Global Evidence Summit

Automating Study Screening for Systematic Reviews Using a Large Language Model

Zhaozhen Xu

Senior Research Associate in Machine Learning

MRC Integrative Epidemiology Unit, University of Bristol

zhaozhen.xu@bristol.ac.uk

Using evidence. Improving lives.

Declaration of Conflict of interest

I have no actual or potential conflict of interest in relation to this presentation.

This project is funded by the **World Cancer Research Fund.**

Collaboration between the **University of Bristol** and **Imperial College London**

Global Evidence Summit

World Cancer Research Fund

UKRI MRC Integrative Epidemiology Unit

University of BRISTOL

Background

The World Cancer Research Fund has an ongoing research programme called Global Cancer Update Programme (CUP Global) Since 2007

- Provide up-to-date systematic reviews to analyse the evidence linking diet, nutrition and physical activity to the risk of, and survival from, cancer
- From 2023, CUP Global with the University of Bristol start seeking Artificial Intelligence (AI) approaches to increase efficiency and reduce error involved in conducting the systematic reviews

➔Using a Large Language Model (LLM) to (semi-)automate the process of identifying primary studies for different systematic review topics

➔Aim: Provide reliable inclusion/exclusion predictions for each study among all existing topics

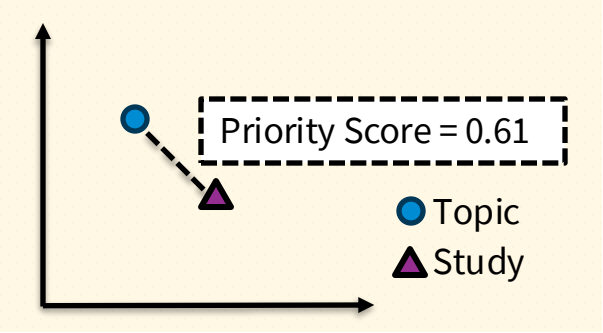
MRC Integrative Epidemiology Unit

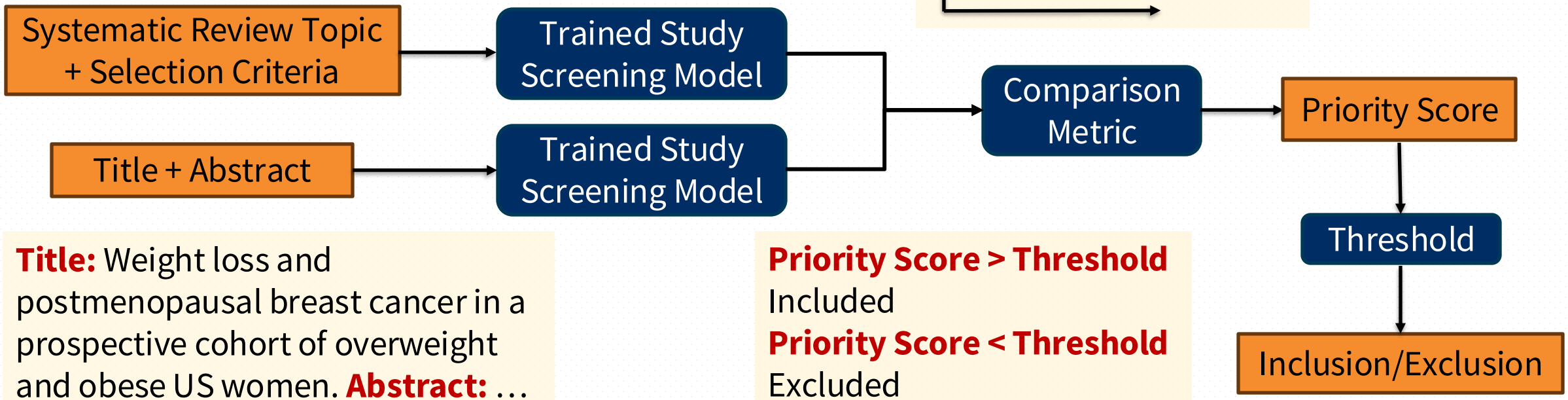
University of BRISTOL

General Study Screening Model

Based on BlueBERT, an LLM pre-trained on general knowledge and biomedical contents

Systematic Review Topic:
Breast Cancer Incidence study
Criteria: including studies …





Title: Weight loss and postmenopausal breast cancer in a prospective cohort of overweight and obese US women. **Abstract:** …

