The University Of Queensland AUSTRALIA | CREATE CHANGE

Improving Systematic Review Literature Search with Information Retrieval

Harry Scells
h.scells@uq.net.au
ITEE, University of Queensland
St Lucia, QLD, Australia

Research developed in collaboration with CSRO eHealth Centre, Bond Institute of Evidence-Based Healthcare

ie lab

Acknowledgements

Thank you to Neusoft and
The University of Queensland for
Sponsoring my Travels

Systematic Reviews (SR)

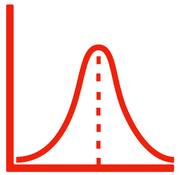
- Similar to a literature review
  - Synthesis and appraisal of **all relevant documents** for a particular **research question**
  - Aim to be **unbiased and comprehensive**
- Must adhere to strict guidelines and protocol
  - Ensures the review could be **reproduced**

Use of Systematic Reviews

- In the **medical domain**, systematic reviews:

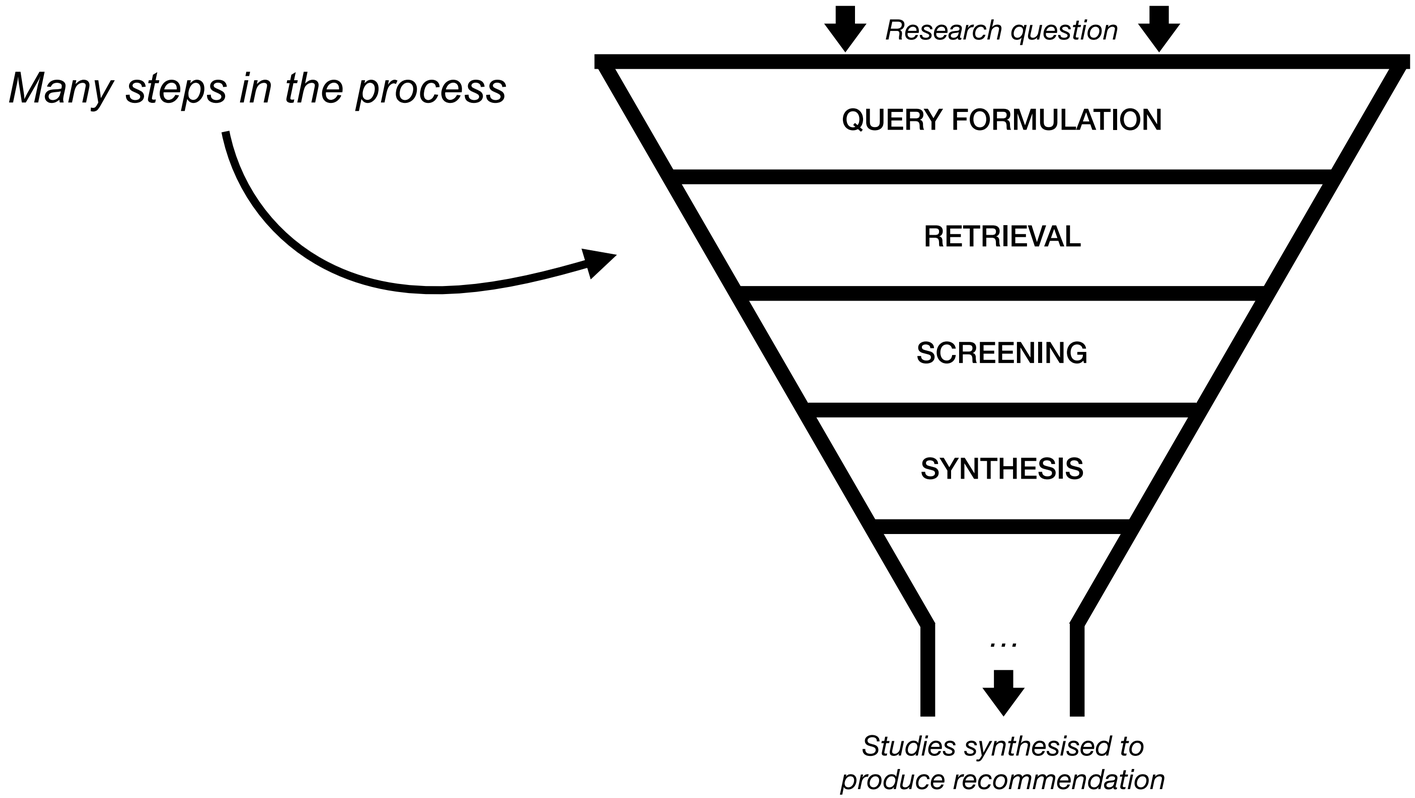
  - **Guide** clinical decisions
    - What actions clinicians should take to treat patients

  - **Inform** institutional practice and policy
    - e.g. Banning smoking in public areas in UK

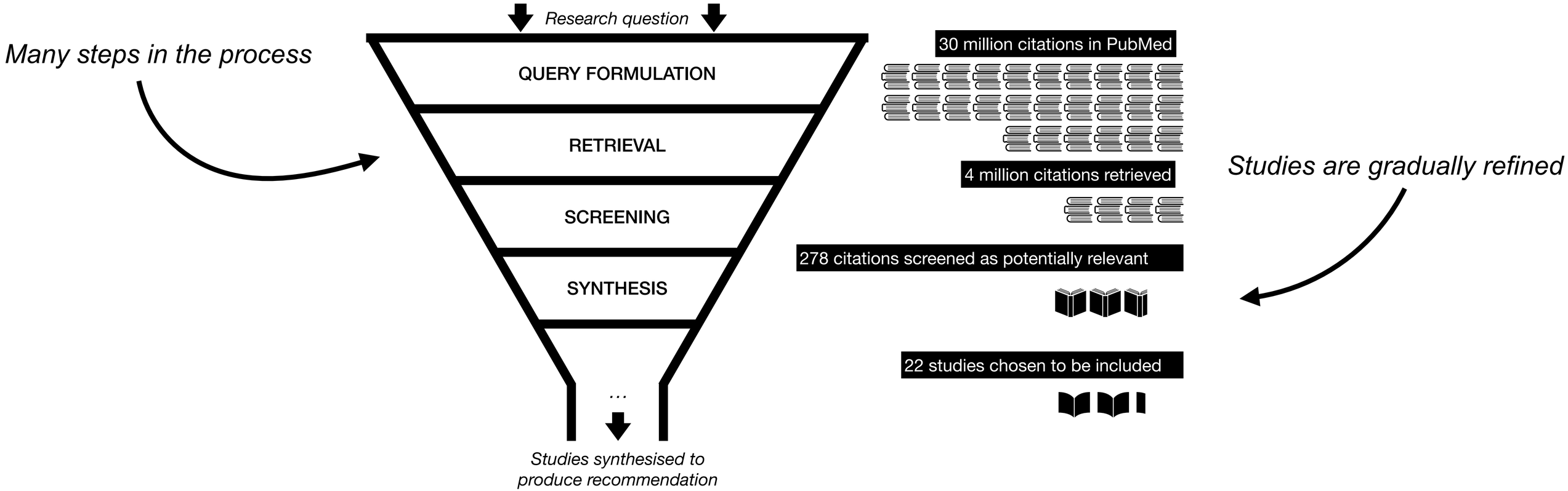


  - **Provide** evidence through comprehensive literature review
    - Cornerstone of evidence based medicine

