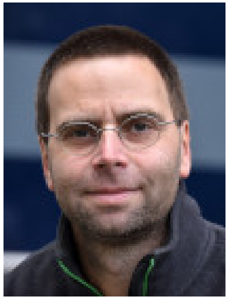
MiBi at BioASQ 2024:
Retrieval-Augmented Generation for Answering Biomedical Questions

September 9, 2024

Jan Heinrich
Merker

Alexander
Bondarenko



Matthias
Hagen

Adrian
Viehweger

Friedrich-Schiller-Universität Jena
Leipzig University

https://webis.de

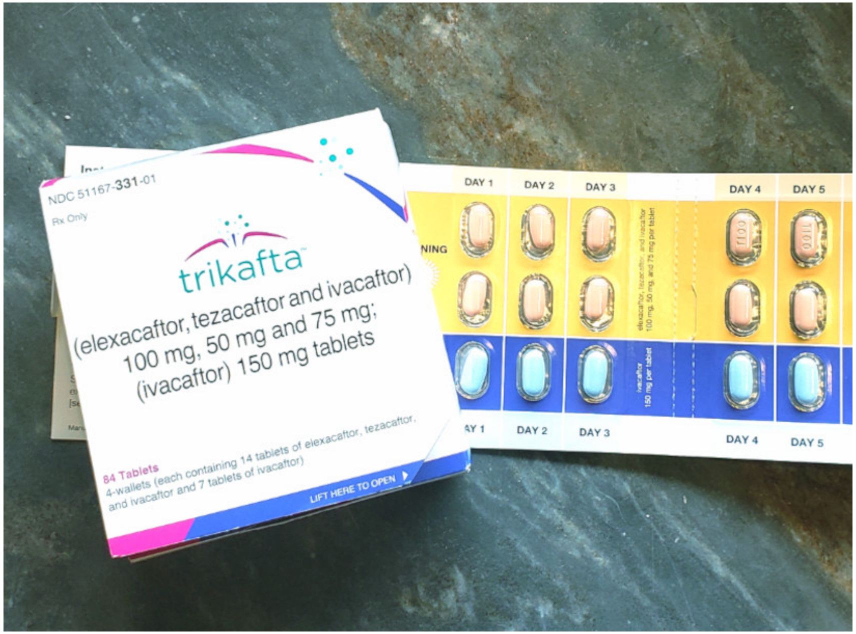
CLEF 2024
GRENOBLE

15th International Conference of the CLEF Association (CLEF 2024)

MiBi at BioASQ 2024

Medical Q&A

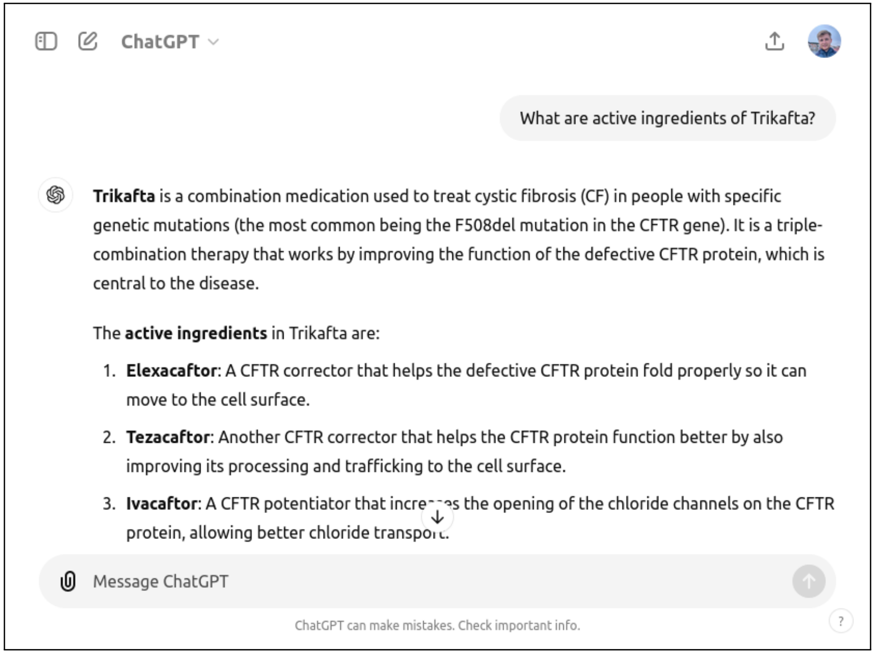
Example: What are active ingredients of Trikafta?


