

Stream 4: Synthesising, appraising and exploring medical evidence

Simon Šuster

with Tim Baldwin, Antonio Jimeno Yepes, Jey Han Lau,
David Martinez Iraola, Yulia Otmakhova, Karin Verspoor

5 November 2020



Evidence-based decision support



clinical scenario



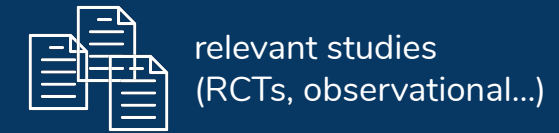
- relate patient characteristics to inclusion/exclusion criteria of guidelines
- use summaries or direct clinical questions

suggest most relevant guideline (IR & NLP)

guidelines
(recommendations for specific questions)



summarisation



exploration
and selection



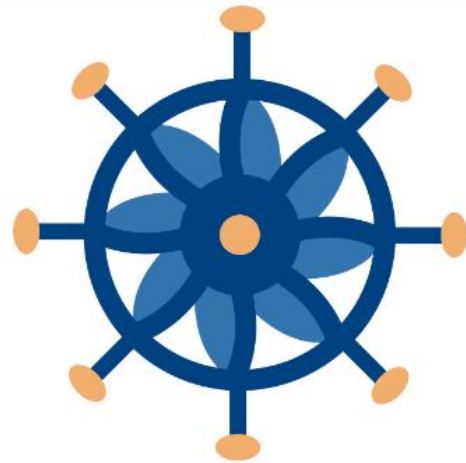
medical literature

clinical question

in **PICO** format:
Population/Problem
Intervention
Comparison
Outcome

IR & NLP

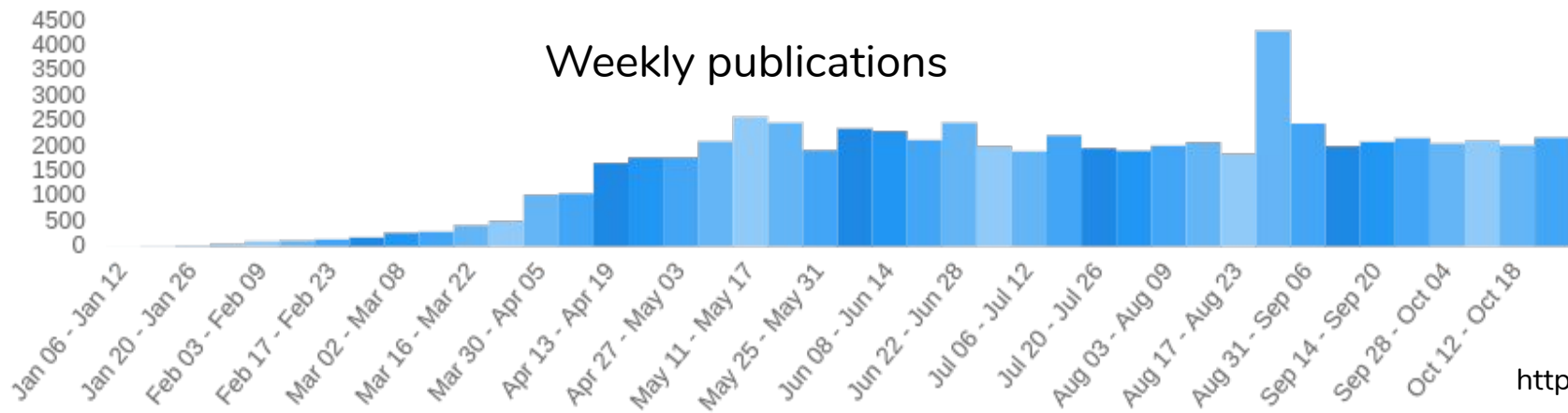
Exploring scientific evidence



COVID-SEE
Scientific Evidence Explorer

Verspoor, K., Šuster, S., Otmakhova, Y., Mendis, S., Zhai, Z., Fang, B., Lau, J.H., Baldwin, T., Yepes, A.J. and Martinez, D., 2020. COVID-SEE: Scientific Evidence Explorer for COVID-19 related research. arXiv preprint arXiv:2008.07880.

Otmakhova, Y., Verspoor, K., Baldwin, T. and Suster, S., 2020. Improved Topic Representations of Medical Documents to Assist COVID-19 Literature Exploration. EMNLP 2020 Workshop NLP-COVID.



<https://www.ncbi.nlm.nih.gov/research/coronavirus/>

Medical literature around COVID-19 is the major hub of research activity!

