

Revolutionizing Literature Reviews with a GPT-Assisted Tool: Enhanced Search, Extraction, and Visualization

Martin Schärer, Hoffman La Roche, Basel, Switzerland
Yingjia Chen, Genentech, South San Francisco, USA

Introduction

Overview of the Problem

Literature reviews are a cornerstone in **pharmacovigilance** and **drug safety**, crucial for identifying drug safety signals and assessing **causal relationships** between drug exposure and adverse events (AEs).

Key Challenges

- **Resource-Intensive:** Manual search, screening, and extraction require significant time and effort, often involving **multiple teams** or **CROs**.
- **Subjective Variations:** Review outcomes can vary based on individual expertise, leading to **inconsistencies** in data interpretation.
- **Information Overload:** Difficulties in visualizing large datasets hinder effective **decision-making** and **communication** with stakeholders.

Objective

This study introduces a **GPT-assisted tool** to **automate** critical aspects of the literature review process, improving **efficiency**, **consistency**, and **accuracy**. This presentation focuses on **drug-event associations**, something routinely assessed in drug safety.

Current Literature Review Process

Step 1: Request and Initial Search

- A **safety scientist** requests a literature review for a specific **drug-event association**.
- The literature review team conducts an initial **broad search** using specialized databases like **Clarivate Dialog**.
- This search yields a **large dataset** of publications, often numbering in the hundreds or thousands.

Step 2: Outsourcing to CROs

- Due to the **large volume** of search results, the dataset is often outsourced to **Contract Research Organizations (CROs)**.
- **CRO reviewers** assess the **full text** of each publication, refining the dataset to identify **relevant studies**.

Step 3: Manual Selection and Summary

- Selected publications are returned to the **drug safety team** as a **compiled sheet**.
- A **drug safety scientist** further reviews the list, performing **manual screening** for relevance and **extracting key information**.

Current Literature Review Process (Continued)

Step 4: Data Extraction and Analysis

- The scientist extracts **critical data points** like **measures of association** (e.g., RR, ROR, HR) for analysis.
- The data is then **compiled and summarized** into reports for decision-making, often without the benefit of **visualizations**.

Challenges in the Current Process

- **Time-Consuming**: Multiple manual steps lead to **delays** in analysis and **slower** for quick turnarounds.
- **Inconsistent Interpretations**: Variations in judgments across different reviewers and teams.
- **Information Overload**: Lack of **effective visualization** makes it difficult to **synthesize** the large volume of data.

Solution Overview

Introducing MetaBooster: A GPT-Assisted Literature Search



- **Automates Key Steps** in the literature review process, reducing manual effort and increasing efficiency.
- Developed using **GPT-4o**, integrated with **Python** and **Streamlit** for an intuitive user interface.
- The example demonstrated is focused on **drug-event associations**.