NDC 51167-331-01
Rx Only
trikafta
(elexacaftor, tezacaftor and ivacaftor)
100 mg, 50 mg and 75 mg;
(ivacaftor) 150 mg tablets
84 Tablets
4-wallets (each containing 14 tablets of elexacaftor, tezacaftor, and ivacaftor and 7 tablets of ivacaftor)
LIFT HERE TO OPEN
DAY 1
DAY 2
DAY 3
DAY 4
DAY 5
NING

MiBi at BioASQ 2024

Baselines

Example: What are active ingredients of Trikafta?



What are active ingredients of Trikafta?

Trikafta is a combination medication used to treat cystic fibrosis (CF) in people with specific genetic mutations (the most common being the F508del mutation in the CFTR gene). It is a triple-combination therapy that works by improving the function of the defective CFTR protein, which is central to the disease.

The **active ingredients** in Trikafta are:

1. **Elexacaftor**: A CFTR corrector that helps the defective CFTR protein fold properly so it can move to the cell surface.
2. **Tezacaftor**: Another CFTR corrector that helps the CFTR protein function better by also improving its processing and trafficking to the cell surface.
3. **Ivacaftor**: A CFTR potentiator that increases the opening of the chloride channels on the CFTR protein, allowing better chloride transport.

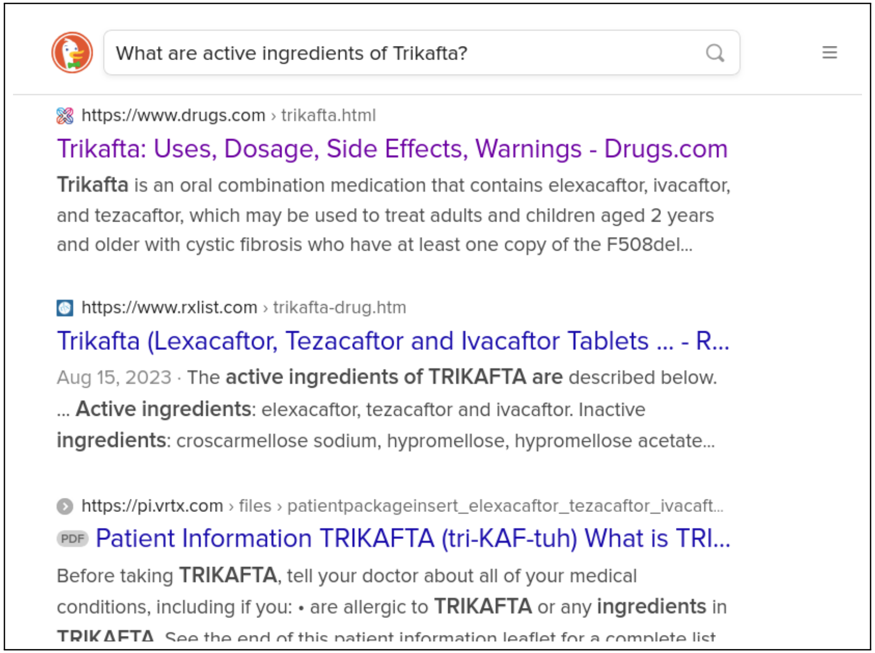
Why not just use GPT . . . ?

(correct ingredients, no dosage, no sources)

MiBi at BioASQ 2024

Baselines

Example: What are active ingredients of Trikafta?



What are active ingredients of Trikafta?

https://www.drugs.com › trikafta.html

Trikafta: Uses, Dosage, Side Effects, Warnings - Drugs.com

Trikafta is an oral combination medication that contains elexacaftor, ivacaftor, and tezacaftor, which may be used to treat adults and children aged 2 years and older with cystic fibrosis who have at least one copy of the F508del...

https://www.rxlist.com › trikafta-drug.htm

Trikafta (Lexacaftor, Tezacaftor and Ivacaftor Tablets ... - R...

Aug 15, 2023 · The **active ingredients of TRIKAFTA are** described below. ... **Active ingredients**: elexacaftor, tezacaftor and ivacaftor. Inactive **ingredients**: croscarmellose sodium, hypromellose, hypromellose acetate...

https://pi.vrtx.com › files › patientpackageinsert_elexacaftor_tezacaftor_ivacaft...