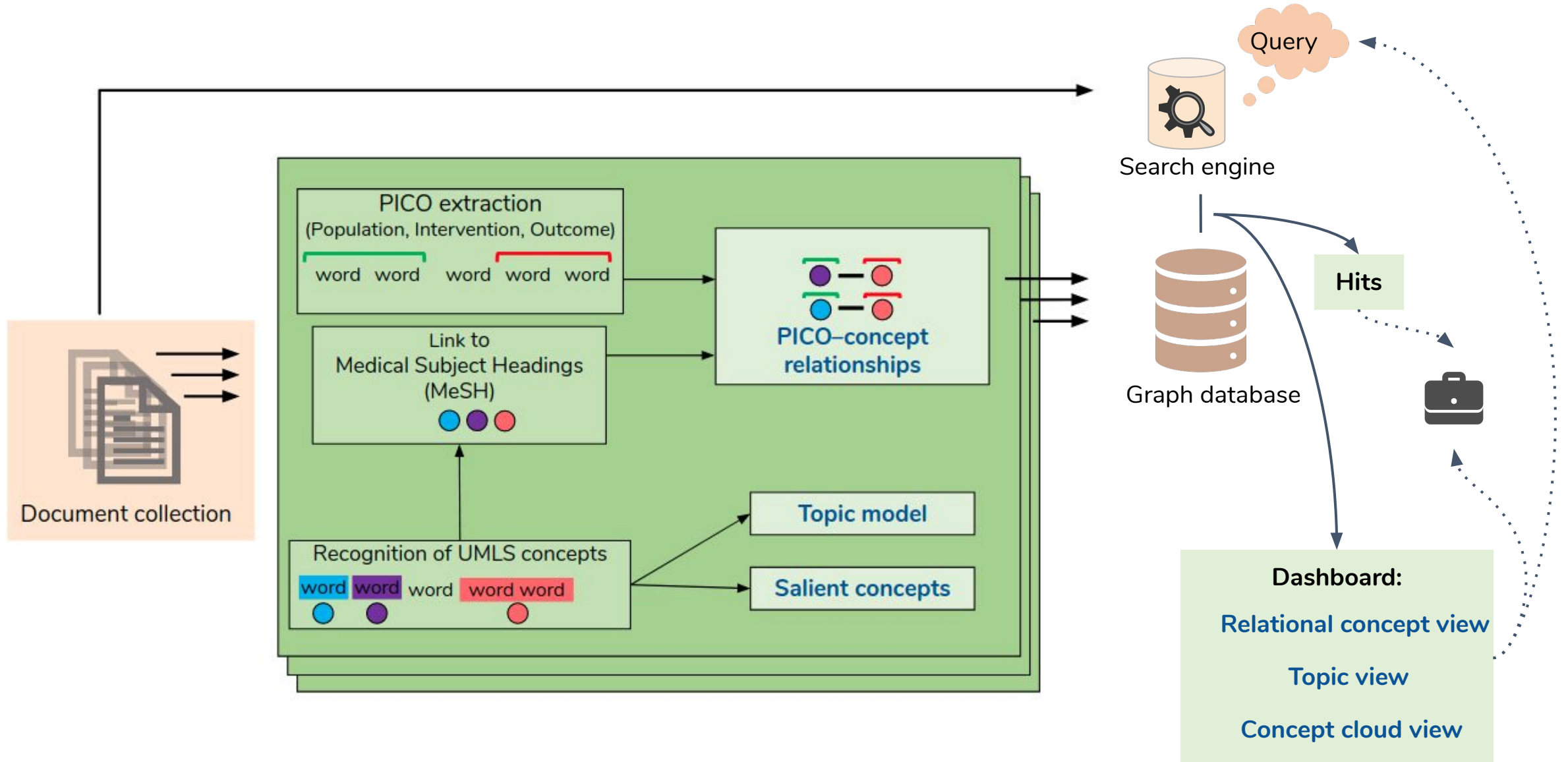
	COVID-SEE (ours)	SciSight	DOC Search	COVID-19 Navigator	LitCOVID	Sem Viz	WellAI	COVID Intel. Search	Le Bras et al. 2020	COVID-19 LOVE	Trialstreamer
Search:											
NL/IR	✓	✗	✓	✗	✓	✗	✗	✓	✗	✗	✓
Concepts	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PICO	✓	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓
Visualisation:											
Concepts	✓	?	?	✗	✗	✓	✗	✗	✗	✗	?
Relations	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗
Topics	✓	✗	✗	✗	✗	✗	✗	✗	?	✗	✗

Our goals

- Design a system to navigate the COVID-19-related literature effectively
- Use NLP methods, including concept recognition, detection of PICO elements and topic modeling, to structure the key information
- Offer visual overviews of document content to facilitate information discovery



System overview



hydroxychloroquine for treatment of Covid-19



We found 44 potential matches:

Filter

Year Published:

2020 (44)

Source:

- Elsevier (9)
- MedRxiv (20)
- Medline (20)
- PMC (24)
- WHO (14)

Clear All

Review: Hydroxychloroquine and Chloroquine for Treatment of SARS-CoV-2 (COVID-19)