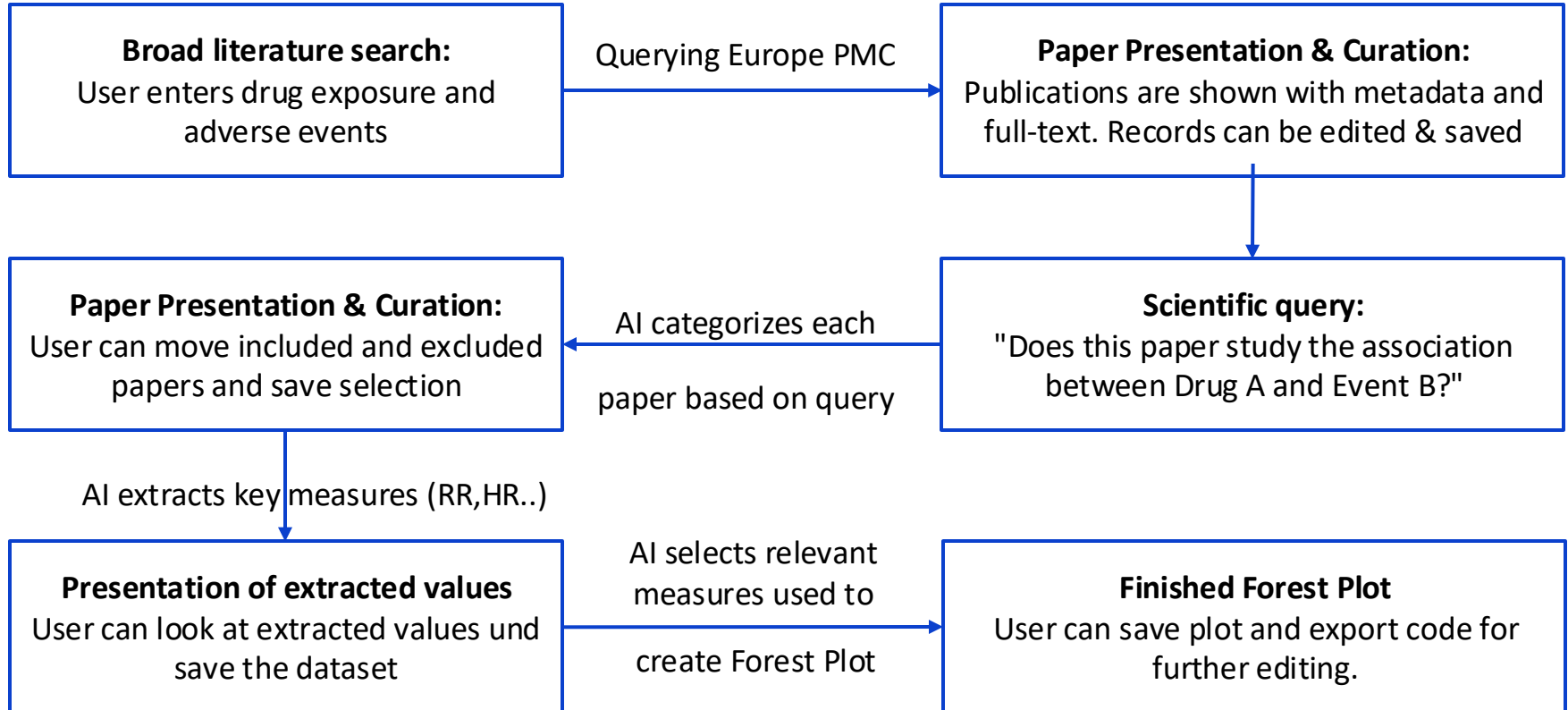
Core Capabilities

- **Automated Search & Retrieval:** Queries Europe PMC and other databases using **customized search strings**, tailored to specific research questions.
- **Categorization of Studies:** Uses GPT to classify publications into “**included**” and “**excluded**” groups based on relevance to predefined criteria.
- **Data Extraction:** Automatically extracts **key measures** like **RR**, **ROR**, and **HR** from relevant studies providing a foundation for literature reviews and supporting meta-analysis.
- **Visual Summaries & Reporting**
- Generates **forest plots** to visualize measures of association, providing a clear, integrated view of findings.
- **Reduces Information Overload** by presenting data in a structured, visual format, aiding decision-making.
- Results can be **saved** and **revisited**, allowing for **iterative analysis** and **updates**.

Technology Stack

Technology	Description
GPT-4o Model	GPT-4o (“Omni”): OpenAI's most advanced model, designed for complex, multi-step tasks. High Context Window: Up to 128,000 tokens, allowing the analysis of long and detailed texts. Cost-Efficiency: \$2.5 per million input tokens, \$10 per million output tokens as of October 2024.
Python and Streamlit	Python 3.9: Chosen for its extensive scientific libraries and NLP support, ideal for integrating GPT. Streamlit: Provides a user-friendly interface for interacting with the AI model, enabling rapid prototyping and visualization. Python and Streamlit work together to allow for interactive querying, data display, and real-time adjustments.
Data Sources & APIs	Europe PMC Web Service: Access to over 33 million publications including PubMed and other major databases. - Automated Queries for drug-event associations, tailored to focus on specific outcomes. - Retrieves metadata, abstracts, and full texts for further analysis. Azure OpenAI Service: Provides API access to the GPT-4o model for text processing and data extraction.