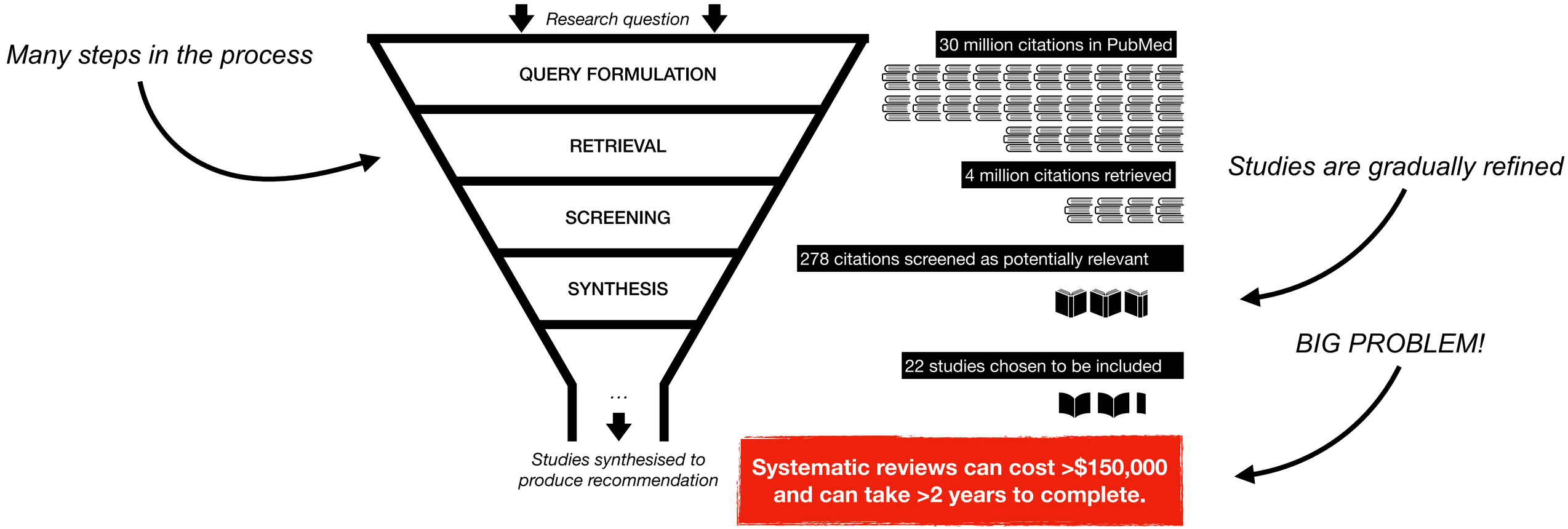
Systematic Review Creation Pipeline



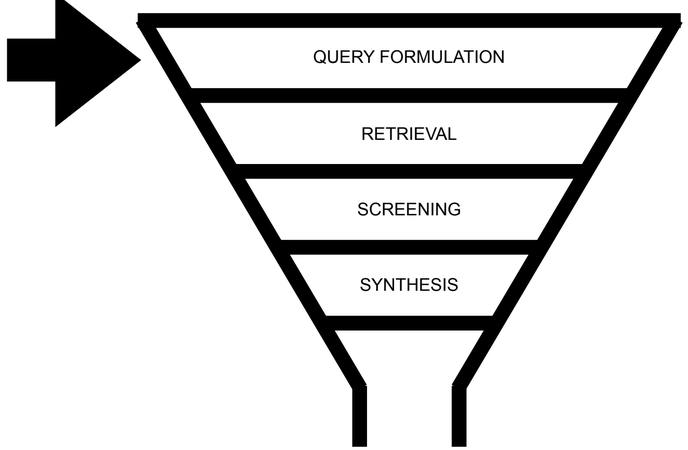
Systematic Review Creation Pipeline



Systematic Review Creation Pipeline



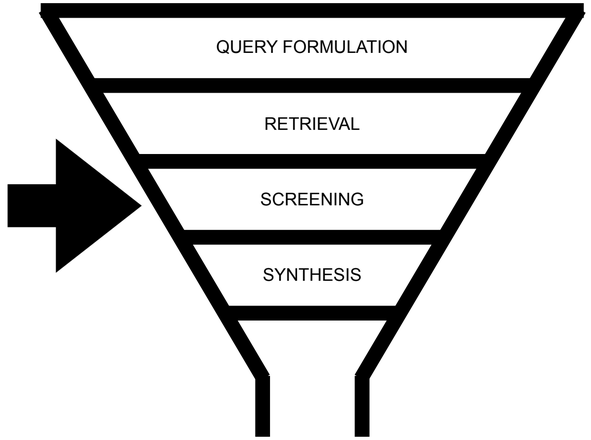
Query Formulation



- First step in the SR creation process is developing a query:
  - Information experts (i.e., librarians) usually formulate queries
  - Currently, multiple ways to go about this [1,2]
- Query is submitted to a medical database (e.g., PubMed)
  - Used to retrieve literature used for synthesis

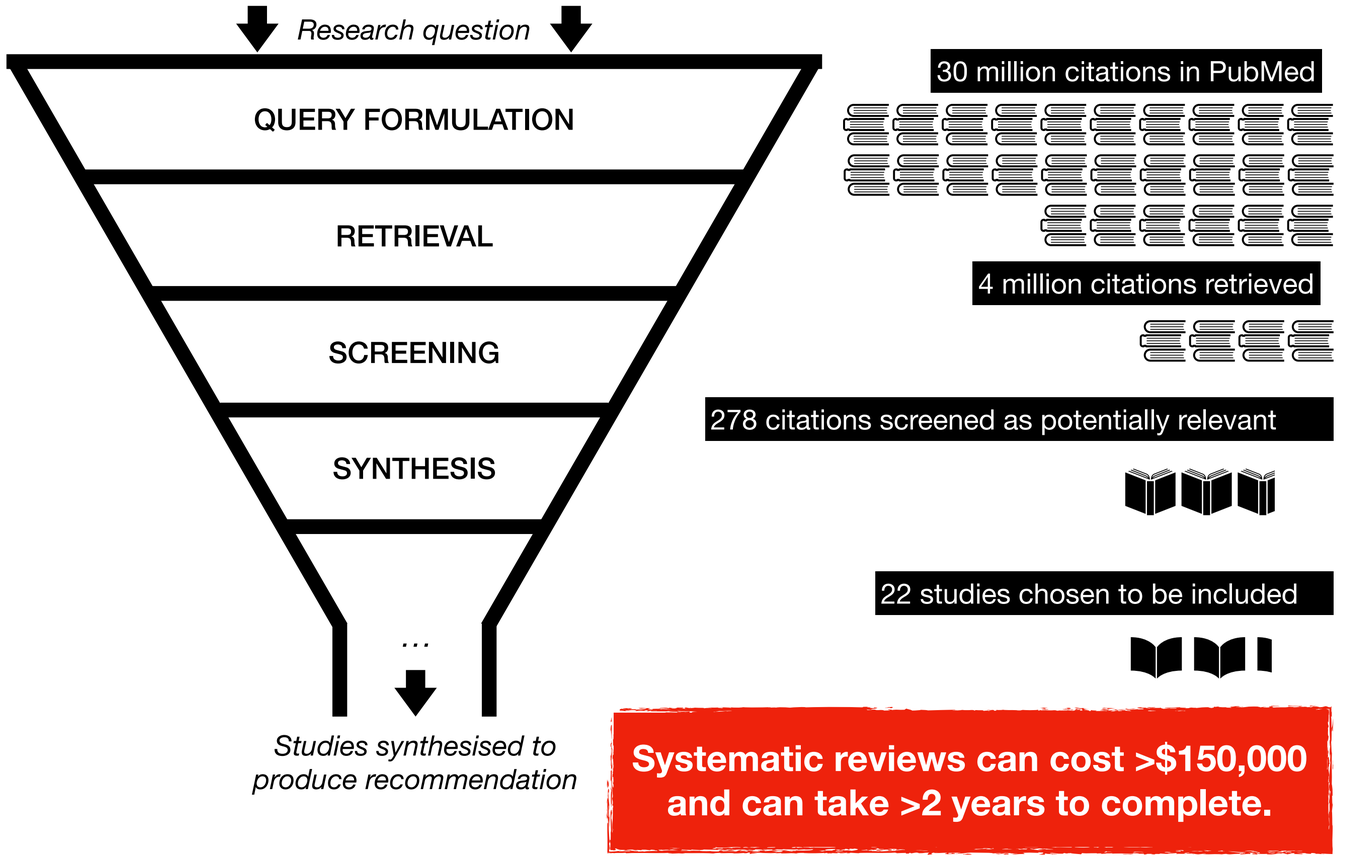
[1] Elke Hausner and Siw Waffenschmidt and Thomas Kaiser and Michael Simon. 2012. Routine development of objectively derived search strategies. Systematic reviews.
[2] Justin Clark. 2013. Systematic Reviewing. Methods of Clinical Epidemiology.

Screening

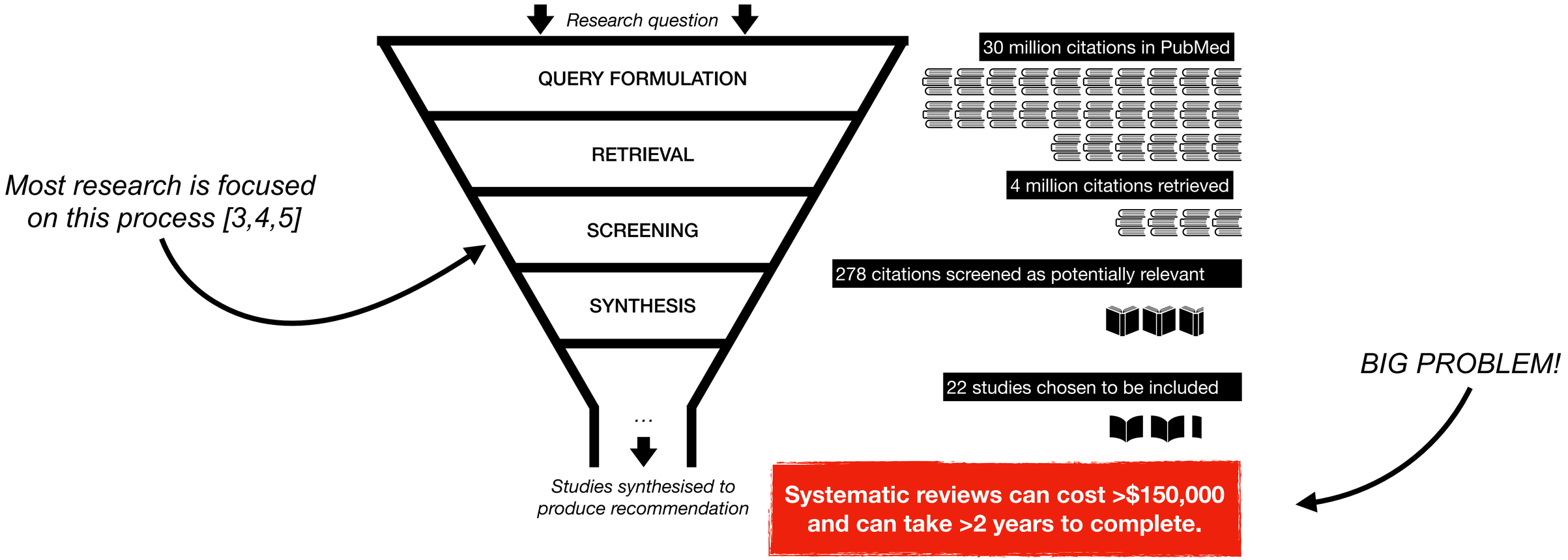


- Next step after query formulation is screening:
  - **All** of the studies that were retrieved by the query are screened
  - Studies that match inclusion criteria defined in the protocol go on to the next step in the process