PDF Patient Information TRIKAFTA (tri-KAF-tuh) What is TRI...

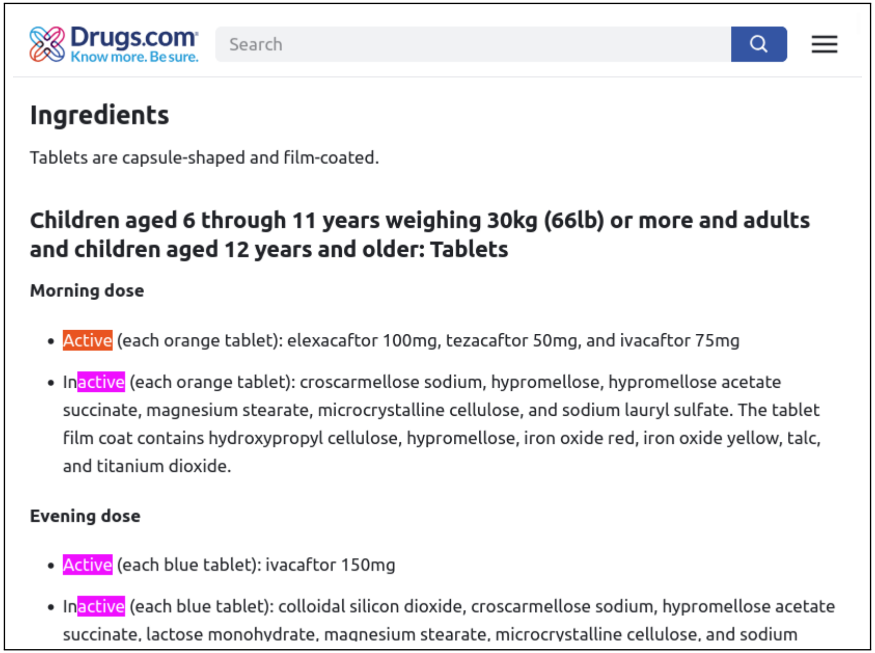
Before taking **TRIKAFTA**, tell your doctor about all of your medical conditions, including if you: • are allergic to **TRIKAFTA** or any **ingredients** in **TRIKAFTA**. See the end of this patient information leaflet for a complete list

… or a quick web search … ?

MiBi at BioASQ 2024

Baselines

Example: What are active ingredients of Trikafta?



Ingredients

Tablets are capsule-shaped and film-coated.

Children aged 6 through 11 years weighing 30kg (66lb) or more and adults and children aged 12 years and older: Tablets

Morning dose

- Active (each orange tablet): elexacaftor 100mg, tezacaftor 50mg, and ivacaftor 75mg
- Inactive (each orange tablet): croscarmellose sodium, hypromellose, hypromellose acetate succinate, magnesium stearate, microcrystalline cellulose, and sodium lauryl sulfate. The tablet film coat contains hydroxypropyl cellulose, hypromellose, iron oxide red, iron oxide yellow, talc, and titanium dioxide.

Evening dose

- Active (each blue tablet): ivacaftor 150mg
- Inactive (each blue tablet): colloidal silicon dioxide, croscarmellose sodium, hypromellose acetate succinate, lactose monohydrate, magnesium stearate, microcrystalline cellulose, and sodium

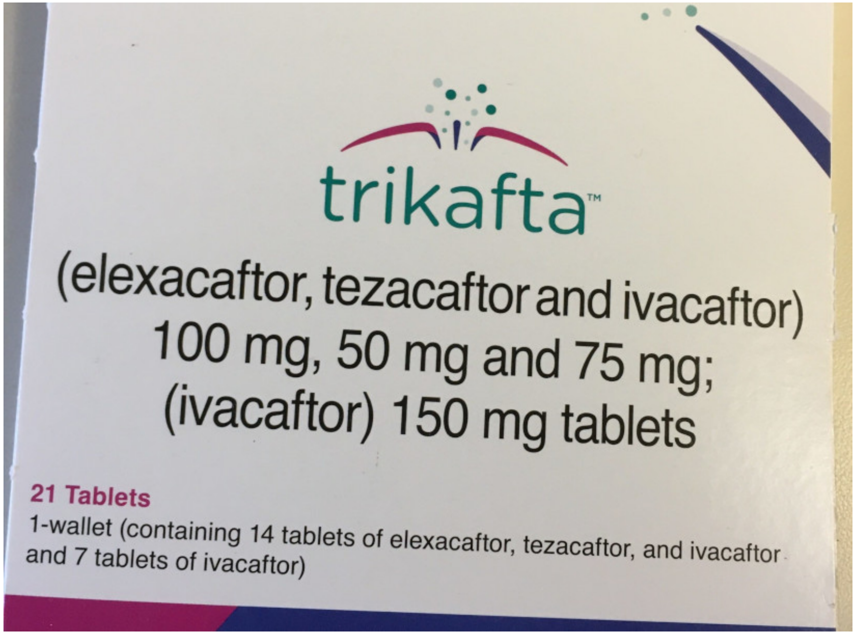
🌐 . . . with Ctrl+F on the first result?