Workflow Overview



Initial Literature Search & Retrieval

MetaBooster

Exposure

Atezolizumab

Outcome

"Immune mediated hepatitis"

Publication Type

Choose an option

Start Date

YYYY-MM-DD

Stop Date

2024-10-03

Query articles

Previously saved Analysis

new3_exp3

	☐ Exclude	☐ Extraction Method	☐ PMID	☐ PMID	☐ Title	☐ Fulltext	☐ Publication Link
1	☐	Abstract + HTML	37,660,741	None	Acute liver failure following immune checkpoint inhibitors.	We report the case of a 64-year-old man admitted to intensive care unit for liver failure.	https://doi.org/10.1016/j.cline.2023.102203
2	☐	Fulltext XML	38,018,074	PMC10990673	Complications of immunotherapy in advanced hepatocellular carcinoma.	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.17998/jlc.2023.11.21#free-full
3	☐	Fulltext XML	36,881,599	PMC9990950	Hepatotoxicity in immune checkpoint inhibitors: A pharmacovigilance study from 20	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1371/journal.pone.0281983
4	☐	Fulltext XML	36,871,392	PMC10163154	Safety and clinical activity of atezolizumab plus erlotinib in patients with non-small-c	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1016/j.esmoop.2023.101160
5	☐	Fulltext XML	38,596,295	PMC10905042	Promising immunotherapy targets: TIM3, LAG3, and TIGIT joined the party.	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1016/j.onton.2024.200773
6	☐	Fulltext XML	37,808,223	PMC10557510	Safety and Efficacy of Atezolizumab and Bevacizumab Combination as a First Line Tr	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.2147/jhc.s347932#free-full-tx
7	☐	Fulltext XML	37,095,751	PMC10122589	Early Experience of TACE Combined with Atezolizumab plus Bevacizumab for Patient	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1155/2023/6353047#free-full
8	☐	Fulltext XML	38,236,364	PMC10954993	Tremelimumab: A Review in Advanced or Unresectable Hepatocellular Carcinoma.	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1007/s11523-023-01026-9
9	☐	Fulltext XML	38,545,481	PMC10965262	Predictors of Immunotherapy Efficacy in Metastatic Non-Small Cell Lung Cancer: Lanq	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.14740/wjon1790#free-full-tee
10	☐	Fulltext XML	35,987,165	PMC9588873	Isatuximab plus atezolizumab in patients with advanced solid tumors: results from a	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1016/j.esmoop.2022.100562

Out of 154 publications in the search result, 120 had the full text automatically extracted from the XML full texts, 33 had to use the publication full text pdf and one extraction was unsuccessful and had to resort to the abstract text only. Manual adjustments were made in the user interface and a final selection of 137 publications were used for further analysis

MetaBooster

Dreg

Exposure

Atezolizumab

Publication Type

Choose an option

Start Date

YYY-MM-DD

Query articles

Outcome

"immune mediated hepatitis"

Stop Date

2024-10-23

Previously saved Analysis.

