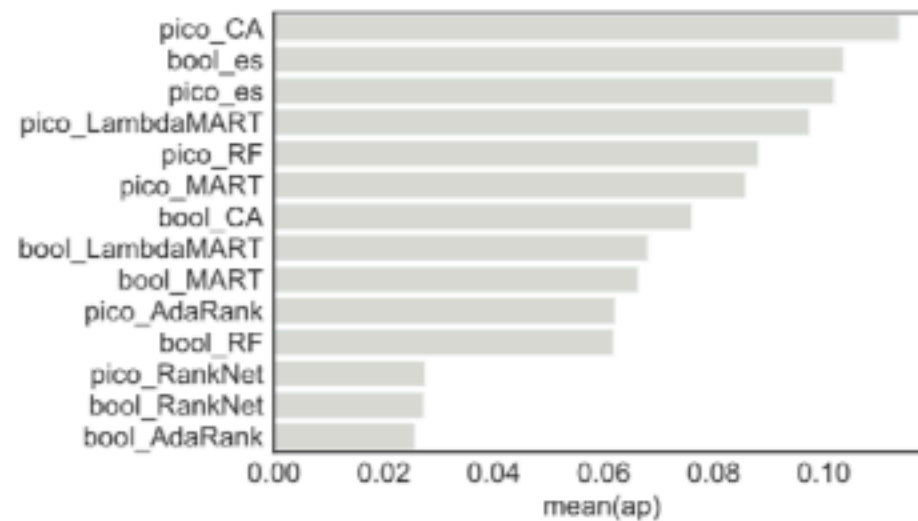


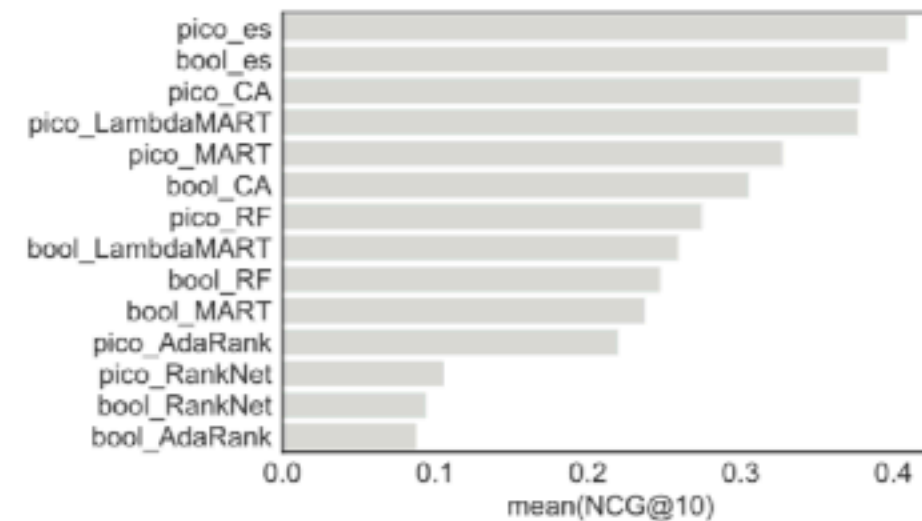
# Task 4: Screening Prioritisation

- Trained a learning to rank models using clinical features (PICO) and limited number of QPPs
  - IDFSum, IDFStd, IDFMax, IDFAvg, P-Count, I-Count, O-Count
- Systematic review queries then transformed manually into Elasticsearch queries
- Train 5 LTR models using these features and evaluate the performance of each
  - MART, AdaRank, Coordinate Ascent, LambdaMART, Random Forests
- Compare the difference between queries with PICO features and non-PICO features

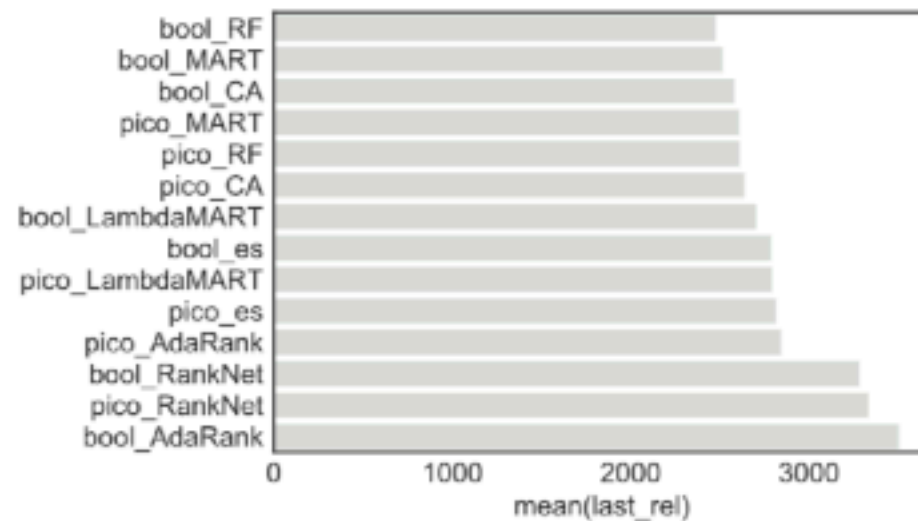
# Task 4: Screening Prioritisation



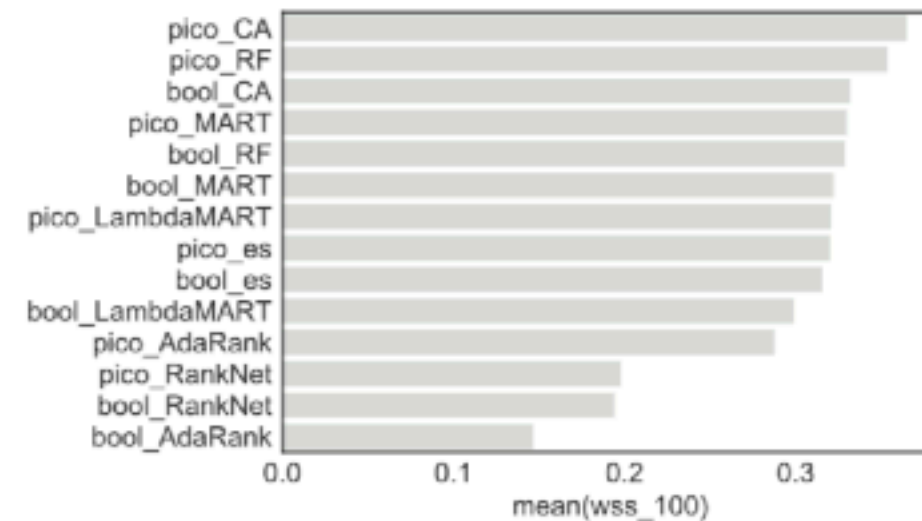
(a) Average Precision (AP) for each run (including baselines: bool\_es and pico\_es).



(b) Normalised Cumulative Gain at position 10 (NCG@10) for each run (including baselines: bool\_es and pico\_es).



(c) Position in the re-ranked list of the last study retrieved (including baselines: bool\_es and pico\_es).



(d) Position in the re-ranked list of the last study retrieved (including baselines: bool\_es and pico\_es).