Authors: Pastick, Katelyn A; Okafor, Elizabeth C; Wang, Fan; Lofgren, Sarah M; Skipper, Caleb P; Nicol, Melanie R; Pullen, Matthew F; Rajasingham, Radha; McDonald, Emily G; Lee, Todd C; Schwartz, Ilan S; Kelly, Lauren E; Lother, Sylvain A; Mitjà, Oriol; Letang, Emili; Abassi, Mahsa; Boulware, David R

Journal: Open Forum Infect Dis, 2020

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a rapidly emerging viral infection causing coronavirus disease 2019 (COVID-19). Hydroxychloroquine and chloroquine have garnered unprecedented attention as potential therapeutic agents against COVID-19 following several small clinical trials,...[+]

Systematic benefit-risk assessment of hydroxychloroquine and chloroquine for COVID-19: Evidence synthesis

Authors: Osborne, V.; Davies, M.; Dhillon, S.; ...

Journal: N/A, 2020

Objectives: Given the current paucity of evidence on the effectiveness of hydroxychloroquine and chloroquine (coronavirus disease). A systematic review and meta-analysis of the benefit-risk balance of these drugs for the treatment of COVID-19.



COVID-SEE
Scientific Evidence Explorer

About

Search ▾

Collections

Visualiser

Query:

Coronavirus AND (Hydroxychloroquine OR Chloroquine) AND Mortality

search

Logical Operators:

() AND OR

Find MESH Terms:

heart ar

Outcome Heart Arrest

Population Heart Arrest

Intervention Heart Arrest

Outcome Heart Arrest, Induced

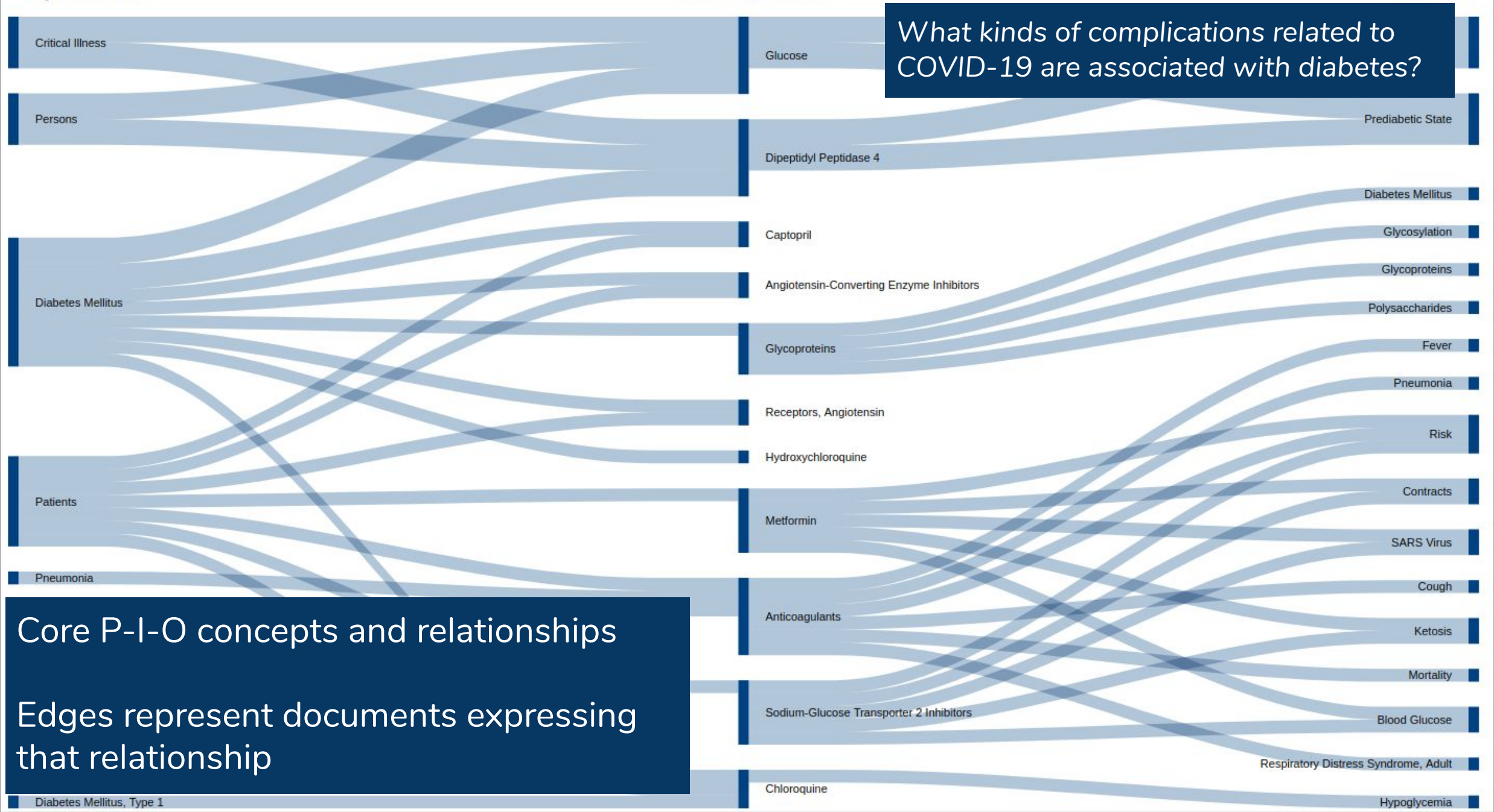
Structured search

Simple search

Population

Intervention

Outcome



Population

Intervention

Outcome



What kinds of complications related to COVID-19 are associated with diabetes?

Diabetes Mellitus → Chloroquine

Related Literature:

Chloroquine and hydroxychloroquine in the treatment of COVID-19 with or without diabetes: A systematic search and a narrative review with a special reference to India and other developing countries

Authors: Singh, Awadhesh Kumar; Singh, Akriti; Shaikh, Altamash; Singh, Ritu; Misra, Anoop
Journal: Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 2020

Abstract Background and aims No drugs are currently approved for Coronavirus Disease-2019 (COVID-19), although some have been tried. In view of recent studies and discussion on chloroquine and hydroxychloroquine (HCQ), we aimed to review existing literature and relevant websites regarding these drugs...

Core P-I-O concepts and relationships
 Edges connect to documents expressing that relationship