Systematic Review Creation Pipeline

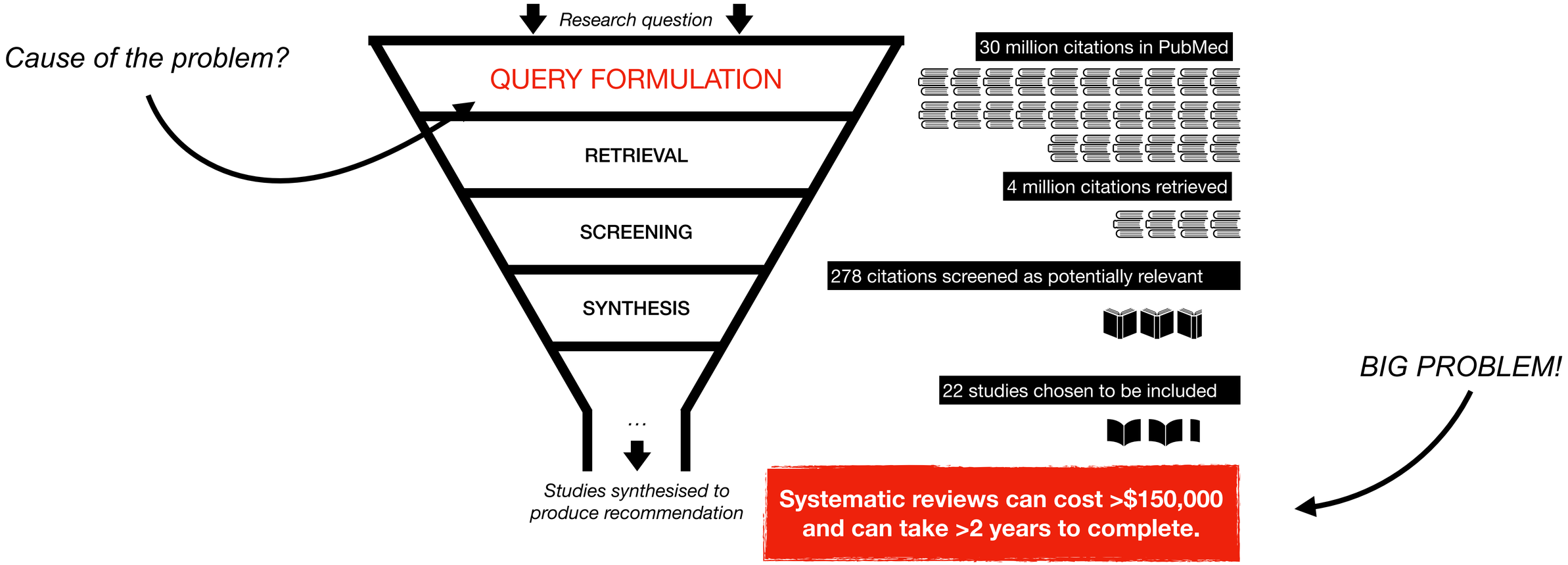


Systematic Review Creation Pipeline

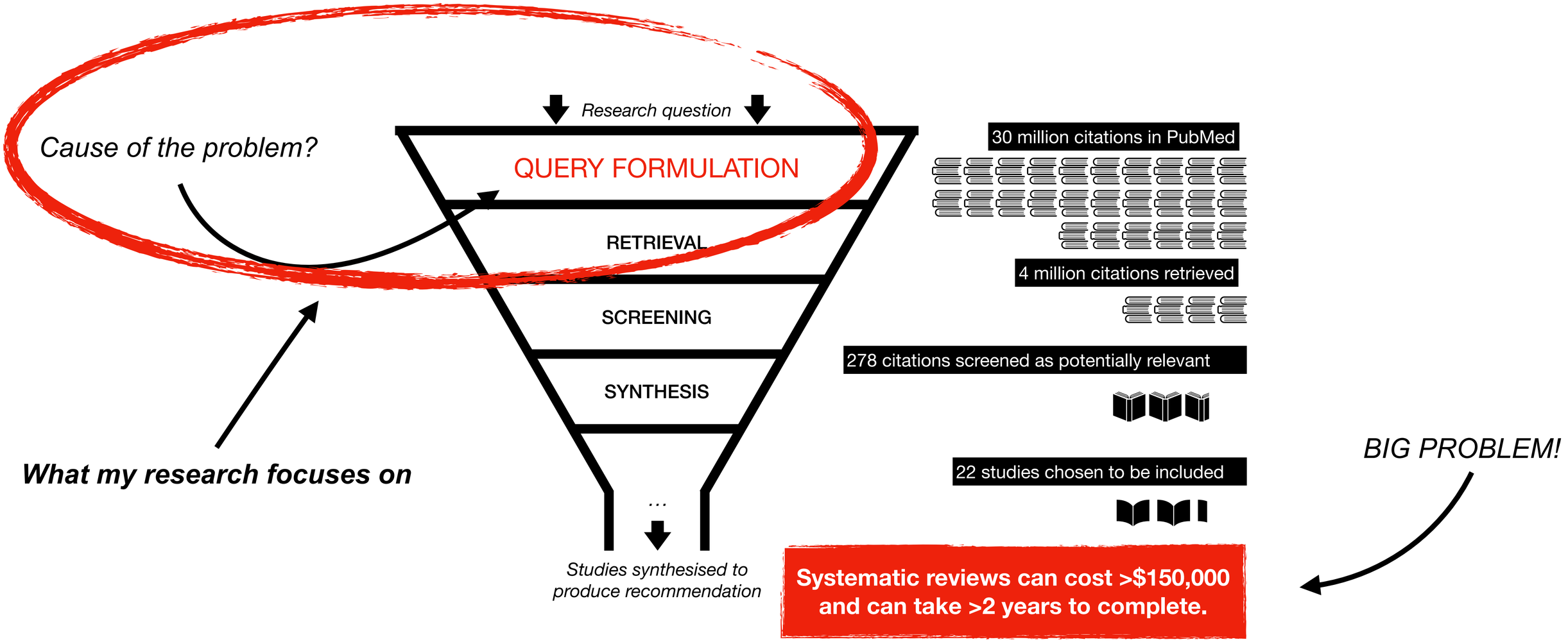


[3] Miwa, M., Thomas, J., O'Mara-Eves, A. and Ananiadou, S., 2014. Reducing systematic review workload through certainty-based screening. Journal of biomedical informatics,.
[4] Olorisade, B.K., de Quincey, E., Brereton, P. and Andras, P., 2016, June. A critical analysis of studies that address the use of text mining for citation screening in systematic reviews. In Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering (p. 14). ACM.
[5] Wallace, B.C., Trikalinos, T.A., Lau, J., Brodley, C. and Schmid, C.H., 2010. Semi-automated screening of biomedical citations for systematic reviews. BMC bioinformatics, 11(1)

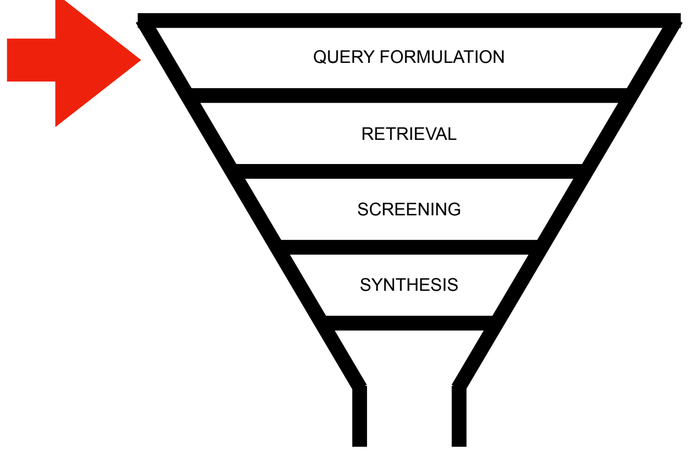
Systematic Review Creation Pipeline



Systematic Review Creation Pipeline



Improving SR Literature Search



- Query formulation impacts all downstream activities of a SR
  - Quality of the SR is **ultimately decided by the query**
- What if instead of automating the screening process?
  - We **start screening** with **better queries** to begin with
- My research is novel:
  - Directly tackles problem at the source

Addressing the Problem

Systematic reviews can cost >$150,000 and can take >2 years to complete.

- **RQ1**: Is it possible to formulate Boolean queries that are more effective than those originally used within search strategies of systematic reviews?

- **RQ2**: If the answer to RQ1 is positive, then: Can alternative, more effective Boolean queries, generated from the original systematic review queries, be automatically selected?

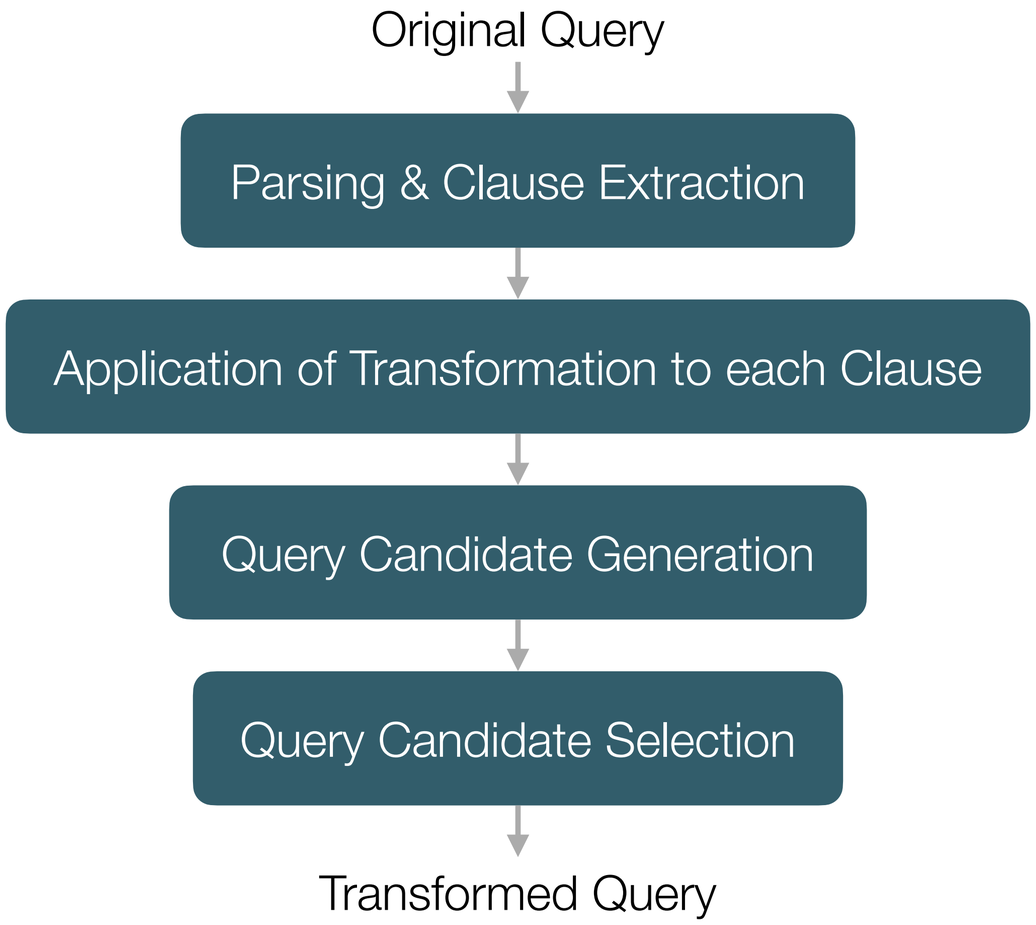
“Better Query”

