Medical Question Answering for Critical Care Medicine

Methods and Preliminary Results on Retrieval

Feb. 26 2025

Overview

- **Outline**
  - Methods (Dataset, Evaluation, Answer Annotation)
  - Results (Retrieval, Figures/ Tables Template)
- **Progress**

Item	Progress	Note
Medical Questions	Done	
Paper Collection	Done (120 papers)	
Paper Annotation	31/120	Expected: Mar. 7
Paper Draft	Methods	Expected: Feb. 28

Objective

We aim to design and execute an **experiment to demonstrate our proposed methodology** for medical question answering for critical care medicine.

Question: Do we need baselines?

Example Question/ Claim:

Steroids improves survival and reversal of shock **in patients with septic shock**.

Supporting the claim:

Paper 1
Study Design and Methodology:
Study Population:
Interventions:
Comparator:
Outcomes:
Strengths and Weaknesses:
Key Findings and Conclusion:

Paper 2
…

Against the claim:

Paper 3
Study Design and Methodology:
Study Population:
…

Dataset

- **Research Papers**
  - 120 research papers related to any medical conditions are collected for model build up and evaluation. (34 papers from WikiJournal, 86 from PubMed, Google Scholar)
- **Medical Questions**
  - 36 yes-no medical questions (12 for ARDS, 10 for sepsis, 6 for cardiac arrest, 7 for delirium, 1 for Sepsis *and* delirium).

Characteristics for Data Collection Variation

- Current Data Collection Status
  - 4 Critical Care Categories: ARDS, Sepsis, Cardiac Arrest, Delirium
  - Approximately 120 papers collected in total
  - Random Papers Current Comparison: high-impact vs. low-impact (cited < 100 times)
- Consideration
  - Proportion
  - Categories

Characteristics for Data Collection Variation

- Key Data Variation Characteristics
  - Paper Selection-based
    - Journal Impact Factor
    - Publication Year
  - Contents-based
    - Treatment vs. Diagnose
    - Positive vs. Negative Findings

Evaluation

- **Relevance**
  - Relevance of retrieved papers measured by precision and recall (aggregated with all medical questions).
- **Factuality**
  - Stances evaluated using overall accuracy, precision and recall.
  - Synthesis (PICO summary) evaluated human rating (subset), ROUGE
- **Consistency**
  - All experiments will be executed 5 times at temperature 0, average results and standard deviation will be reported.