Selected Topic: 4

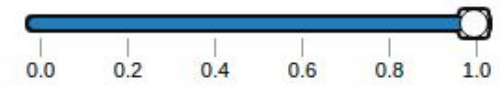
Previous Topic

Next Topic

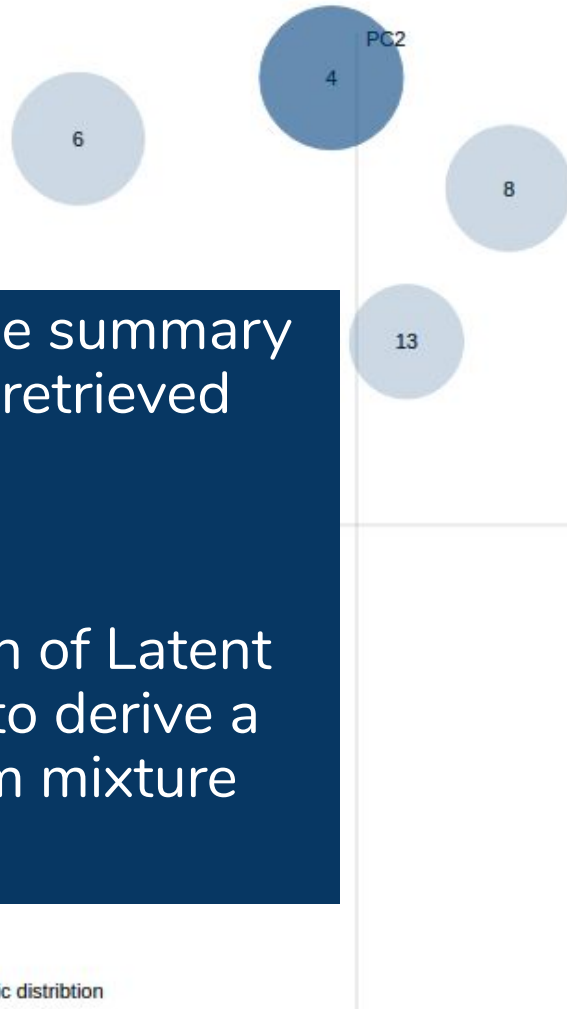
Clear Topic

Slide to adjust relevance metric:⁽²⁾

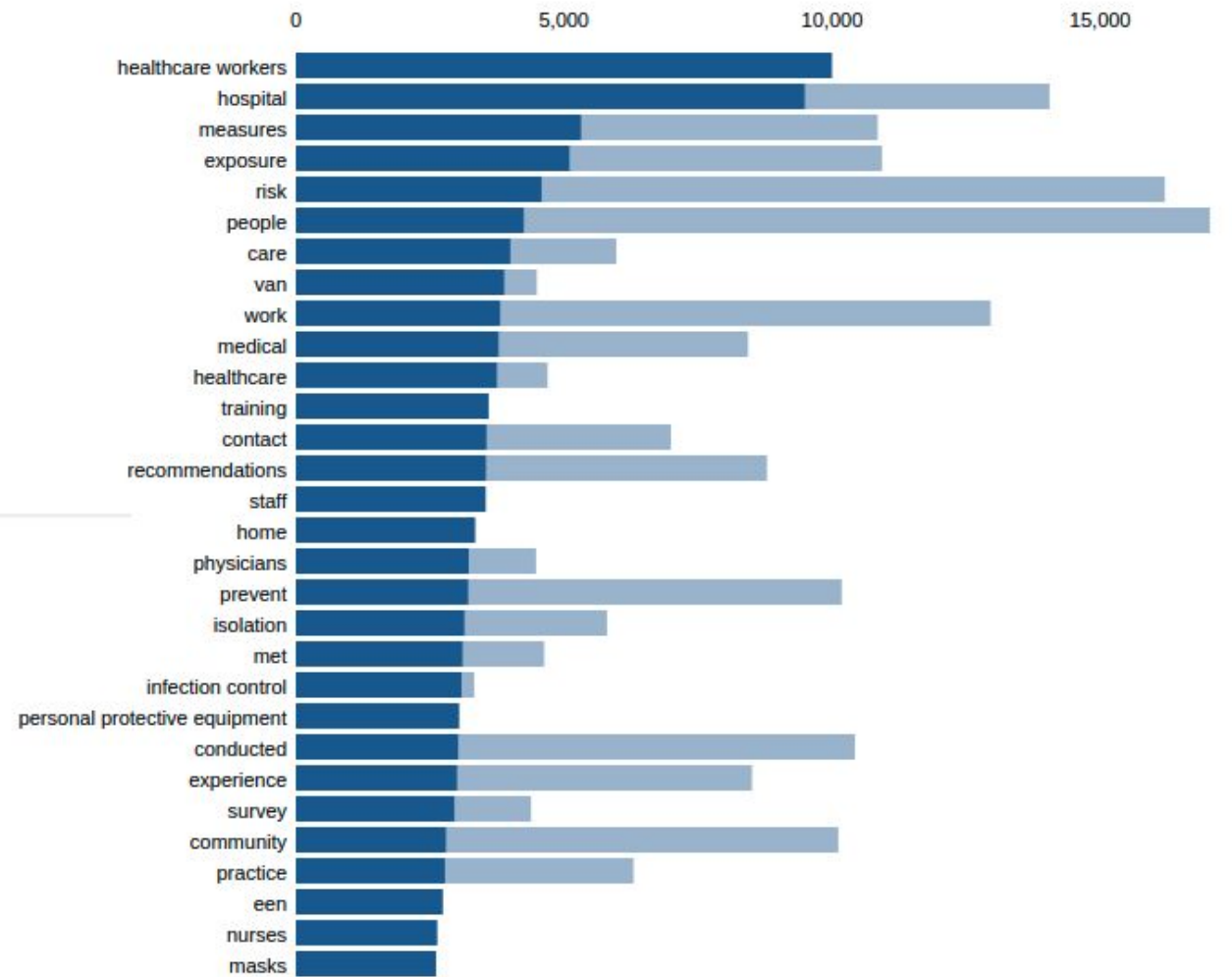
$\lambda = 1$



Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 4 (6.7% of tokens)



Overall term frequency
 Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
 2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

Marginal topic distribution



19

Topic models provide summary of salient themes in retrieved collection

Based on application of Latent Dirichlet Allocation to derive a document/topic/term mixture model

Term Clouds

Current Article:

Prevention and management of COVID-19 among patients with diabetes: an appraisal of the literature.



respiratory rate
burden problem mucosal sub described
countries enter septic europe
additional shock writing enter septic united states
upper respiratory tract unknown later state estimated

Current Article:

Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: the CORONADO study



retrospective meta-analysis well
affected prevalence a factor icu severe out united states
learning high prevalence net
list pneumonia network triple italy
chinese severity

Current Article:

Cardiovascular comorbidities and complications associated with coronavirus disease 2019

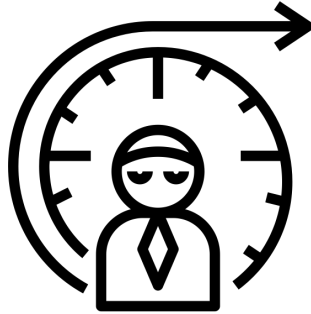


severity hand underlying share affected affect involved
affected patients cardiovascular system
global lead rapid economy
ongoing cardiovascular field countries
hypertension mortality lung