(correct ingredients and dosage, good source, but takes longer)

MiBi at BioASQ 2024

Baselines

Example: What are active ingredients of Trikafta?



trikafta™

(elexacaftor, tezacaftor and ivacaftor)
100 mg, 50 mg and 75 mg;
(ivacaftor) 150 mg tablets

21 Tablets

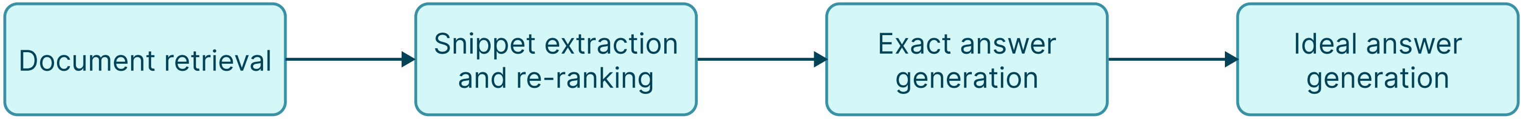
1-wallet (containing 14 tablets of elexacaftor, tezacaftor, and ivacaftor and 7 tablets of ivacaftor)

🙄 . . . and what about this?

(correct ingredients and dosage, good source, fastest?!)

MiBi at BioASQ 2024

Goal: RAG for Medical Questions



Stages

❑ Document retrieval → Find relevant medical articles (from PubMed).

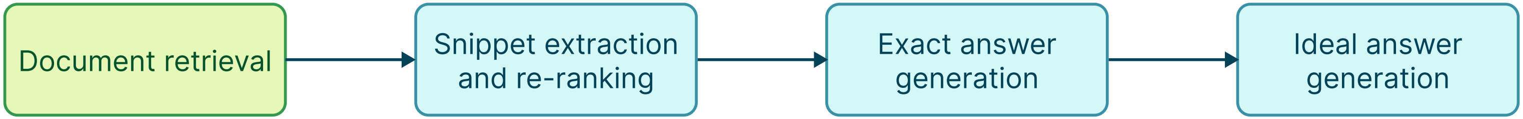
❑ Snippet extraction and re-ranking → Extract snippets and rank by relevance.

❑ Answer generation → Generate exact answer and “ideal” summary answer.

❑ RAG → Combine retrieval- and generation-focused components.

MiBi at BioASQ 2024

Approaches: Document Retrieval



Goal: Find relevant medical articles (from PubMed).

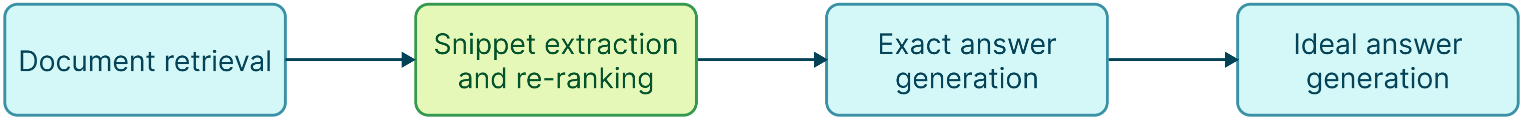
- PubMed search API
  - Re-rank with BM25, MiniLM, and MPNet

- Custom BM25 index with metadata (Elasticsearch)
  - Match abstract, title, and MeSH terms
  - Disallow non-peer-reviewed publication types

→ Do we need to wory about indexing?