Retrieval results

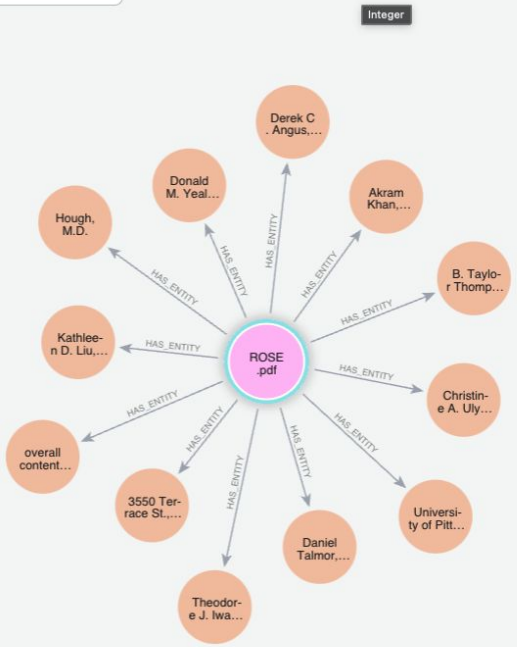
	Question number	Question	Retrieved FileName	Chunk Text	Page Number	Position	Similarity
0	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	LSPA.pdf	clinical trials have shown that early mobilization in both medical and surgical critically ill patients is safe and associated with increased ventilator-free days and improved physical function at hospital discharge.22–25 Early mobilization is limited by use of deep sedation and development of delirium, which can be minimized through the use of scale- based targeted light sedation is implemented early on.26 After reviewing this literature in 2013, the Society of Critical Care Medicine (SCCM)'s	3	20	0.878212
1	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	FMWSCPARDS.pdf	significantly lower cumulative fluid balance by 5,074 mL over 7 days than FACTT Liberal. In subjects without baseline shock, in whom the fluid protocol was applied throughout the duration of the study, management with FACTT Lite resulted in an equivalent cumulative fluid balance to FACTT Conservative. FACTT Lite had similar clinical outcomes of ventilator-free days, ICU-free days, and mortality as FACTT Conservative and significantly greater ventilator-	6	57	0.858627
2	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	FMWSCPARDS.pdf	be easily understood and implemented by physician and nursing staff in the ICU. Conclusions Although the FACTT Lite protocol had a greater cumulative fluid balance than FACTT Conservative, the results of our study indicate that the FACTT Lite protocol is safe and has equivalent ventilator-free days, ICU-free days, acute kidney injury, and adjusted 60-day mortality to FACTT Conservative. FACTT Conservative has improved ventilator-free days,	8	75	0.849876
3	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ARDSSRDRFMS.pdf	subphenotypes have remarkably similar prevalence across the cohorts, and similar natural histories and clinical outcomes. Table 3. Clinical Outcomes by ARDS Subphenotype Subphenotype 1 (n = 727) Subphenotype 2 (n = 273) P Value 60-d mortality, % 21 44 ,0.0001 90-d mortality, % 22 45 ,0.0001 Ventilator-free days, median 19 3 ,0.0001 Definition of abbreviation:	6	93	0.844384
4	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ETSDMV.pdf	[17,18,22,23]. The main side effects of dexmedetomidine are bradycardia, hypotension and the potential for withdrawal symptoms upon discontinuation of long-term therapy [17,18]. When compared to other sedatives, dexmedetomidine has been shown to result in a more awake and interactive patient, a lower incidence of delirium, more ventilator free days, and less days in the ICU [17–	3	26	0.843922
5	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ESCNBC.pdf	undergoing at least 2 days and up to 3 consecutive days of NMBAs (NMBA treatment), within 48 h from commencement of IMV were compared with subjects who did not receive NMBAs or only upon commence- ment of IMV (control). The primary objective in the PS-matched cohort was comparison between groups in 90-day in-hospital mortality, assessed through Cox proportional hazard modeling. Secondary objectives were comparisons in the numbers of ventilator-free days (	1	8	0.843073

Retrieval results

	Question number	Question	Retrieved FileName	Chunk Text	Page Number	Position	Similarity
0	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	LSPA.pdf	clinical trials have shown that early mobilization in both medical and surgical critically ill patients is safe and associated with increased ventilator-free days and improved physical function at hospital discharge.22–25 Early mobilization is limited by use of deep sedation and development of delirium, which can be minimized through the use of scale- based targeted light sedation is implemented early on.26 After reviewing this literature in 2013, the Society of Critical Care Medicine (SCCM)'s	3	20	0.878212
1	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	FMWSCPARDS.pdf	significantly lower cumulative fluid balance by 5,074 mL over 7 days than FACTT Liberal. In subjects without baseline shock, in whom the fluid protocol was applied throughout the duration of the study, management with FACTT Lite resulted in an equivalent cumulative fluid balance to FACTT Conservative. FACTT Lite had similar clinical outcomes of ventilator-free days, ICU-free days, and mortality as FACTT Conservative and significantly greater ventilator-	6	57	0.858627
2	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	FMWSCPARDS.pdf	be easily understood and implemented by physician and nursing staff in the ICU. Conclusions Although the FACTT Lite protocol had a greater cumulative fluid balance than FACTT Conservative, the results of our study indicate that the FACTT Lite protocol is safe and has equivalent ventilator-free days, ICU-free days, acute kidney injury, and adjusted 60-day mortality to FACTT Conservative. FACTT Conservative has improved ventilator-free days,	8	75	0.849876
3	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ARDSSRDRFMS.pdf	subphenotypes have remarkably similar prevalence across the cohorts, and similar natural histories and clinical outcomes. Table 3. Clinical Outcomes by ARDS Subphenotype Subphenotype 1 (n = 727) Subphenotype 2 (n = 273) P Value 60-d mortality, % 21 44 ,0.0001 90-d mortality, % 22 45 ,0.0001 Ventilator-free days, median 19 3 ,0.0001 Definition of abbreviation:	6	93	0.844384
4	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ETSDMV.pdf	[17,18,22,23]. The main side effects of dexmedetomidine are bradycardia, hypotension and the potential for withdrawal symptoms upon discontinuation of long-term therapy [17,18]. When compared to other sedatives, dexmedetomidine has been shown to result in a more awake and interactive patient, a lower incidence of delirium, more ventilator free days, and less days in the ICU [17–	3	26	0.843922
5	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	ESCNBC.pdf	undergoing at least 2 days and up to 3 consecutive days of NMBAs (NMBA treatment), within 48 h from commencement of IMV were compared with subjects who did not receive NMBAs or only upon commence- ment of IMV (control). The primary objective in the PS-matched cohort was comparison between groups in 90-day in-hospital mortality, assessed through Cox proportional hazard modeling. Secondary objectives were comparisons in the numbers of ventilator-free days (	1	8	0.843073

```
neo4j$ MATCH (n:Chunk)-[r]->(m:__Entity__) RETURN n, r, m
```