Concept clouds allow for quick overview of salient concepts in a document

Based on UMLS concepts

Weighting of terms based on relative distinctiveness

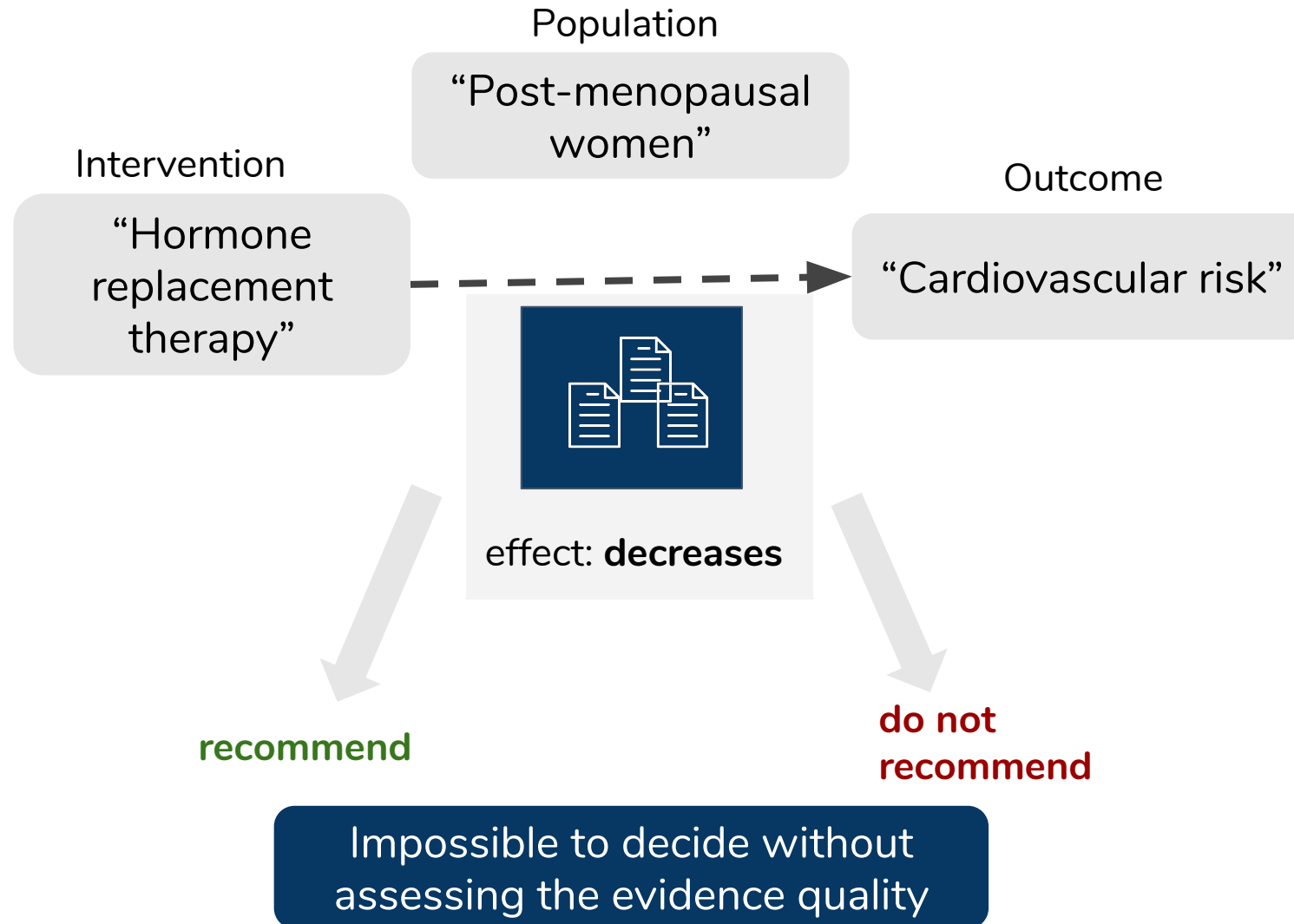


- Refine structured search and PICO visualisation
- Recommend articles from similar topic distributions
- Evaluate with an existing information-retrieval test collection
- Carry out a user study to better understand exploration and design