- Better Query = Performs better for evaluation measures
- In this work the measures we use are:
- Precision, Recall,
- Fβ0.5, Fβ1, Fβ3,
- WSS

WSS = “work saved over sampling”; compares # of non relevant that have not been retrieved, those that have been retrieved, and recall

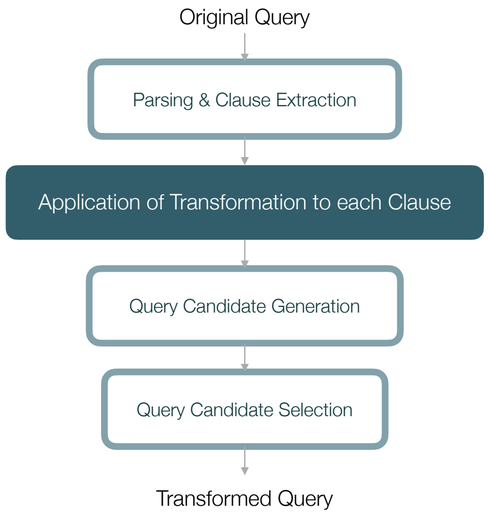
$$WSS = \frac{N - \texttt{NumRet}}{N} - (1 - \texttt{Recall})$$

Methodology


Original Query
Parsing & Clause Extraction
Application of Transformation to each Clause
Query Candidate Generation
Query Candidate Selection
Transformed Query

Transforming Boolean queries

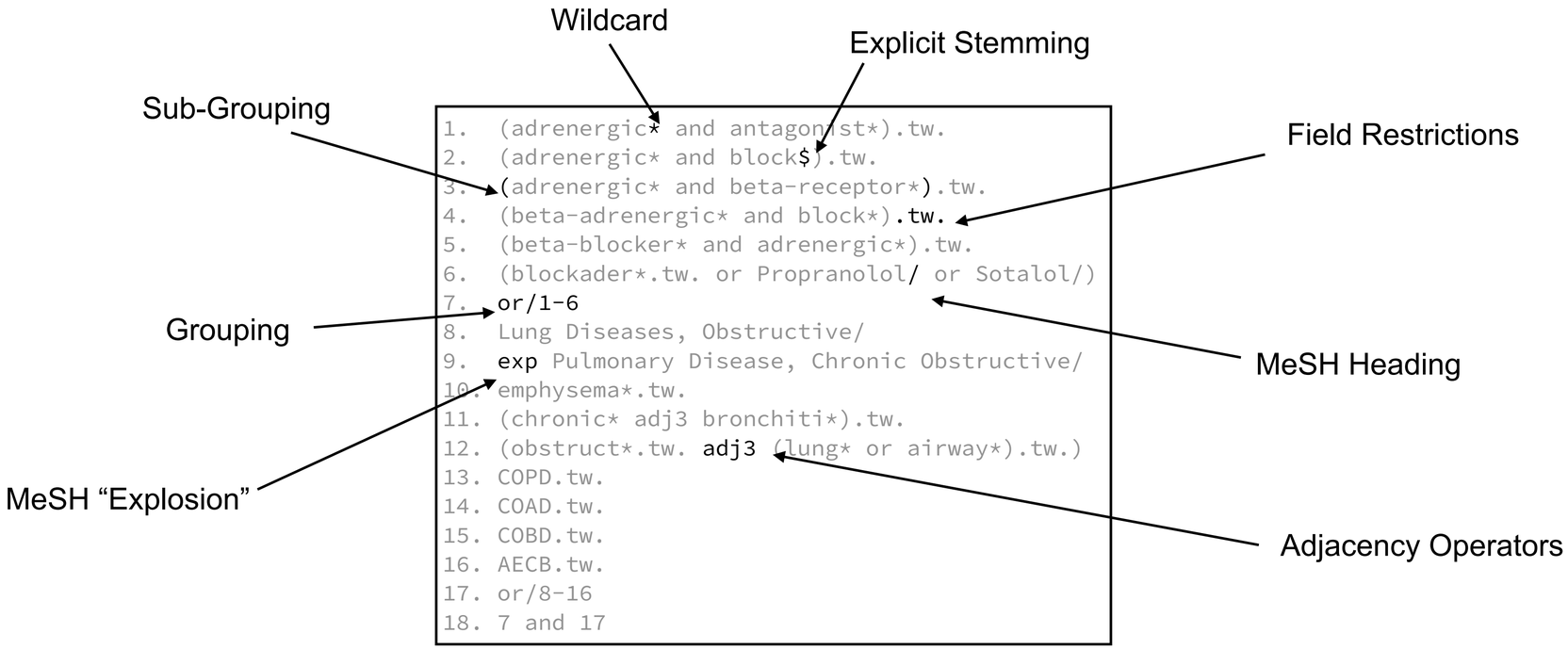
- What ways can Boolean queries be modified?
- Syntactic modifications:
  - Logical Operator Replacement
  - Adjacency Range / Adjacency Replacement
  - MeSH Explosion
  - Field Restrictions



What does a Boolean query look like?

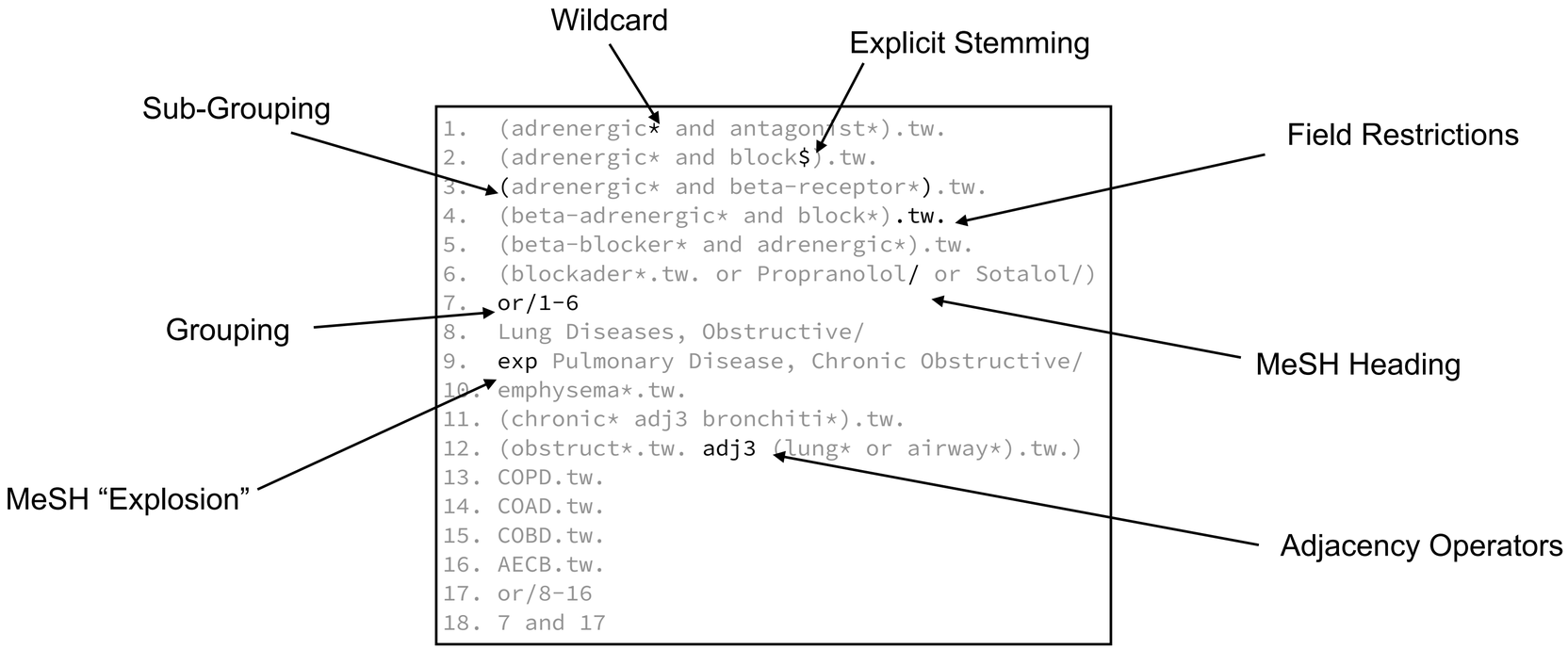
```
1.  (adrenergic* and antagonist*).tw.
2.  (adrenergic* and block$).tw.
3.  (adrenergic* and beta-receptor*).tw.
4.  (beta-adrenergic* and block*).tw.
5.  (beta-blocker* and adrenergic*).tw.
6.  (blockader*.tw. or Propranolol/ or Sotalol/)
7.  or/1-6
8.  Lung Diseases, Obstructive/
9.  exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* adj3 bronchiti*).tw.
12. (obstruct*.tw. adj3 (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 and 17
```

What does a Boolean query look like?



MeSH = **M**edical **S**ubject **H**eadings — an ontology of medical concepts
MeSH Explosion = Subsumption — consider all children

How do we transform Boolean queries?



…WHAT DO THESE TRANSFORMATIONS LOOK LIKE?