Graph | Table | RAW



Node details

Chunk

Key	Value
<id>	4:e2698d19-f2ed-4c2d-b3dc-8d517ab5cfe1:13
fileName	"ROSE.pdf"
content_offset	491
page_number	1
length	608
id	"47eaf2e0e1ddcf84e0ce8a8798c83c562bd471ee"
embedding	[-0.05561273545026779, 0.04924878105521202, 0.07040516287088394, 0.07194343209266663, 0.07987554371356964, -0.007318001240491867, -0.05179267376661301... Show all
text	". Hough, M.D., Theodore J. Iwashyna, M.D., Ph.D., Akram Khan, M.D., Kathleen D. Liu, M.D., Ph.D., Daniel Talmor, M.D., M.P.H., B. Taylor Thompson... Show all
position	2

Two hyperparameters

Cypher query that returns the nodes with the highest similarity scores above the **threshold** and **limit** output nodes

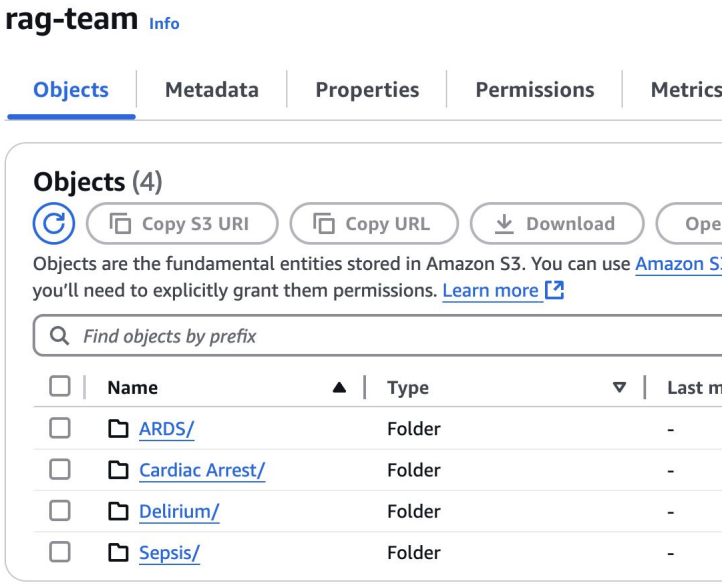
```
def enhanced_chunk_finder(
    graph,
    query: str,
    limit: int = 20,
    similarity_threshold: float = 0.8,
    max_hops: int = 1
     ) -> List[Tuple[str, str, int, int, float]]:
```