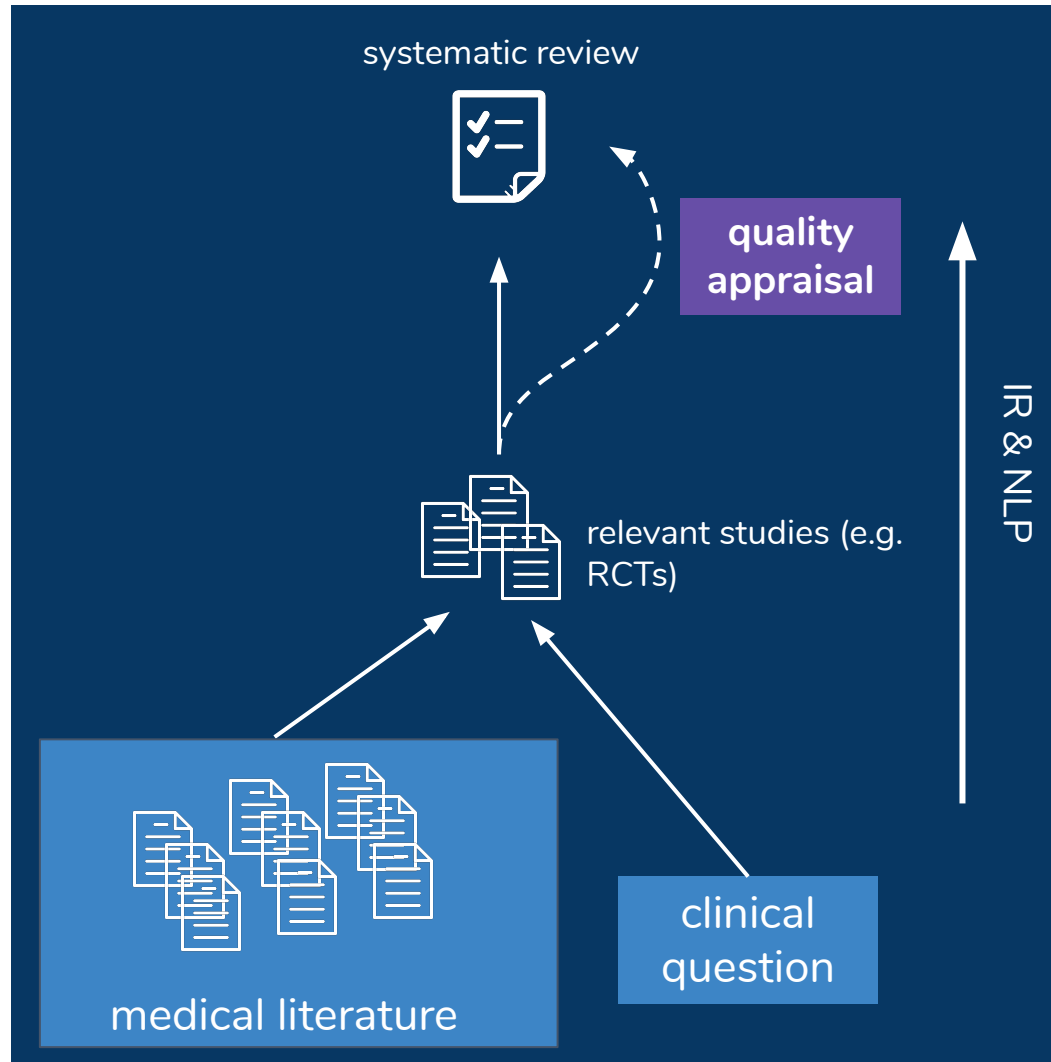
Automated quality appraisal of medical evidence



Quality assessment in evidence synthesis



Quality assessment in evidence synthesis



What confidence should we have in the estimate (based on flaws in trial design, conduct, analysis or reporting)?

Systematic reviews are often out of date:

- Extremely labour intensive (typically take 1-2 years)
- High and increasing publication rate of primary studies

Can we expedite the reviewing process with ML?

- So far, strong focus on search and screening
- In quality appraisal, work limited to risk-of-bias detection for individual studies,

e.g.



Constructing a dataset for automated quality appraisal



~8,000
systematic reviews

extract data related to quality
appraisal from summaries
of findings

14,000
outcomes rated
for quality (with
justification)

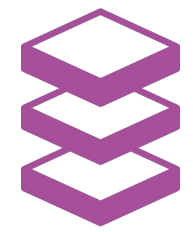
Examples of Cochrane Database of Systematic Reviews abstracts:

- Interventions for restoring patency of catheter lumens** (Version published: 18 April 2015)
- Pharmacological interventions for cognitive decline in people with Down syndrome** (Version published: 29 October 2015)
- Mediterranean-style diet for the primary and secondary prevention of cardiovascular disease** (Version published: 13 March 2019)
- Wheat flour fortified with iron alone compared to unfortified wheat flour for reducing anaemia and improving iron status in populations** (Version published: 13 April 2015)



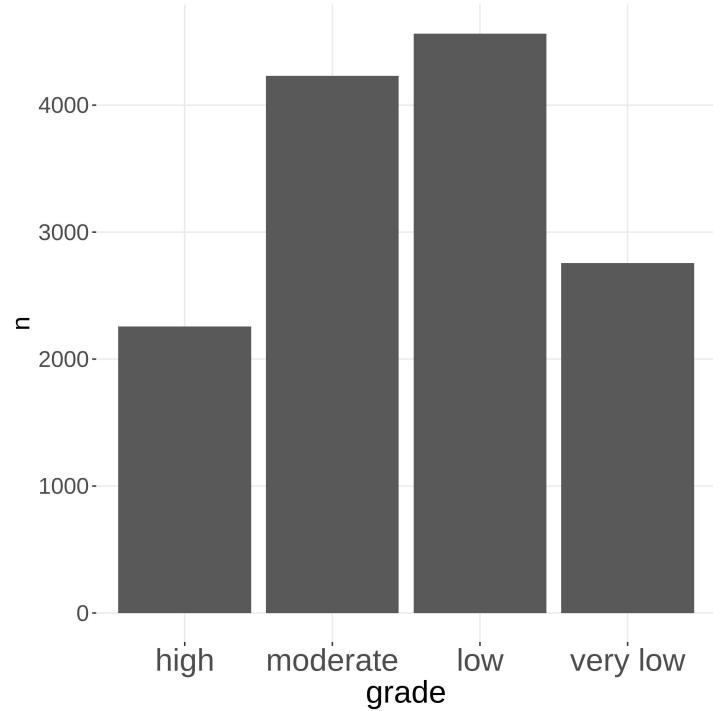
Summary of findings for the main comparison. Comparison 1: donepezil versus placebo