Logical Operator Replacement

```
1.  (adrenergic* and antagonist*).tw.
2.  (adrenergic* and block$).tw.
3.  (adrenergic* and beta-receptor*).tw.
4.  (beta-adrenergic* and block*).tw.
5.  (beta-blocker* and adrenergic*).tw.
6.  (blockader*.tw. or Propranolol/ or Sotalol/)
7.  or/1-6
8.  Lung Diseases, Obstructive/
9.  exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* adj3 bronchiti*).tw.
12. (obstruct*.tw. adj3 (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 and 17
```

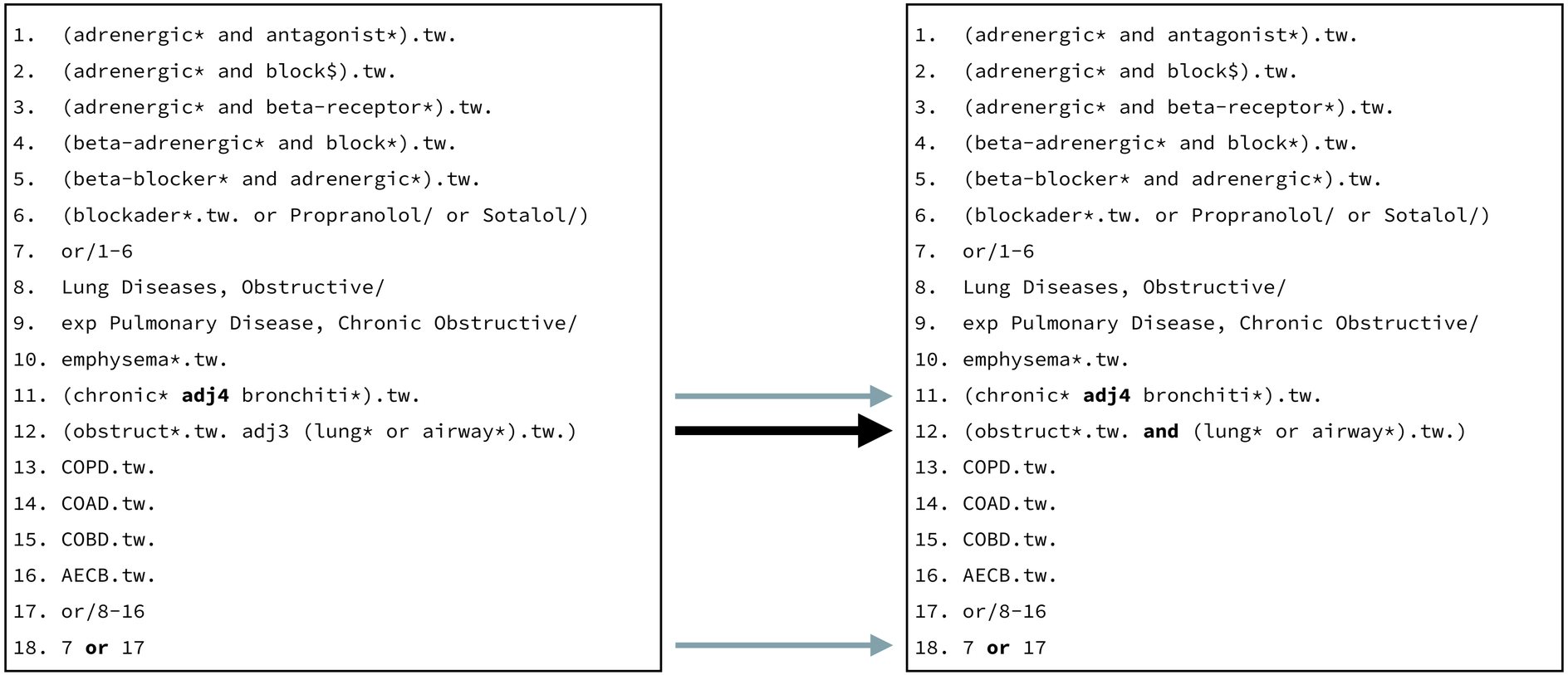
```
1.  (adrenergic* and antagonist*).tw.
2.  (adrenergic* and block$).tw.
3.  (adrenergic* and beta-receptor*).tw.
4.  (beta-adrenergic* and block*).tw.
5.  (beta-blocker* and adrenergic*).tw.
6.  (blockader*.tw. or Propranolol/ or Sotalol/)
7.  or/1-6
8.  Lung Diseases, Obstructive/
9.  exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* adj3 bronchiti*).tw.
12. (obstruct*.tw. adj3 (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 or 17
```

Adjacency Range

```
1.  (adrenergic* and antagonist*).tw.
2.  (adrenergic* and block$).tw.
3.  (adrenergic* and beta-receptor*).tw.
4.  (beta-adrenergic* and block*).tw.
5.  (beta-blocker* and adrenergic*).tw.
6.  (blockader*.tw. or Propranolol/ or Sotalol/)
7.  or/1-6
8.  Lung Diseases, Obstructive/
9.  exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* adj3 bronchiti*).tw.
12. (obstruct*.tw. adj3 (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 or 17
```

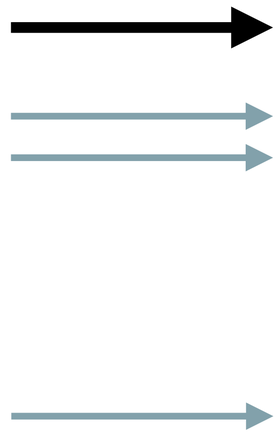
```
1.  (adrenergic* and antagonist*).tw.
2.  (adrenergic* and block$).tw.
3.  (adrenergic* and beta-receptor*).tw.
4.  (beta-adrenergic* and block*).tw.
5.  (beta-blocker* and adrenergic*).tw.
6.  (blockader*.tw. or Propranolol/ or Sotalol/)
7.  or/1-6
8.  Lung Diseases, Obstructive/
9.  exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* adj4 bronchiti*).tw.
12. (obstruct*.tw. adj3 (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 or 17
```

Adjacency Replacement



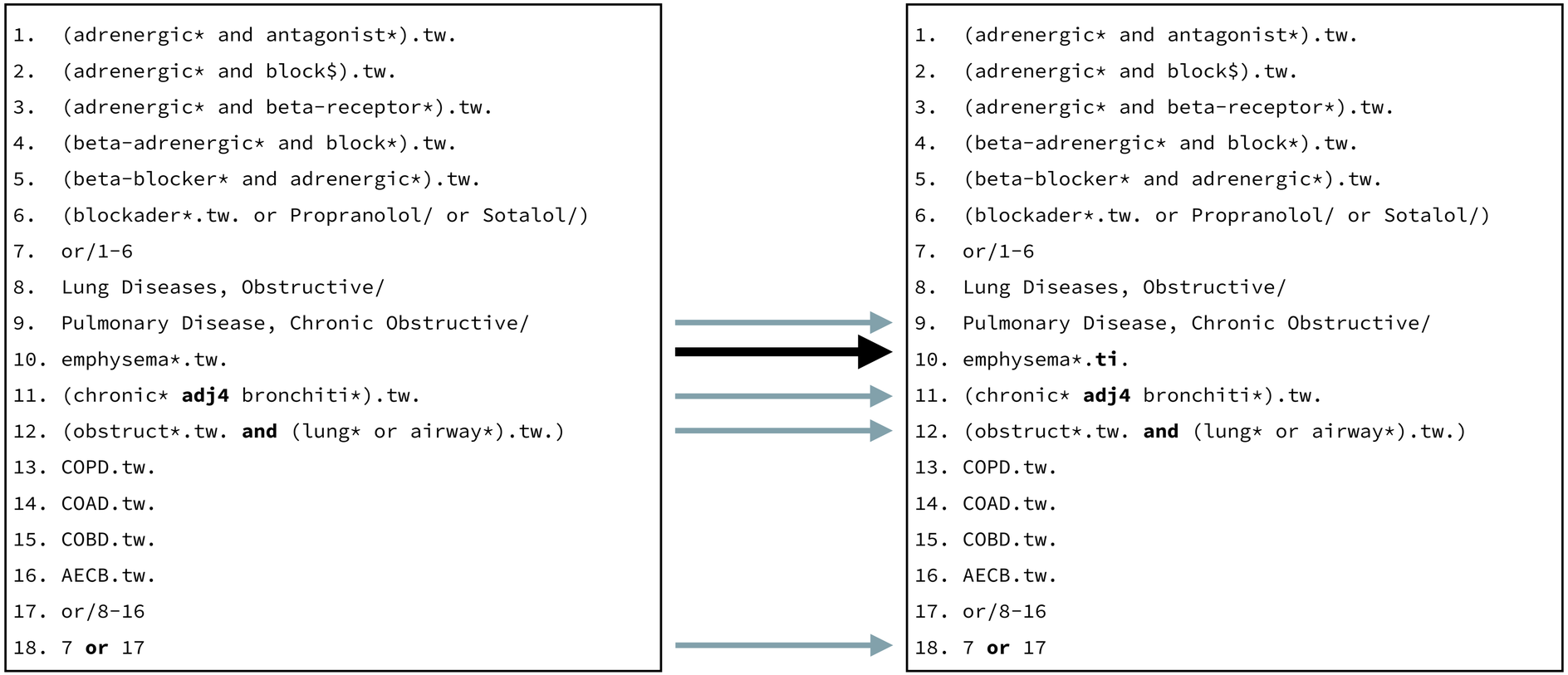
MeSH Explosion

1. (adrenergic* and antagonist*).tw.
2. (adrenergic* and block$).tw.
3. (adrenergic* and beta-receptor*).tw.
4. (beta-adrenergic* and block*).tw.
5. (beta-blocker* and adrenergic*).tw.
6. (blockader*.tw. or Propranolol/ or Sotalol/)
7. or/1-6
8. Lung Diseases, Obstructive/
9. exp Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* **adj4** bronchiti*).tw.
12. (obstruct*.tw. **and** (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 **or** 17