PCV43

	Fulltext	Publication Link	Author
us Bevacizumab for Patient-	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1155/2023/6353047#free-full-text	Wang K, Zhu H, Yu H, Cheng Y, Xiang Y, Cheng Z, Li Y, Li T, Wang D, Zhu Z
ipato cellular Carcinoma.	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.1007/s11523-023-01026-9#free-full-text	France NL, Blair HA
small Cell Lung Cancer: Lung	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch	https://doi.org/10.14740/wjon1790#free-full-text	Moratelli Pellitero A, Zapata García M, Gascon Ruiz M, Arbones-Mainar JM
olid tumors: results from a	atezolizumab (IsaRx000a0; Atez) in immunotherapy-naïve patients with advanced solid tumors who received at least one previous line of treatment for their advanced disease. Types of tumors included platinum-resistant/refractory epithelial ovarian cancer (EOC), recurrent GBM, unresectable HCC, and platinum-refractory recurrent squamous cell carcinoma of the head and neck (SCCHN). Cohort selection for this trial was based on immune responsiveness, with HCC and SCCHN being moderately immune-responsive to anti-PD-1/PD-L1 monotherapy, and GBM and EOC being immune-resistant cancers for which PD-1/PD-L1 axis blockade alone has not been shown to be effective. <p>sec id="sec1.1">title=Patients and Methods</p>sec id="sec1.1">title=Study design and patients</p>id="p0550">This was an open-label, multicenter, phase III study (NCT03637764) designed to evaluate the safety, preliminary efficacy, pharmacokinetics (PK), and pharmacodynamics of IsaRx000a0. Atez in patients with advanced solid malignancies. The study comprised two parts. Phase 1, the safety run-in, characterized safety and tolerability, and assessed the	https://doi.org/10.1016/j.esmoop.2022.100562#free-full-text	Simonelli M, Garralda E, Eskens F, Gil-Martin M, Yen CJ, Obermannova R, C
vanced Hepatocellular Car		https://doi.org/10.3390/cancers30090628#free-full-text	Foy V, McNamara MG, Valle JW, Lamarca A, Edeline J, Hubner RA
erapy in Patients with Can		https://doi.org/10.3390/cancers16040795#free-full-text	Lasagna A, Sacchi P
int inhibitors: Current upd		https://doi.org/10.3389/fphar.2022.107746#free-full-text	Liu Z, Zhu Y, Xie H, Zou Z
d hepatocellular carcinoma		https://doi.org/10.1177/17588359221080298#free-full-text	de Castro T, Jochheim LS, Bathon M, Welland S, Scheiner B, Shmanko K, F
b, as Assessed for Causality		https://doi.org/10.1155/2022/5090200#free-full-text	Tzadok R, Levy S, Aouizerate J, Shibolet O
ic atezolizumab and bevac		https://doi.org/10.3389/fimmu.2023.1107542#free-full-text	Wang K, Xiang XJ, Yu HM, Cheng YQ, Liu ZH, Zhong JY, Feng S, NI QZ, Zhu

Update Data

Save Article Set

pcv3

AI-Driven Categorization

Query

Analyzing this paper, is it possible to extract the ratio values (RR, HR, OR) for either individual ICIs (such as Atezolizumab) or for groups of ICIs (such as PD-L1 inhibitors, PD-1 inhibitors, CTLA-4 inhibitors) in the context of "Immune-mediated hepatitis" where

Included

☐	☐ Move to excluded	☐ Title	☐ Inclusion/Exclusion Reason	☐ Fulltext
12	☐	The ABC of Immune-Mediated Hepatitis during Immunotherapy in Patients with Canc	Analyzing the provided text, it is possible to extract some ratio values (RR, HR, OR) for	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
36	☐	Incidence rate and treatment strategy of immune checkpoint inhibitor mediated hep	Yes, it is possible to extract the ratio values (RR, HR, OR) for groups of ICIs such as PD-	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
65	☐	Common Immune-Related Adverse Events of Immune Checkpoint Inhibitors in the G	To determine if the ratio values (Risk Ratio (RR), Hazard Ratio (HR), Odds Ratio (OR)) f	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
78	☐	Hepatotoxicity in patients with solid tumors treated with PD-1/PD-L1 inhibitors alone	Analyzing the provided paper, it is possible to extract ratio values (RR) for individual I	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch

Excluded

☐	☐ Move to included	☐ Title	☐ Inclusion/Exclusion Reason	☐ Fulltext
4	☐	Safety and clinical activity of atezolizumab plus erlotinib in patients with non-small-c	The paper provides detailed information on the safety and efficacy of combining atez	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
5	☐	Promising immunotherapy targets: TIM3, LAG3, and TIGIT joined the party.	The article provided details various clinical trials and results related to immune che	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
6	☐	Safety and Efficacy of Atezolizumab and Bevacizumab Combination as a First Line Tre	The provided document is a comprehensive review article discussing the safety and e	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
7	☐	Early Experience of TACE Combined with Atezolizumab plus Bevacizumab for Patient	The article provided is a research study that evaluates the efficacy and safety of TACE	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
8	☐	Tremelimumab: A Review in Advanced or Unresectable Hepatocellular Carcinoma.	The provided article focuses on tremelimumab and its combination with durvalumab	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
9	☐	Predictors of Immunotherapy Efficacy in Metastatic Non-Small Cell Lung Cancer: Lung	<Answer>No</Answer> <Explanation>This paper does not provide specific ratio valu	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
10	☐	Isatuximab plus atezolizumab in patients with advanced solid tumors: results from a	The paper provided is a detailed description of a clinical study that evaluates the con	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
11	☐	Current Evidence for Immune Checkpoint Inhibition in Advanced Hepatocellular Carc	The provided document contains a detailed review of immune checkpoint inhibitors	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
13	☐	Immune-mediated hepatitis induced by immune checkpoint inhibitors: Current upda	The provided paper is a comprehensive review on immune-mediated hepatitis (IMH)	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
14	☐	Atezolizumab and bevacizumab in patients with advanced hepatocellular carcinoma	The provided document is a research article that primarily discusses the efficacy and	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch
15	☐	Analysis of Efficacy Following a Single Dose of Atezolizumab, or Atezolizumab Plus Durvalumab	This provided article does not provide ratio values (RR, HR, OR) for either individual I	<DOCTYPE article PUBLIC "-//NLM//DTD JATS (Z39.96) Journal Archiving and Interch

new3

Total session cost: \$12.81

Data Extraction for Meta-Analysis

Analyzed Papers

Paper ID	Source Data	Disease Population	Time Span	Sample Size	Measure of Association	Measure of Association with 95% CI	
78	ns alone	The dataset used in this study was sourced from the Pub	The disease population used in this study consisted of patients with solid t	2 to 12 weeks following initiation of IC1 treatme	11,634	RR	[RR] PD-1/PD-L1 inhibitor vs Chemotherapy, 1
65	in the G:	The dataset used in this study was obtained from the US	The disease population used in this study included patients who experien	July 1, 2014, to December 31, 2020	70,330	Reporting Odds Ratio (ROR)	[ROR] IC1 vs Any drug, Colitis: 18.1 (17.2, 19.1)
12	9th Canc	The dataset used in this study is from the FAERS (FDA Ad	The disease population in this study comprises patients with cancer who a	The time span of the included observations in th	213	OR	[IC] IC1s, Liver failure: 1.24 (1.06, 1.42) [ROR] 1
36	ited hep	In the article, the dataset was sourced from four databas	The disease population used in this study consists of oncology patients wi	January 1, 2010, to December 31, 2021	18,515	RR	[RR] IC1 vs Control, Overall HAE: 1.40 (not specified) [RR] IC1 vs Control, Severe HAE: 2.55 (not specified) [RR] IC1 + ST vs ST, Overall HAE: 1.41 (1.17, 1.71) [RR] IC1 vs Placebo, Overall HAE: 2.26 (1.22, 4.19) [RR] IC1 vs ST, Overall HAE: 1.09 (0.88, 1.36) [RR] CTLA-4 vs Control, Overall HAE: 1.53 (0.98, 2.37) [RR] PD-1 vs Control, Overall HAE: 1.32 (1.04, 1.67) [RR] PD-1 vs Control, Overall HAE: 1.37 (1.18, 1.58) [RR] CTLA-4 vs Control, Severe HAE: 2.95 (1.03, 8.42) [RR] PD-1 vs Control, Severe HAE: 2.31 (1.69, 3.15) [RR] PD-1 vs Control, Severe HAE: 1.52 (1.11, 2.08)