Outcomes	Absolute effects* (95% CI)	Relative effect (95% CI)	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
Cognitive abilities (Severe Impairment Battery; SIB)	The mean change in cognitive abilities in the intervention groups was 0.52 points higher (0.27 lower to 1.31 higher)	-	165 (3 RCTs)	Low ¹	-
Behavioural problems (various scales)	The mean change in behavioural problems the intervention groups was 0.42 points higher (0.06 lower to 0.89 higher)	-	157 (3 RCTs)	Low ²	-
Adverse events	Study population	OR 0.32 (0.16 to 0.62)	192 (4 RCTs)	Low ¹	-



Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework

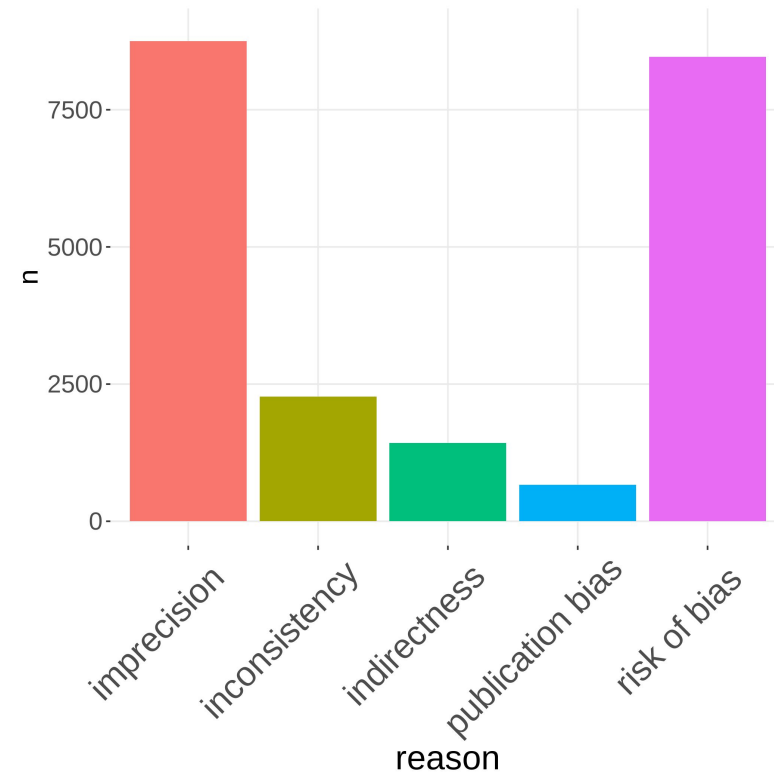


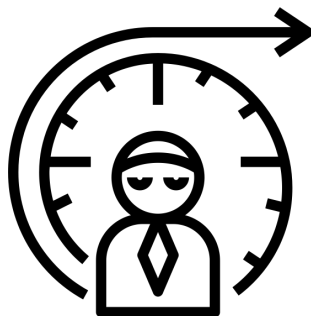


Medical evidence is of high quality only $\frac{1}{6}$ of the time

Half of the time, evidence is of (very) low quality

When medical evidence is downgraded, it is because of...



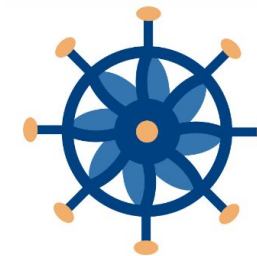


- Neural models for text classification/regression with extension to support numerical and categorical feature types (e.g. fine-tuning pre-trained language models for scientific texts)
- Provide explainable quality ratings: pointers to study flaws when downgrading
- Semi-automatically derived quality ratings for supervision: How subjective are they?

Summary

Evidence exploration

- System for navigating COVID-19-related publications
- Leveraging IR and NLP techniques for detecting concepts, PICO categories and topics



COVID-SEE
Scientific Evidence Explorer

Quality assessment of medical evidence

- New dataset from Cochrane Systematic Reviews
- A broad-scope quality rating task based on GRADE