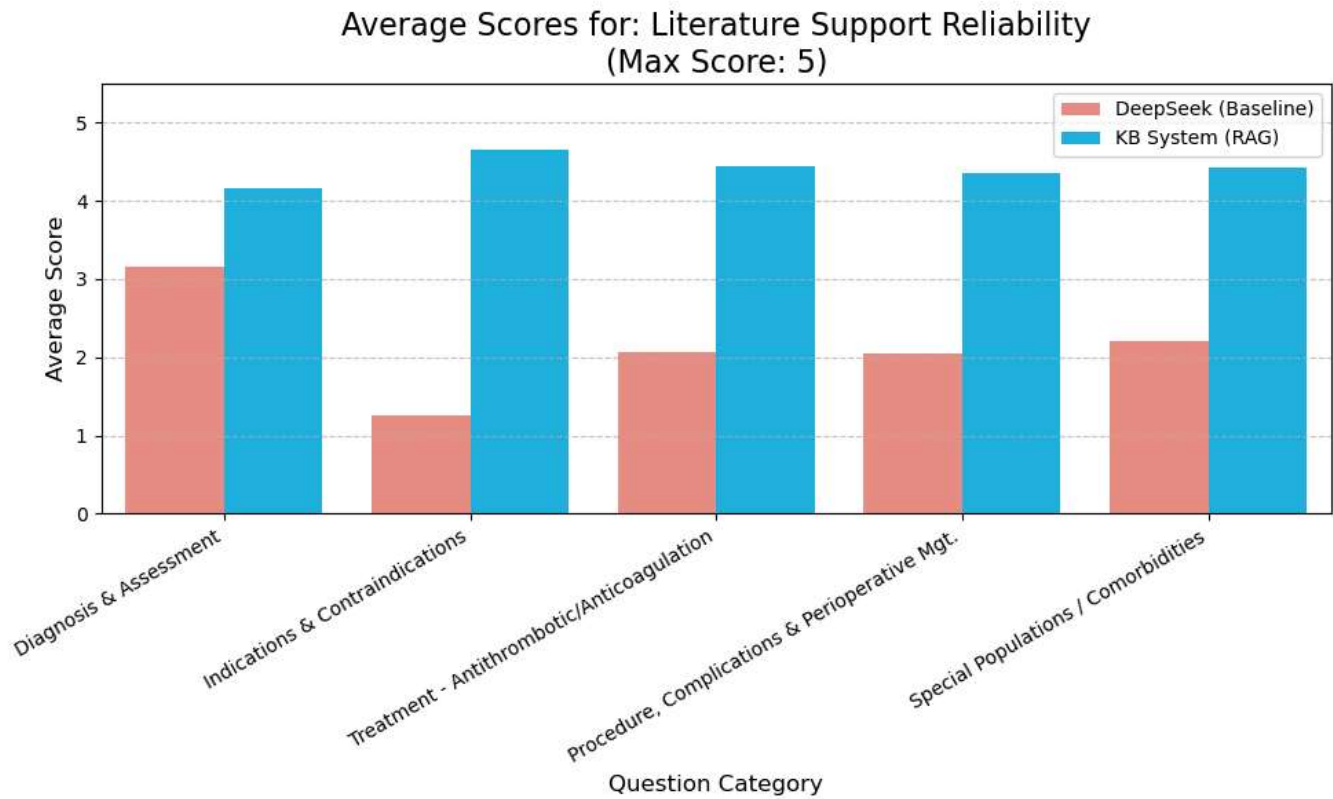
下图比较了两个系统在“答案相关性” (Answer Relevance) 指标上的平均得分表现。



简要分析: 在答案相关性方面，知识库系统（KB System (RAG)）在各类别中的平均得分普遍高于或持平于DeepSeek基础模型，表明其生成的答案通常能更贴合问题主旨。

3. 文献支撑可靠性 (Literature Support Reliability) 得分对比

下图展示了在“文献支撑可靠性” (Literature Support Reliability) 指标上，两个系统的平均得分情况。



简要分析: 对于文献支撑可靠性，知识库系统（KB System (RAG)）在所有问题类别上的平均得分均显著大幅领先DeepSeek基础模型，这直观反映了其在提供有据可循、信息可靠答案方面的核心能力。