Priority Score > Threshold
Included
Priority Score < Threshold
Excluded

MRC Integrative Epidemiology Unit

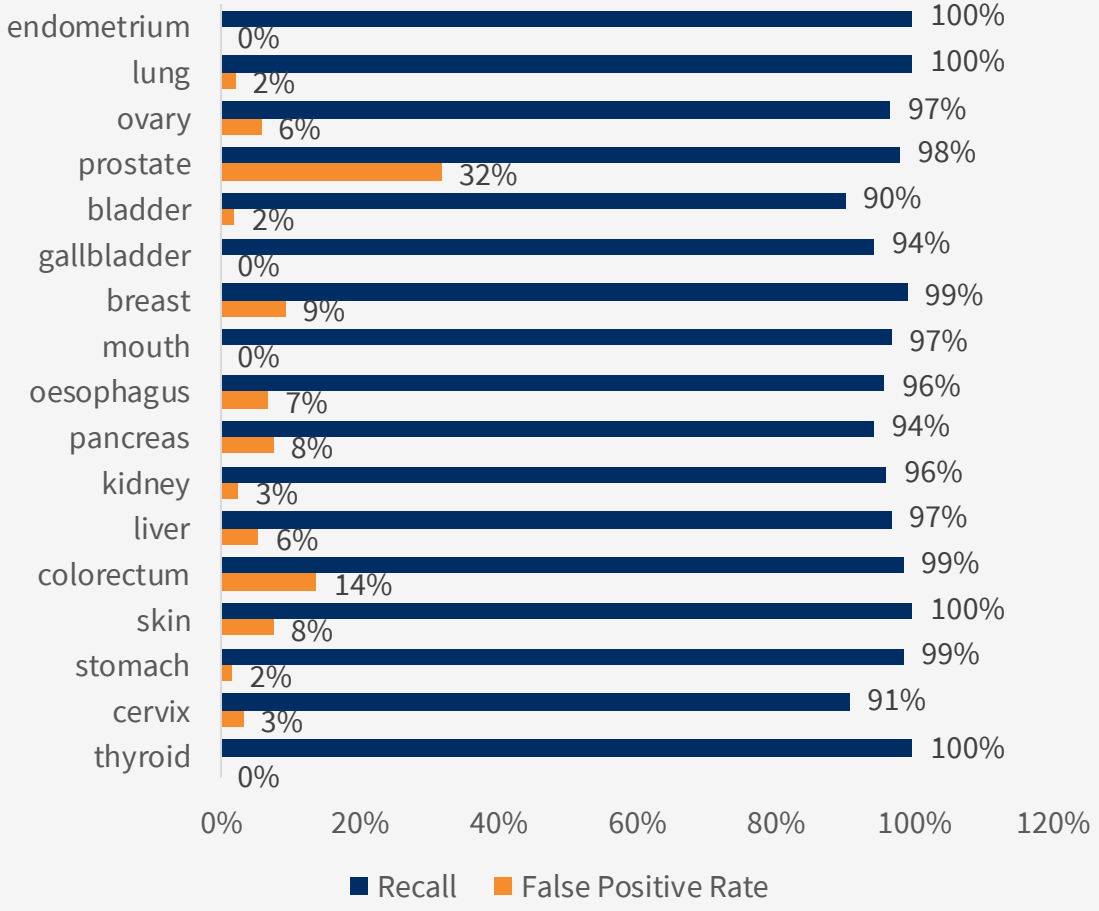
University of BRISTOL

Results on Cancer Incidence Study Screening

- We train and evaluate our general model on 17 cancer incidence review topics
- The primary studies are manually reviewed by the CUP-Global Team
- The model has more than 90% of Recall among all topics on the test set
- In the meanwhile, the false positive rate is under 10% for most of the topics

➔Identify most of the primary studies without including too many false positive studies to review

➔Excluded the studies on title abstract level, whereas humans need to review the full text



MRC Integrative Epidemiology Unit

University of BRISTOL

Future Work and Acknowledgement

Ongoing

Analysing and improving our study screening model on a more general domain

Future Work

Automating other processes of systematic review:

- Data extraction
- Risk of bias analysis

University of Bristol team

Core team: Yi Liu, Zhaozhen Xu, Tom Gaunt

Advisor: Louise Millard, Richard Martin, Julian Higgins, Philippa Davies, Maria Sobczyk-Barad

Imperial College London team

Doris Chan, Kostas Tsilidis, Eduardo Seleiro, Lam Teng, Georgios Markozannes

World Cancer Research Fund

MRC Integrative Epidemiology Unit

University of BRISTOL

Global Evidence Summit

Open to collaboration! Please contact:

Zhaozhen Xu

zhaozhen.xu@bristol.ac.uk

Using evidence. Improving lives.