Finished

Save analyzed article set

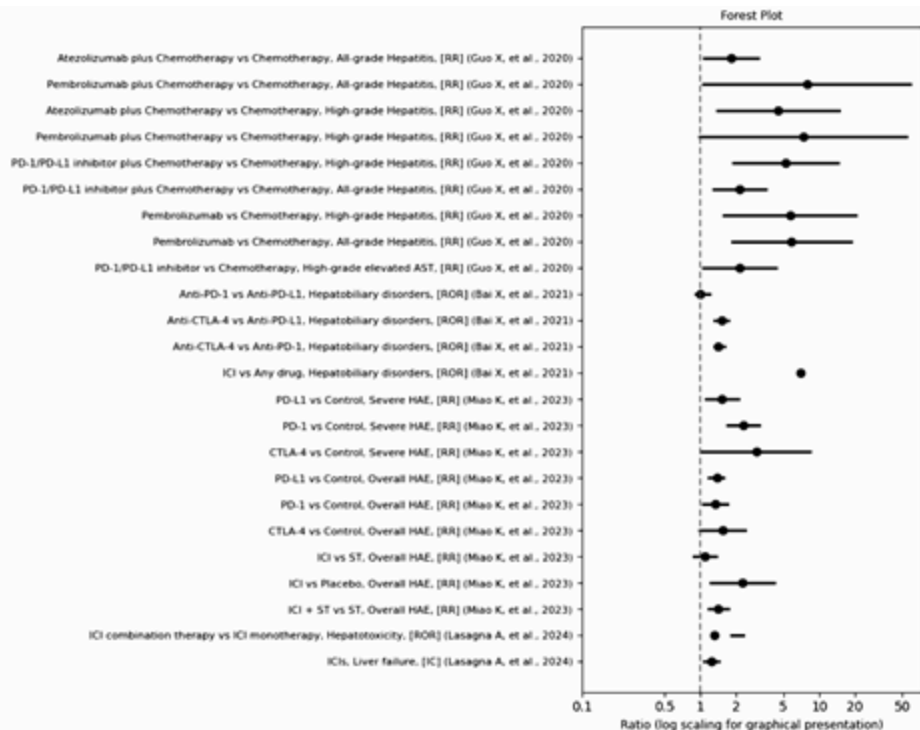
Make Forest Plot

- **Six prompts were used to extract 8 values:**
 Source Data, Disease Population, Time Span, Sample Size, Measure of Association,
 Explanation of Measure of Association, Interpretation of Measure of Association, Measure of Association with 95% CI
- **Each prompt is applied to each selected publication**
- **Simple extractions can be done in bulk (1 prompt, several extractions)**
- **More complex tasks like extraction of values in formatted form need single prompt using examples**

Example Prompts used

Description	Prompt
Sample Size (Simple value extraction)	In the article above, what is the sample size? Put the number in the tags <Answer></Answer>
Measure of association (Extraction of several values at once)	In the article above, First describe the measure of association that was used in this study. Then do the following: 1) In tags <MA></MA> put just the term for the measure of association, e.g. RR, OR etc. 2) In tags <EX></EX> put a short explanation of the measure of association used 3) In tags <IN></IN> put an interpretation of the used measure of association
Actual formatted values (Complex extraction of values using example format)	In the article above, can you find the ratios (RR, ROR, HR, OR) and their 95% confidence interval for a drug and adverse event? Here is an example: 1) ROR for Aspirin vs Advil and Fever is 3.5 with 95% CI (2.6, 9.8). 2) ROR for Aspirin vs Tylenol and Headache is 4.1 with 95% CI (4.0, 4.2) 3) HR for Paracetamol vs Aspirin and Anaemia is 5.5 with 95% CI (4.1, 5.2) In this case write the answer as "<Answer>[ROR] Aspirin vs Advil, Fever: 3.5 (2.6, 9.8) [ROR] Aspirin vs Tylenol, Headache: 4.1 (4.0, 4.2) [HR] Paracetamol vs Aspirin, Anaemia: 5.5 (4.1, 5.2)</Answer>" Now put the answer for the paper above in this format: <Answer></Answer>