量化结果总结 (Summary of Quantitative Results):

综合以上三个核心指标的平均得分数据，知识库增强系统（KB System (RAG)）在所有评估的问题类别和各项核心评估维度上，均一致地展现出优于DeepSeek基础模型的性能。特别是在“与原文片段的一致性”和至关重要的“文献支撑可靠性”方面，知识库系统的优势尤为突出。这些量化结果初步表明，引入专业知识库能有效提升大语言模型在TAVI/R医学领域专业问答的准确性、相关性与整体可靠性。

• **3.2 知识库的独特功能**

本知识库系统在提供精准问答的基础上，集成了以下特色功能，旨在提升用户获取、理解及验证专业医学信息的效率和可靠性：

i. 原文文献链接与直接访问

系统在生成答案时，能够清晰展示其内容所依据的核心原始文献列表（如下图所示）。用户可以通过这些链接直接访问和下载完整的PDF格式指南、专家共识或研究报告，方便进行深入学习或获取更全面的上下文信息。

四、随访与长期管理

术后1/3/6个月评估：超声心动图检测左室功能，优化药物（如停用β受体阻滞剂）。
2024中国专家共识强调：所有TAVR患者出院前应进行24小时Holter筛查迟发型传导障碍。
注：最终决策需个体化权衡血栓/出血风险，尤其对高龄衰弱患者可放宽PPM指征。

- 心尖途径经导管主动脉瓣置换术围术期管理规范专家建议.pdf
- 单纯主动脉瓣反流经股动脉主动脉瓣置换中国专家共识2023.pdf
- 经导管主动脉瓣置换术中国专家共识（2020更新版）.pdf
- 中国经导管主动脉瓣置换术临床路径专家共识（2021版）.pdf
- 经导管主动脉瓣置换术临床实践指南.pdf

ii. 答案来源片段可视化追溯

为增强答案的可信度并方便快速核对，本系统提供了强大的原文追溯功能。在展示答案的同时，系统能够通过旁置的文档预览器或页面缩略图（如下图所示），高亮或直接定位到答案中关键信息点在原始文献内的具体章节、段落或图表。这一功能使用户可以迅速验证信息的准确性，并理解其在原始文献中的具体语境。

知识库 聊天 搜索

在评估预备行TAVR的主动脉瓣置换术前，需对主动脉根部及瓣环解剖进行评估。需通过CT精确测量瓣环最大径、主动脉窦直径、左室流出道、二尖瓣反流程度、左室收缩功能和舒张功能、跨瓣血流动力学与左室功能等。除常规跨瓣压差和有效瓣口面积外，还需关注冠状动脉开口高度、左室流出道有无室间隔缺损、二尖瓣反流程度、有无二尖瓣狭窄、有无二尖瓣前瓣钙化、左室大小和功能、三尖瓣形态和功能、反流程度、估测肺动脉收缩压、有无心包积液、积液部位及积液量。

Document Previewer

重度 AS 患者（COR I，LOE A），但 TAVR 的适应症需综合评估。

中华超声影像学杂志 2023 年 10 月第 32 卷第 10 期 Chin J Ultrasonogr.

表 1 主动脉瓣狭窄的 TAVR 术前超声心动图检查要点

检查项目	检查内容
主动脉瓣和主动脉根部	二叶式(分型)或三叶式瓣叶钙化程度及分布
跨主动脉瓣血流动力学	跨主动脉瓣压差和有效瓣口面积
主动脉瓣环径	最大径和最小径、周长、面积
左室流出道	内径、有无室间隔缺损明显肥厚
主动脉根部径线和钙化评估	主动脉窦直径和高度、窦管交界直径和钙化程度、冠状动脉开口位置 and 高度, 预估堵塞风险
二尖瓣结构和功能	二尖瓣反流程度、有无二尖瓣狭窄、有无二尖瓣前瓣钙化
左室结构和功能	左室壁运动评估、排除心腔血栓、左室质量、左室壁厚情况、收缩功能和舒张功能
右心结构和功能	右室大小和功能、三尖瓣形态和功能、反流程度、估测肺动脉收缩压
其他	有无心包积液、积液部位及积液量