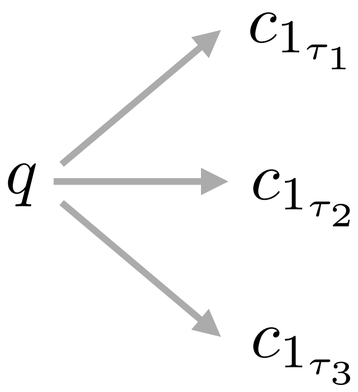


1. (adrenergic* and antagonist*).tw.
2. (adrenergic* and block$).tw.
3. (adrenergic* and beta-receptor*).tw.
4. (beta-adrenergic* and block*).tw.
5. (beta-blocker* and adrenergic*).tw.
6. (blockader*.tw. or Propranolol/ or Sotalol/)
7. or/1-6
8. Lung Diseases, Obstructive/
9. Pulmonary Disease, Chronic Obstructive/
10. emphysema*.tw.
11. (chronic* **adj4** bronchiti*).tw.
12. (obstruct*.tw. **and** (lung* or airway*).tw.)
13. COPD.tw.
14. COAD.tw.
15. COBD.tw.
16. AECB.tw.
17. or/8-16
18. 7 **or** 17

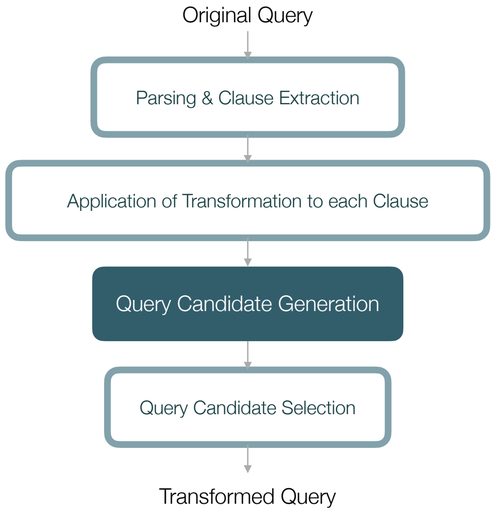
Field Restrictions



Generating Candidates



- Given an **input query**
- Generate candidates based on **transformations**
- **Lots of new queries** - good and bad
- Next step: select the best candidate(s)

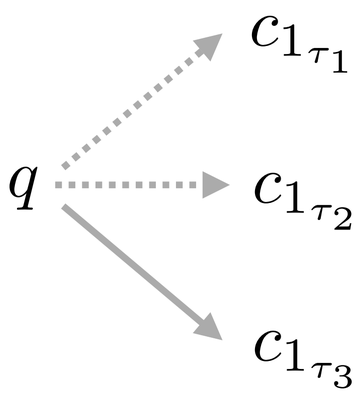


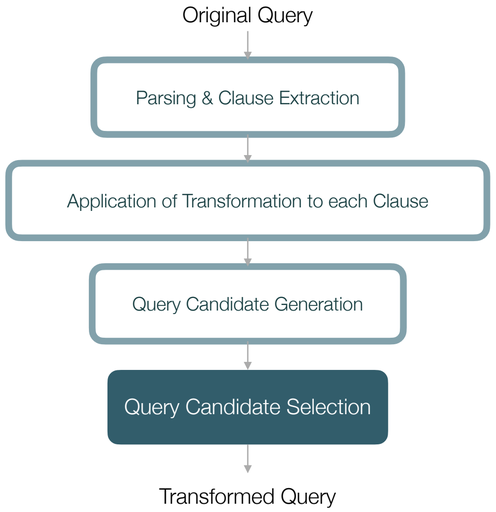
Candidate Selection

- **How** to select the “best” generated candidate(s)?
  - **Optimise evaluation measures!**
- Four candidate selection methods:

Relevance Feedback
- Greedy
- Oracle

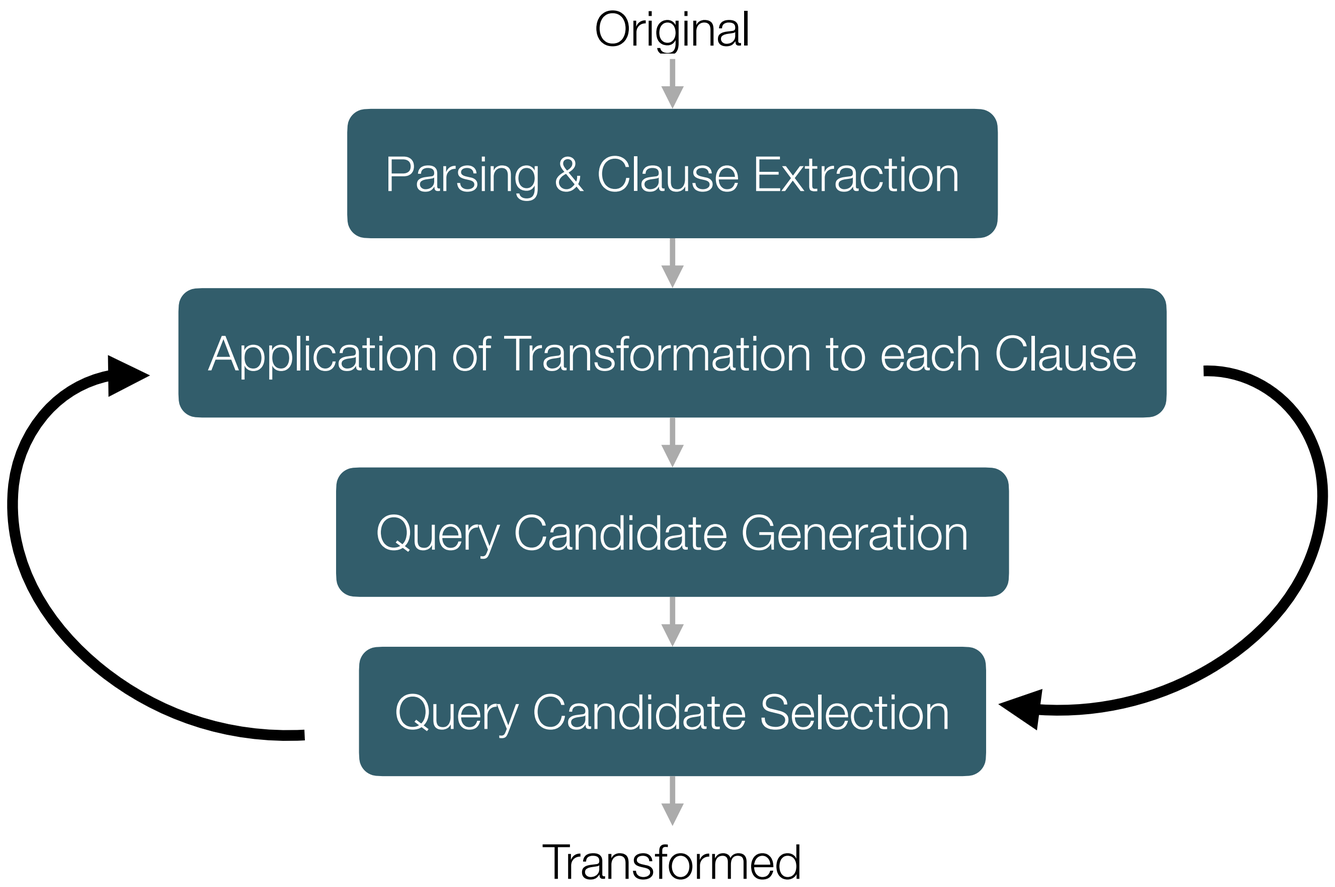
Automatic
- Classification
- Ranking





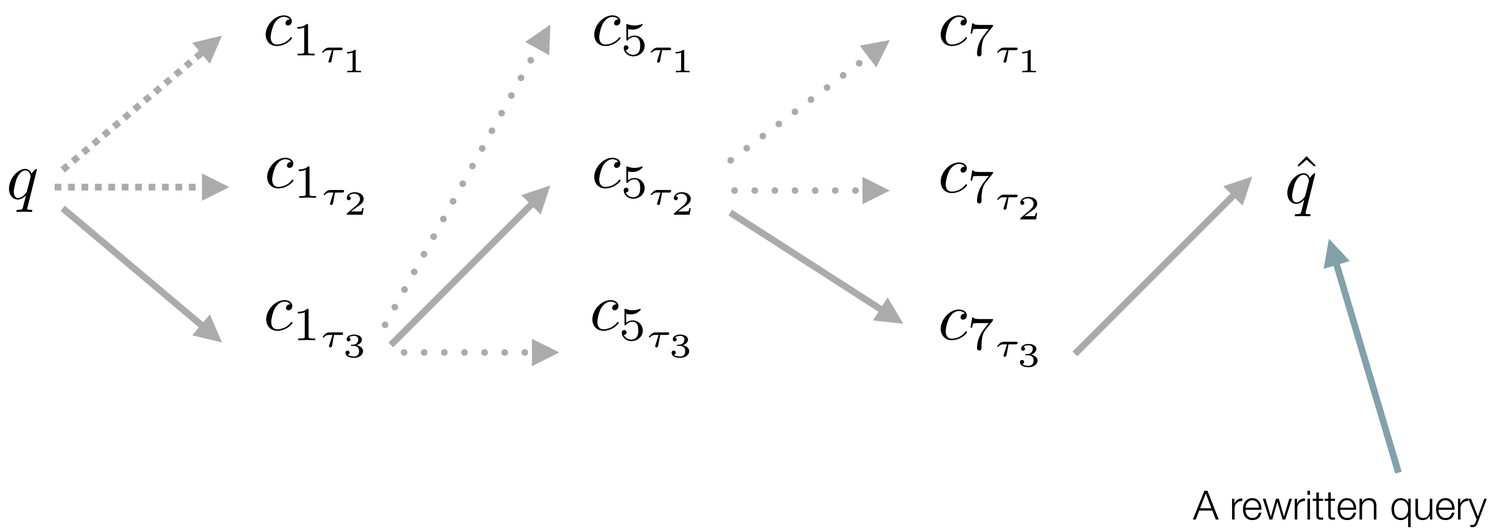
Query Chains

- **Repeated application** of transformations
- Each application generates **candidates**
- **Follow the path** of the best candidate(s) n times
  - Output: **transformed query**