MiBi at BioASQ 2024

Approaches: Snippet Extraction and Re-Ranking



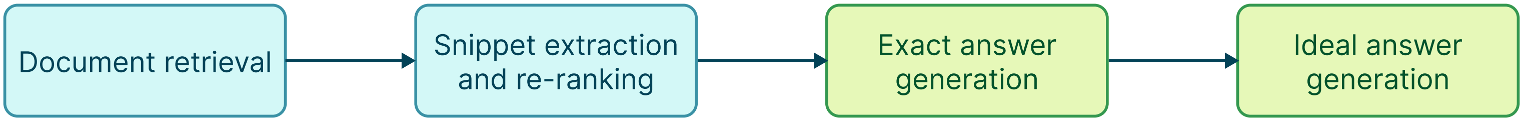
Goal: Extract concise snippets from the article's abstract (or title).
Rank extracted snippets by relevance to the question.

- ❑ Using LLMs
  - Chain-of-thought 3-shot prompt (GPT-3.5-turbo)
  - No re-ranking
- ❑ Rule-based
  - Split abstract in sentences
  - Candidates: full title + sentence n-grams (up to 3 sent.) from abstract
  - Re-rank with TAS-B and duoT5

→ Are LLMs better at snippet extraction?

MiBi at BioASQ 2024

Approaches: Answer Generation with LLMs



Goal: Generate exact (e.g., yes-no) answer and “ideal” summary answer.

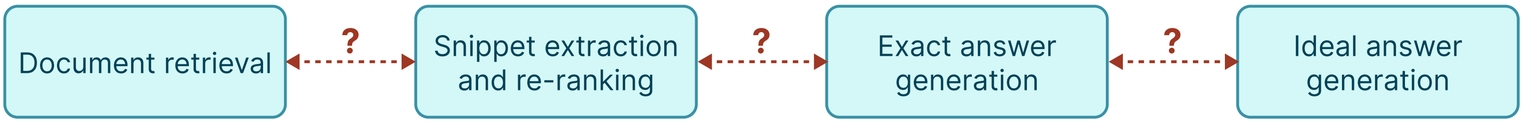
- Few-shot prompting with function calling
  - Manual prompts per question/answer type (e.g., yes-no / exact)
  - Context: top-3 abstracts or all (top-10) snippets
  - GPT-3.5-turbo and GPT-4

- Modular “programming” with DSPy
  - Automatic prompts via DSPy (signature of in-/outputs are Python classes)
  - Context: abstracts, snippets, previous answer
  - Mixtral-7B

→ Do we need manual prompts?

MiBi at BioASQ 2024

Approaches: RAG paradigms



Goal: Combine retrieval- and generation-focused components.

- ❑ Retrieve-then-generate (exact → ideal → documents → snippets)
- ❑ Generate-then-retrieve (documents → snippets → exact → ideal)
- ❑ GtRtG / RtGtR (e.g., exact → ideal → documents → snippets → exact → ideal)
- ❑ Let the LLM decide (DSPy, Mixtral-8x7B)

→ Which paradigm to use when?

MiBi at BioASQ 2024

Results

- 42 submitted runs, different systems per phase and batch
- Retrieval:
  - PubMed search API struggles with question-like queries
  - Enhancing index with metadata pays off for domain-specific retrieval
- Snippet extraction and re-ranking:
  - Neither GPT- nor rule-based snippet extraction competitive
- Answer generation:
  - Snippets instead of abstracts as context less “confusing” for LLMs
  - Models: GPT-4 $>>$ GPT-3.5 $>$ Mixtral-7B
  - No difference by prompting strategy (manual vs. DSPy)
- RAG paradigm:
  - With ground truth: RtG, GtRtG
  - Without ground truth: GtR, GtRtG

MiBi at BioASQ 2024

Summary

- Mixed results, but some recommendations:
  - Put work into first-stage retrieval
  - Show snippets to LLMs, not long texts
  - Use the latest LLMs (e.g., GPT-4)
  - GtRtG seems to work well with/without ground-truth evidence

- Limitation: comparability across test batches

- Future work: systematic evaluation of grounded RAG paradigms

Code and Data

`github.com/webis-de/CLEF-24`

MiBi at BioASQ 2024

Summary

- Mixed results, but some recommendations:
  - Put work into first-stage retrieval
  - Show snippets to LLMs, not long texts
  - Use the latest LLMs (e.g., GPT-4)
  - GtRtG seems to work well with/without ground-truth evidence

- Limitation: comparability across test batches

- Future work: systematic evaluation of grounded RAG paradigms

Code and Data

`github.com/webis-de/CLEF-24`

Thank you & merci!