Forest Plot Generation



Generated Forest Plot

Finished

Save everything with forest plot

Download Forest Plot Python code

- Relevant measures are selected from the compiled results (previous slide) by applying a further prompt specifying scientific background
- To get consistent results chunks of data are processed (one prompt per publication's extracted measures)
- Results are more consistent by asking the engine to explain reasoning for inclusion
- Combined answers of included measures are then fed to another prompt summarizing and formatting the results into python readable code:


```
[
  ("Regimen 1, Adverse Event 1, [RR]", "Zanuso V, et al.",
   2023, 0.66, 0.52, 0.85),
  ("Regimen 2, Adverse Event 1, [HR]", "Pirozzi A, et al.",
   2023, 0.65, 0.53, 0.81)
  ...
]
```
- Plot is generated and code used can be saved and downloaded for further editing

Results Summary

Improved Efficiency

- **Time Savings:** MetaBooster significantly reduced the time spent on **manual literature screening**, with AI-driven categorization taking just **7 minutes** for over **130 publications**.
- **Automated Data Extraction:** Streamlined the process of extracting **key measures** (e.g., **RR, ROR, HR**) from selected studies, saving **substantial amount of time** compared to traditional methods.

Enhanced Accuracy & Consistency

- **Minimized Subjectivity:** By applying **consistent criteria** across all publications, MetaBooster reduced **variations in judgment** that typically arise in manual reviews.

Effective Data Visualization

- **Forest Plot Generation:** Automatically generated **forest plots** provided a **clear visual representation** of measures of association, aiding in meta-analysis and **decision-making**.

Lower Costs

- Reduced dependency on **CROs** and **manual labor** results in **cost savings** for drug safety assessments.

Limitations

Prompt Design Challenges

- **Time-Intensive Optimization:** Achieving high accuracy required multiple iterations to **refine prompts**, particularly for complex queries.
- **Variability in Responses:** Prompts that worked well in one context sometimes required **adjustments** in others, adding to **development time**.

Dependence on Expert Oversight

- **Human Judgment Still Essential:** While automation reduced manual effort, **expert review** remains critical for interpreting **nuanced data** and ensuring **relevance**.

Model Limitations

- **AI Generalization:** GPT-4o's understanding is broad but may miss **domain-specific subtleties**, especially in highly specialized research topics.
- **Cost Considerations:** Running complex queries on a large number of publications can become **costly**, though still lower than outsourcing to **CROs**.

Conclusion

Key Takeaways

- **Streamlined Literature Reviews:** MetaBooster has demonstrated significant improvements in **efficiency** and **accuracy**, automating time-consuming tasks like **screening**, **data extraction**, and **visualization**.
- **Balancing Automation with Expertise:** While AI can handle much of the **manual workload**, **expert oversight** remains critical for interpreting results and making **final decisions**.

The Future: Improved with New AI Models

- **Adaptability to New Models:** MetaBooster's architecture is designed to **seamlessly integrate** new AI advancements, ensuring **better performance** and **accuracy** as OpenAI releases more advanced models.

Learnings on Prompting

- **Layered Approach to Complexity:** Complex tasks, like verifying extracted data, benefit from **multi-step prompting**, breaking down tasks into **manageable steps** for better accuracy. Asking for too much in one step deteriorates quality of answers.



Doing now what patients need next