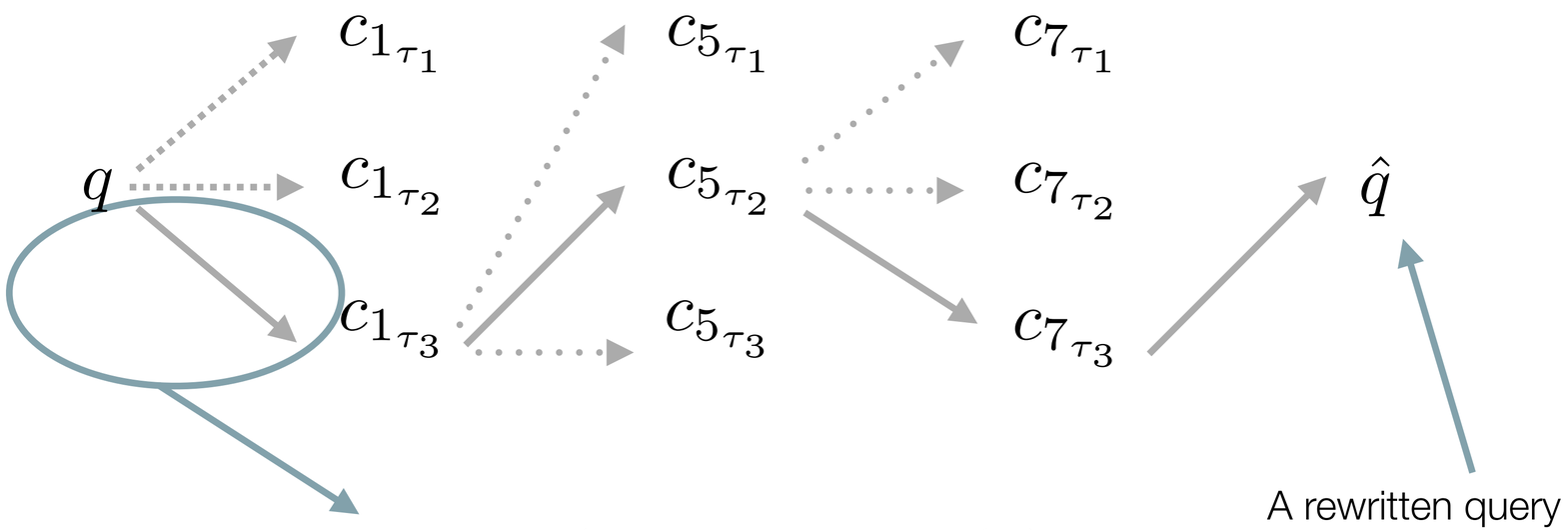
Query Transformation Chain

Repeated application of Candidate Generation and Candidate Selection
= **Query Transformation Chain**



Query Transformation Chain

Formalisation of **candidate selection function:** $f(\hat{q})$



Selection: $q^* = \underset{\hat{q} \in \hat{\mathbb{Q}}_q}{\operatorname{argmax}} f(\hat{q})$

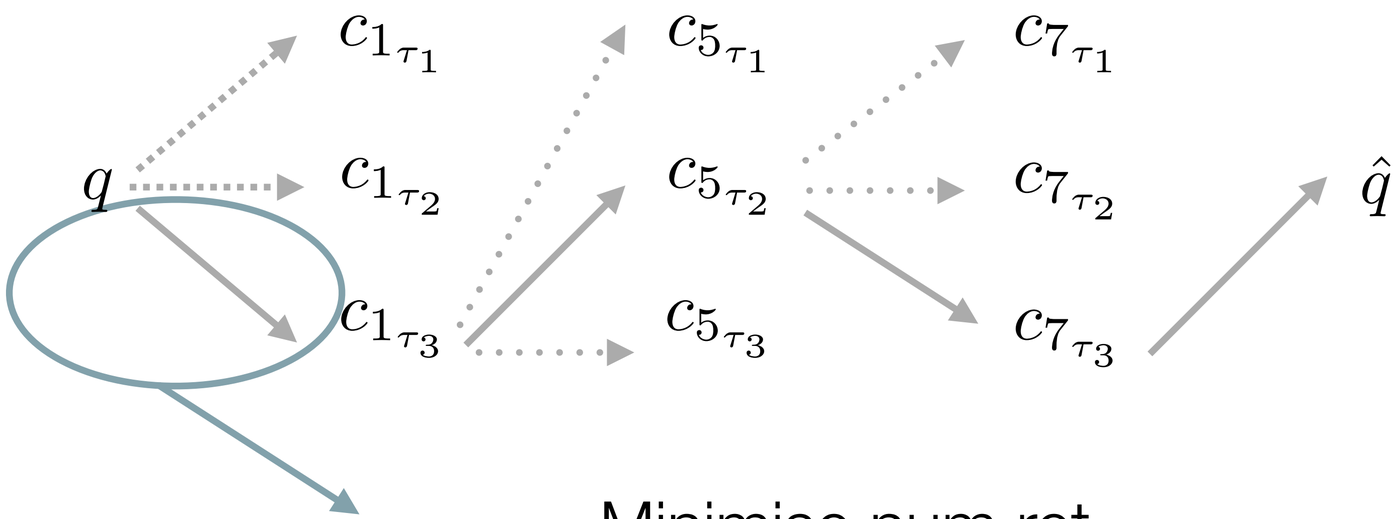
Where:

$\hat{\mathbb{Q}}_q = c1_{\tau_1}, c1_{\tau_2}, c1_{\tau_3} \ldots c1_{\tau_n}$

$\hat{q}$ = a transformed candidate query

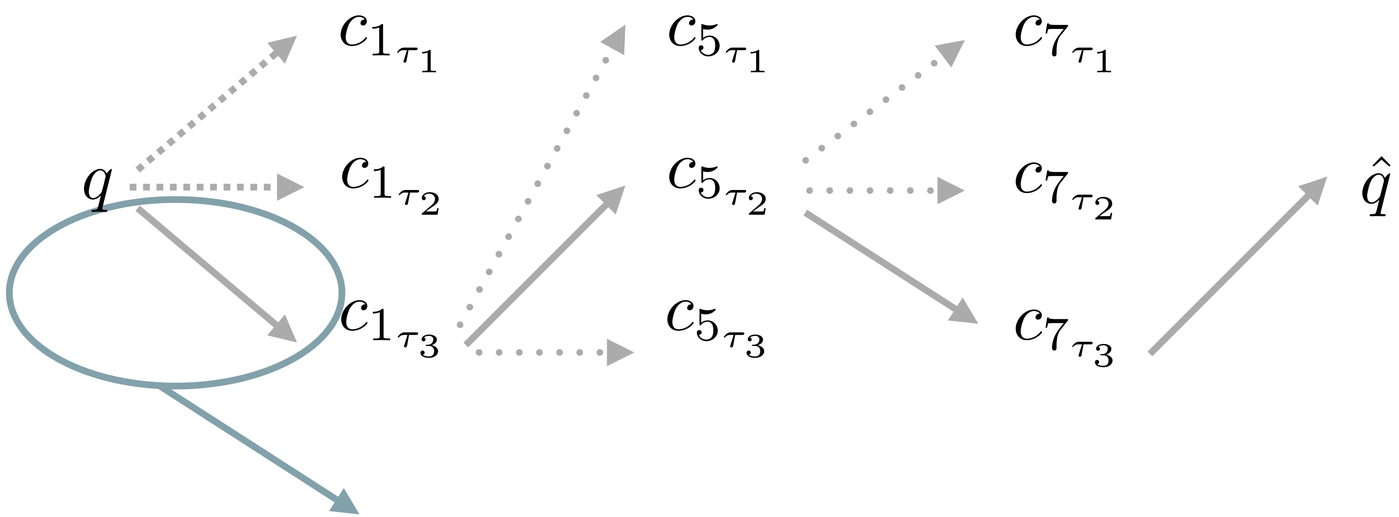
Greedy Candidate Selection

- Minimises total citations retrieved
- Maximises relevant retrieved
- Uses relevance assessments



Oracle Candidate Selection

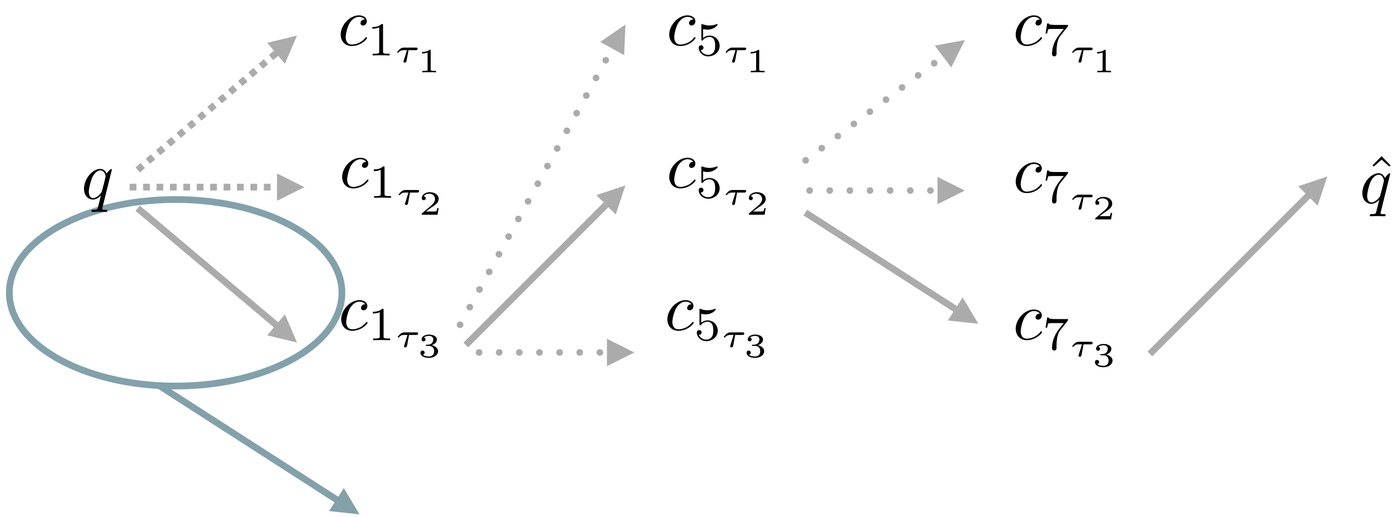
- Maximise specified eval. measure
- Uses relevance assessments