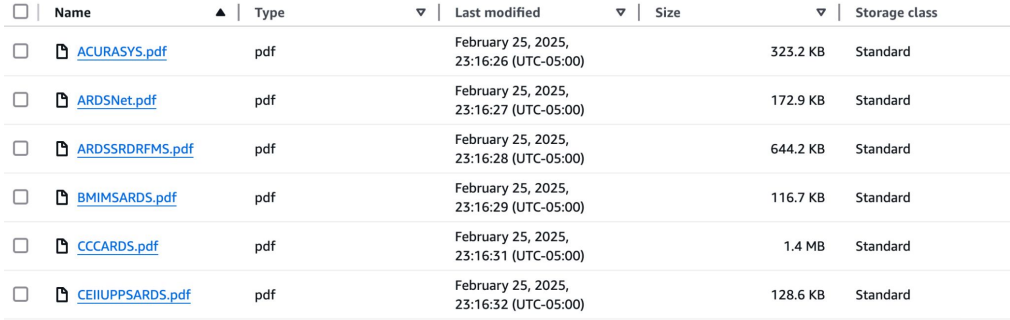
Average similarity score

	Question Number	Question	Retrieved Files	Avg Similarity
0	1	Does early administration of neuromuscular blocking agents increases the ventilator free days?	[ETSDMV.pdf, APROCCHSS.pdf, ESCPARDS.pdf, ESCNBC.pdf, ARDSSRDRFMS.pdf, ACURASYS.pdf, SPICE III.pdf, LSPA.pdf, ARDSNet.pdf, DPSMVAS.pdf, FMWSCPARDS.pdf]	0.845444
1	2	Do patients with severe ARDS being treated with neuromuscular blocking agents have increased muscle weakness?	[CHEST.pdf, ACURASYS.pdf, NBSARDS.pdf, LSPA.pdf, EDPMARDSLPMV.pdf, CEIIUPPSARDS.pdf, ROSE.pdf]	0.869704
2	3	In patients with moderate to severe ARDS, does early use of continuous neuromuscular blockade improve mortality?	[CHEST.pdf, ESCPARDS.pdf, ACURASYS.pdf, ENB.pdf, NBSARDS.pdf, LSPA.pdf, EDPMARDSLPMV.pdf, CEIIUPPSARDS.pdf, ROSE.pdf, TOF-ARDS.pdf]	0.857568
3	4	Do patients with moderate-to-severe ARDS have a significance difference in mortality rate beween patients who recieved an early and continous cisatracurium infusion than those with usual care approach with lighter sedation targets?	[CHEST.pdf, ESCPARDS.pdf, ACURASYS.pdf, ENB.pdf, LSPA.pdf, NBSARDS.pdf, EDPMARDSLPMV.pdf, CEIIUPPSARDS.pdf, ROSE.pdf]	0.823973
4	5	Among patients with ALI/ARDS, does a conservative fluid management strategy improves lung function and decrease ventilator days compared to liberal strategy?	[ACURASYS.pdf, FACTT.pdf, OSCILLATE.pdf, FMWSCPARDS.pdf, ROSE.pdf, TOF-ARDS.pdf]	0.816809

Connection Neo4j with S3 Bucket

- Ask for Valid Credential
- Manually uploaded dataset





Next Steps

Date	Description
Feb. 28	Methods Writeup (M) Experimental results update (model refinement).
Mar. 7	Synthesis Experiment (paradigms in synthesis). Introduction write up. Document Annotation
Mar. 14	Troubleshooting. (Spring Break)
Mar. 19	JAMIA deadline
Mar. 20	In Class Presentation
Mar. 21	Model refinement and troubleshooting. Result write up. Draft symposium presentation.
Mar. 28	Finalize symposium presentation. Discussion write up.
Apr. 4	Capstone Symposium

Backup

Ground Truth Annotation I

- Each research paper was **annotated** in terms of **relevance and stance against *all*** in medical questions (yes-no) or claim. The annotations can be used for **automatic output evaluation**

Relevance Annotation

	Q1	Q2	…
Paper 1	1 (relevant)	0 (irrelevant)	…
Paper 2	…	1 (relevant)	…
…	…	…	…

Stance Annotation

	Q1	Q2	…
Paper 1	1 (supporting)	0 (neutral)	…
Paper 2	…	-1 (against)	…
…	…	…	…

BACK

Ground Truth Annotation II

- For each research paper, we leveraged large language models (specifically, Claude from AWS Bedrock), to **generate summaries**. **After editing and proofreading by physicians** (KN, MA), the summaries will serve as **ground truth** for automatic/ human evaluation.

	Study Design and Methodology	Study Population	Interventions	…	Key Findings and Conclusion
Paper 1			…		
Paper 2			…		
…	…	…	…		

Preliminary Results (Relevance)

	Precision	Recall	Accuracy
Overall			
ARDS (n questions)			
Sepsis (o questions)			
Septic Shock (p questions)			
Delirium (q questions)			

* Results derived from ZZ Papers (XX for ARDS, YY for Sepsis, ZZ for Septic Shock, WW for delium, KK for others)

** Temperature = 0, average over XX trials. *** Annotation