注：TAVR 为经导管主动脉瓣置换术

file:///C:/Users/27134/AppData/Local/Temp/crossnote202555-18608-13ty0mg.ozwy.html

4/10

二、永久起搏器植入的指征与处理建议

(1) 明确植入指征

术后出现以下任一情况需永久起搏器（PPM）：①



通常情况下，患者对慢性AR的耐受性较好，出现症状时往往心脏储备功能很差，常有左心室扩大心功能下降，患者左心收缩功能的恢复往往较慢，部分患者术后左心室重构无法逆转，建议术前加强评估及与患者及家属充分告知 [52-54]。建议术后加强指南推荐的标准药物特别是可改善预后的药物。存在心脏房室传导阻滞需要植入起搏器者，推荐希氏束起搏或左束支起搏等生理性起搏方式 [55-56]。对于有左束支传导阻滞者，若心功能持续不能恢复，可以考虑行心脏再同步化治疗 [57]。建议术后1个月（必要时3个月）、6个月、12个月之后每年对患者进行随访，随访内容包括患者症状、不良事件、生活质量评分及用药情况等，随访的检查包括超声心动图、心电图及相关的生化检查。



单纯主动脉瓣反流经股动脉主动脉瓣置换中国专家共识2023.pdf

通过上述功能，本知识库系统不仅致力于提供准确的医学问题解答，更强调信息来源的透明化与可验证性，有效支持了专业、高效的临床决策和研究工作。

• 3.3 代表性案例

案例：[TAVI术后无OAC适应证患者的常规抗血栓方案]

1. 问题原文 (Question):

TAVI手术后，对于不存在其他明确的长期口服抗凝药物（OAC）治疗适应证的患者，常规的抗血栓治疗方案为何？

2. 系统回答对比:

A. Chat系统回答 (有知识库辅助 - RAGFlow + DeepSeek V3):

(以下为Chat系统结合知识库给出的回答)

对于 **TAVI术后无长期口服抗凝药（OAC）适应证** 的患者，常规抗血栓治疗方案基于最新国际指南和专家共识总结如下：

1. 标准抗血小板方案

- **低血栓风险患者：**
 - **单药抗血小板治疗（SAPT）：**阿司匹林（75~100 mg/天）终身维持，推荐等级 **I类（A级证据）**。
- **较高血栓风险患者（如合并冠心病、既往PCI史等）：**
 - ****双联抗血小板治疗（DAPT）3₆个月**：**阿司匹林（75~100 mg/天）+氯吡格雷（75 mg/天），随后转为阿司匹林单药终身维持。

2. 特殊人群调整

- 亚临床瓣叶血栓或高栓塞风险者：
 - 部分中心建议延长DAPT至6个月，或短期华法林抗凝（INR 2.5，3个月），但需个体化评估出血风险。

3. TAVI联合近期PCI患者的策略

- PCI后1年内需DAPT者：
 - DAPT方案同冠心病管理（通常阿司匹林+氯吡格雷6~12个月），后续转为阿司匹林单药。

4. NOACs的争议与限制

- 避免NOACs联合抗血小板药物：
利伐沙班等NOACs在无OAC适应证的患者中可能增加出血风险，目前指南明确反对常规使用。

核心推荐对比表

患者特征	初始方案 (0~6个月)	≥6个月后长期方案	证据等级
低血栓风险	SAPT (阿司匹林)	SAPT持续	I类 (A)
高血栓风险/合并PCI	DAPT (阿司匹林+氯吡格雷)	→ SAPT	IIa类 (B-NR)
亚临床瓣叶血栓	DAPT延长或短期华法林	→ SAPT	IIb类 (C-LD)

