

Combining Query Reformulations to Perform Rank Fusion for Precision Medicine

Giorgio Maria Di Nunzio, Stefano Marchesin, Maristella Agosti

Department of Information Engineering, University of Padua, Italy

Objectives

- 1) How different query reformulations affect the results and whether the findings obtained in previous years remain valid
- 2) Study whether combining different query reformulations based on expansion/reduction techniques prove effective in a highly specific scenario

Precision Medicine Task 1: Scientific Abstracts

The 2019 Precision Medicine task 1 focuses on the problem of identifying the most relevant articles for the *treatment*, *prevention*, and *prognosis* of the disease under the specific conditions for the given patient.

BM25 + BL + QR + RF

Search Engine:

- ElasticSearch

Concept Mapper:

- MetaMap

Indexing:

- Index documents by: **docid** and **text**

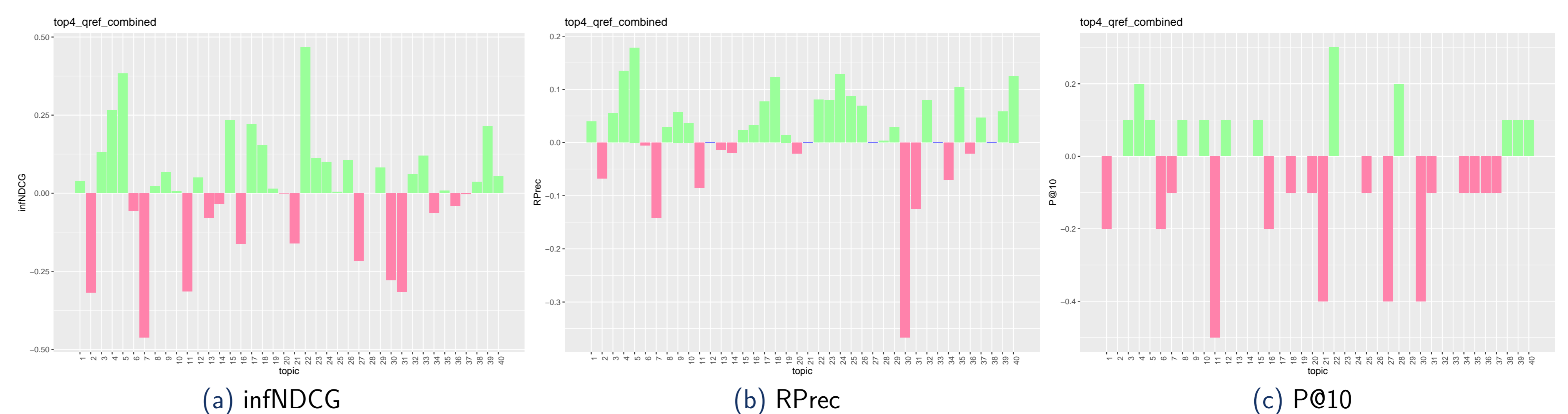
Query Reformulations (QR):

0. Keep each original query as baseline (BL)
1. Expand each original query with term variants of concepts belonging to *neop* semantic type
2. Expand each original query with term variants of concepts belonging to *neop* and *comd* semantic types
3. Expand each original query with term variants of concepts belonging to *neop* and *gngm* semantic type

Rank Fusion (RF):

- Combine the rankings obtained with each original query and the three query reformulations using CombSUM with Min Max normalization

Results



Precision Medicine Task 2: Clinical Trials

The 2019 Precision Medicine task 2 focuses on the problem of identifying the most relevant clinical trials for which a target patient is eligible.

BM25 + QR + F + RF

Search Engine:

- Whoosh

Concept Mapper:

- MetaMap

Indexing:

- Index documents by: **docid**, **text**, **max_age**, **min_age**, and **gender**

Query Reformulations (QR):

1. Reduce original queries by removing gene mutations from **gene** field and then expand reduced queries with term variants of concepts belonging to *neop* semantic type
2. Expand original queries that do not mention "lymphoma" or "leukemia" with the term "solid"
3. Expand reduced queries that do not mention "lymphoma" or "leukemia" with the term "solid"

Filtering (F):

- Filter trials by **min_age**, **max_age**, **gender**

Rank Fusion (RF):

- Combine the rankings obtained with the three query reformulations using CombSUM with Min Max normalization

Results