Where $E(\hat{q})$ is a specified evaluation measure

Classification Candidate Selection

- Uses SVM classifier

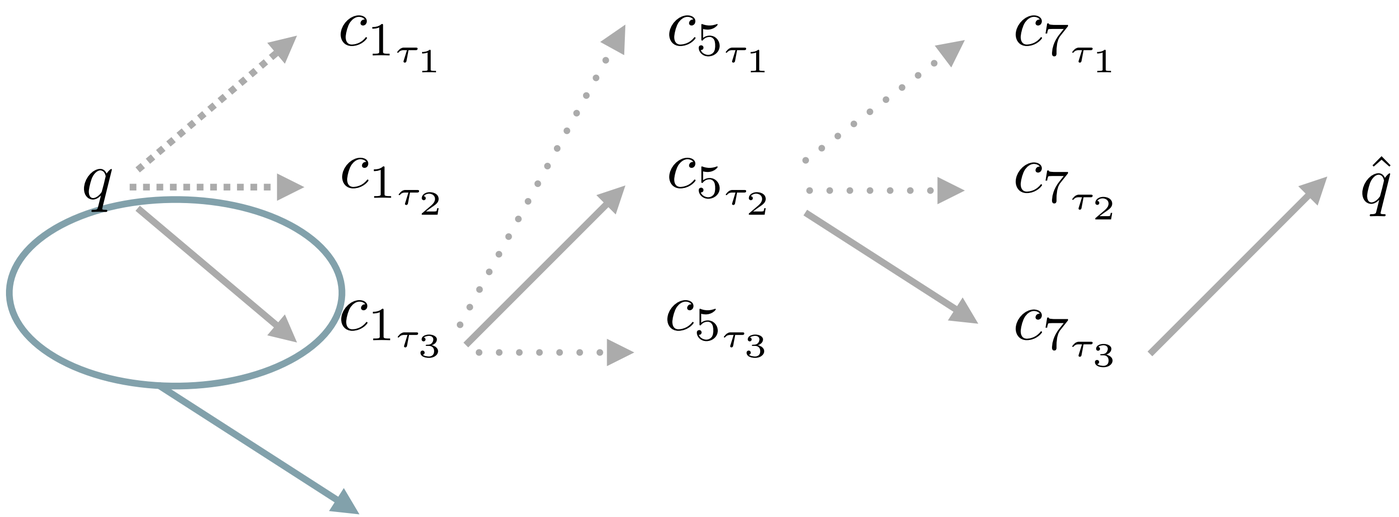


Predict if the new query improves; binary classification
(see paper[6] for details & features)

[6] Harrisen Scells and Guido Zuccon. 2018. Generating Better Queries for Systematic Reviews. The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.

Ranking Candidate Selection

- Uses Learning to Rank (SVMRank)



Pairwise learning to rank
(see paper[6] for details & features)

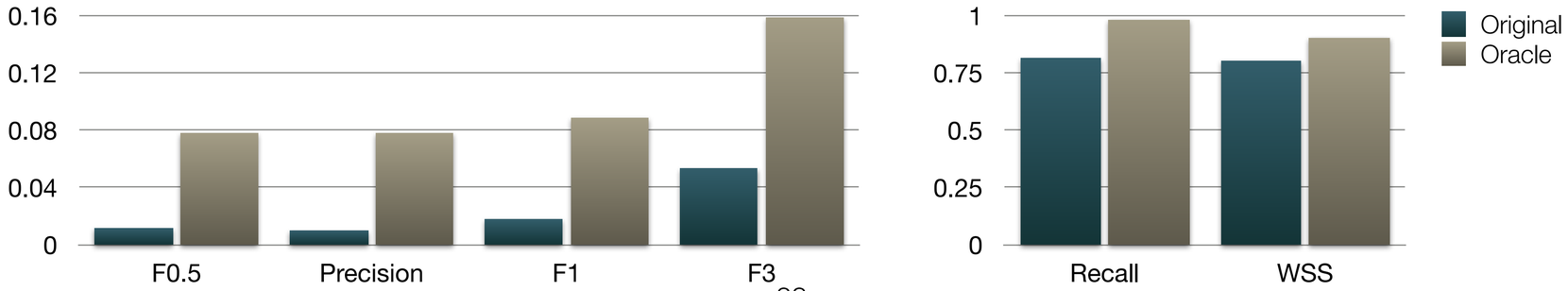
[6] Harrisen Scells and Guido Zuccon. 2018. Generating Better Queries for Systematic Reviews. The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.

Summary

Can Better Queries be Automatically Generated?

RQ1: Is it possible to formulate Boolean queries that are more effective than those originally used within search strategies of systematic reviews?

- **Better queries are generated** with these methods
- Introduced a **trade-off**
  - e.g. optimising precision degrades recall
- Syntactic transformations, **significant improvements**
  - See papers for much more detailed results!



Can Better Queries be Automatically Selected?

RQ2: If the answer to RQ1 is positive, then: Can alternative, more effective Boolean queries, generated from the original systematic review queries, be automatically selected?

- Classifier and Ranker **selected better queries**
  - Not significantly better, but ~ >100-500% improvement
- Recall-based measures **more difficult** to optimise
- Both **not as good as oracle**
  - But not significantly worse
- **Room for improvement!**

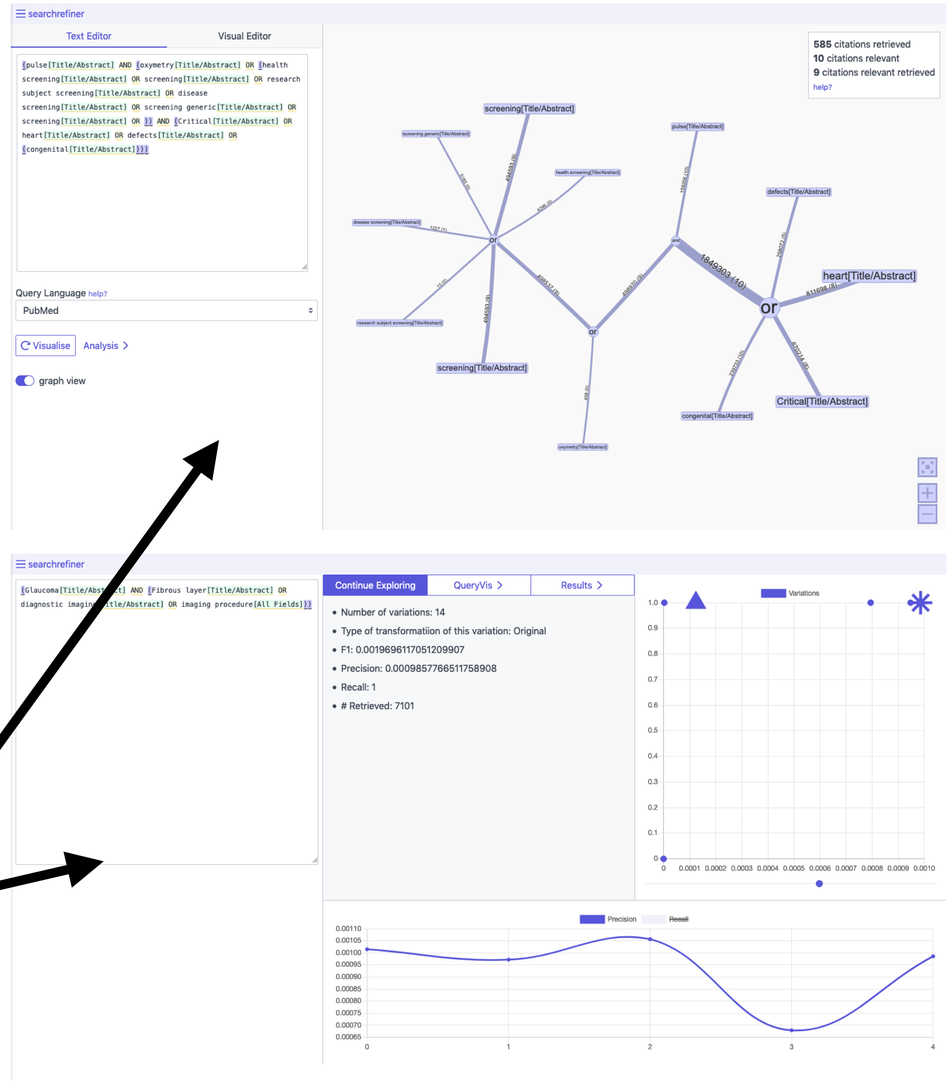
Can Better Queries be Automatically Selected?

- Follow-up study[7] found that syntactic transformations have a larger effect on query performance than semantic transformations
  - Query expansion and query reduction did not help as much as transformations like changing Boolean operators or field restrictions
  - Queries transformed using query expansion or reduction were not ranked highly by the LTR model

[7] Harrisen Scells and Guido Zuccon and Bevan Koopman. 2019. Automatic Boolean Query Refinement for Systematic Review Literature Search. Proceedings of WWW.

Conclusions & Future Work

- Takeaways:
  - Better Boolean queries are **possible**
  - These queries can be **automatically** identified
- Next steps:
  - **Better Sampling** methods (training data = expensive)
  - Is ML necessary? QPP?
  - **End goal**: integration into tools to assist reviewers



Improving Systematic Review Literature Search