注：所有决策需结合多学科团队评估及定期影像学随访监测瓣膜功能。

引用来源文献：

- 经导管主动脉瓣植入术后抗血栓治疗中国专家共识.pdf
- 瓣膜性心脏病患者经导管主动脉瓣置换术后抗血栓管理的最佳证据总结.pdf

B. DeepSeek V3 直接回答 (无知识库辅助):

(以下为DeepSeek V3在无知识库辅助下直接给出的回答)

对于经导管主动脉瓣植入术（TAVI）后无其他长期口服抗凝药物（OAC）适应证的患者，目前的抗血栓治疗方案主要基于最新指南和临床研究证据，具体如下：

1. 常规抗血小板治疗（无OAC适应证患者）

- **单药抗血小板治疗（SAPT）：**
 - **推荐方案：**阿司匹林 75-100 mg/天，长期维持。
 - **依据：**多项研究（如 POPULAR TAVI、ARTE 等）表明，单用阿司匹林在降低出血风险方面优于双联抗血小板（DAPT），且不增加血栓事件。
- **双联抗血小板治疗（DAPT）：**
 - **适用情况：**仅限部分高血栓风险患者（如近期冠状动脉支架植入术），但证据有限。
 - **方案：**阿司匹林 + 氯吡格雷（75 mg/天）3-6个月，后改为单药。
 - **注意：**DAPT可能增加出血风险，目前指南多倾向于避免常规使用。

2. 特殊人群调整

- **合并冠心病但无需OAC者：**
 - 若近期（1年内）冠脉事件或支架植入，可短期DAPT（1-3个月），后过渡至SAPT。
 - 稳定性冠心病患者可直接选择SAPT。
- **出血高风险患者：**
 - 优先选择阿司匹林单药，甚至考虑缩短疗程（如3个月后停药）。

3. 指南推荐汇总

- **2021 ACC/AHA/SCAI瓣膜病指南：**
 - 无OAC指征者，首选阿司匹林单药（Class IIa）。
 - DAPT仅用于合并其他适应证（如近期冠脉介入）（Class IIb）。
- **2022 ESC瓣膜病管理指南：**
 - 明确推荐阿司匹林单药（Class I），反对常规DAPT（Class III）。

4. 抗凝治疗的例外

- **需OAC的情况：**房颤、静脉血栓栓塞、机械瓣等患者，应使用OAC（如华法林或DOACs）而非抗血小板治疗。
- **TAVI术后血栓高风险：**目前不推荐无OAC适应证者常规加用抗凝药（如ENVISAGE-TAVI AF试验提示出血风险增加）。

总结

对于无OAC适应证的TAVI患者，**长期单用阿司匹林**是标准方案，DAPT仅限特定情况。临床决策需个体化评估血栓/出血风险，并参考最新指南动态调整。

简要分析与对比小结:

两系统（知识库增强的RAGFlow与DeepSeek V3直接问答）均能就专业问题提供较为详尽的回复。核心差异在于，知识库增强系统（RAGFlow）的回答能够提供源自特定文献的直接佐证（如原文片段或明确的文献引用），因此其答案在专业领域的可信度与可靠性上显著更高。相较之下，DeepSeek V3的直接回答虽然内容可能同样丰富，但因缺乏此类直接、可验证的文献支撑，其信息的权威性和在特定知识体系内的准确性难以得到同等级别的保障。

4. 结论

本次针对TAVI/R领域知识库问答系统的评估，旨在对比基于RAGFlow和DeepSeek V3的知识库增强系统（以下简称“知识库系统”）与DeepSeek V3基础模型直接问答的性能表现。

通过对20个覆盖不同专业领域的临床问题进行测试，并结合“与原文片段的一致性”、“答案相关性”和“文献支撑可靠性”三个核心量化指标（各项满分5分）进行评估，主要结论如下：

- 知识库系统性能更优：**量化数据显示，在所有三个核心评估指标上，“知识库系统”的平均得分均一致性地高于或持平于“DeepSeek V3基础模型”。尤其在“文献支撑可靠性”和“与原文片段的一致性”方面，知识库系统的优势表现得尤为突出，这表明其在提供准确且有据可查的专业答案方面能力更强。
- 特色功能提升实用性：**知识库系统所具备的原文链接及片段定位等独特功能（如第3.2节所述），进一步增强了答案的透明度和可信度，为用户深入查证提供了便利。

综上所述，本次评估初步验证了在TAVI/R这一高度专业的医学领域中，通过构建和应用领域知识库来增强大语言模型，是提升问答系统答案质量、准确性和可靠性的有效途径。

附录

TAVI/R领域评估问题集 (共20题)

一、诊断与评估 (Diagnosis & Assessment)

- 在对拟行TAVR的主动脉瓣狭窄患者进行术前评估时，除常规超声心动图参数外，还有哪些特定的诊断性检查项目或需特别关注的评估指标？

2. 在TAVR手术的术前规划阶段，多排计算机断层扫描（MSCT）检查能够为临床决策提供哪些方面的关键信息？
3. 在评估TAVR手术患者风险时，除了传统的STS评分和EuroSCORE II，是否有新增或特别推荐的风险评估模型系统？

二、适应证与禁忌证 (Indications & Contraindications)

4. 对于无临床症状但已诊断为重度主动脉瓣狭窄，且左心室射血分数（LVEF）维持在55%以上的患者，在何种具体情况下应考虑启动介入治疗？
5. 根据相关共识，哪些患者群体符合TAVR手术的绝对适应证标准？这些标准是否均需同时满足？

三、治疗策略 - 抗栓/抗凝 (Treatment Strategy - Antithrombotic/Anticoagulation)

6. TAVR手术完成后，若患者无心房颤动等明确需要长期抗凝治疗的基础疾病，其术后抗血栓治疗方案通常如何规划？
7. 就主动脉瓣生物瓣置换术后而言，对于经评估出血风险不高的患者群体，其常规的术后抗凝药物治疗策略是怎样的？
8. 接受二尖瓣机械瓣膜置换术的患者，其长期华法林抗凝治疗的国际标准化比值（INR）目标范围应维持在何水平？
9. 对于已接受机械瓣膜置换并长期服用华法林的患者，若需进行拔牙等小规模口腔手术，其当前的华法林治疗方案是否需要作出调整？
10. TAVI手术后，对于不存在其他明确的长期口服抗凝药物（OAC）治疗适应证的患者，常规的抗血栓治疗方案为何？
11. 若在TAVI术后监测中发现瓣膜血栓形成，应如何进行治疗管理？对于临床症状明显的血栓与亚临床状态的瓣叶血栓，处理原则上是否存在差异？
12. TAVI手术后，哪些患者特征或相关因素预示其发生缺血性事件的风险较高？相对地，哪些因素与出血并发症的风险增加相关？
13. 接受长期华法林抗凝治疗的机械瓣膜术后患者，若需进行有创操作（非紧急），其抗凝桥接治疗的一般原则是什么？

四、手术操作、并发症及围术期管理 (Procedure, Complications & Perioperative Management)

14. 在TAVR手术的规划与实施中，若术前评估提示患者存在较高的冠状动脉阻塞风险，有哪些可供选择的预防性措施或处理预案？
15. 就心脏外科手术而言，是否存在关于在术后特定时间窗内（例如6小时以内）拔除气管插管的推荐性策略？此类早期拔管策略的安全性如何，对患者术后恢复可能带来哪些益处？
16. TAVR术后最常见的心脏传导阻滞并发症有哪些类型？对于发生需要永久起搏器植入的房室传导阻滞，通常的处理建议是什么？
17. 在心脏外科手术前，实施多模式的术前康复（Prehabilitation）的目标是什么？它可能包含哪些方面的内容？

五、特殊人群/共病管理 (Special Populations / Comorbidity Management)

18. TAVR患者在围手术期管理中，临床上常见的、需要重点关注并进行相应处理的合并症主要包括哪些？
19. 对于合并晚期肾病且需要维持性血液透析的患者，TAVR手术是否仍为可行的治疗选项？若可实施，在围手术期造影剂的使用上有无特殊考量？
20. 对于拟行TAVR的老年患者，临床上常使用哪些工具或量表来评估其衰弱状态？