

An evaluation of trajectory analysis for disease risk assessment: a scoping review

Freya Pollington, Prof. Spiros Denaxas, Prof. Ken Li, Prof. Georgios Lyratzopoulos, Dr. Becky White

Epidemiology of Cancer Healthcare and Outcomes, UCL Department of Epidemiology and Public Health



Background

Patient trajectory analysis explores clinical event associations, characterises disease trajectories, and aims to enhance risk prediction for diagnosis, readmission, and mortality. A review encompassing temporal statistical, ML and DL models is needed to identify potential applications and assess improvements over static models.

Aims

- Which model architectures are commonly used to model patient trajectories?
- Are primary and secondary care settings equally represented in the evidence base?
- Do temporal models outperform static baselines?
- What disease outcomes have been explored?

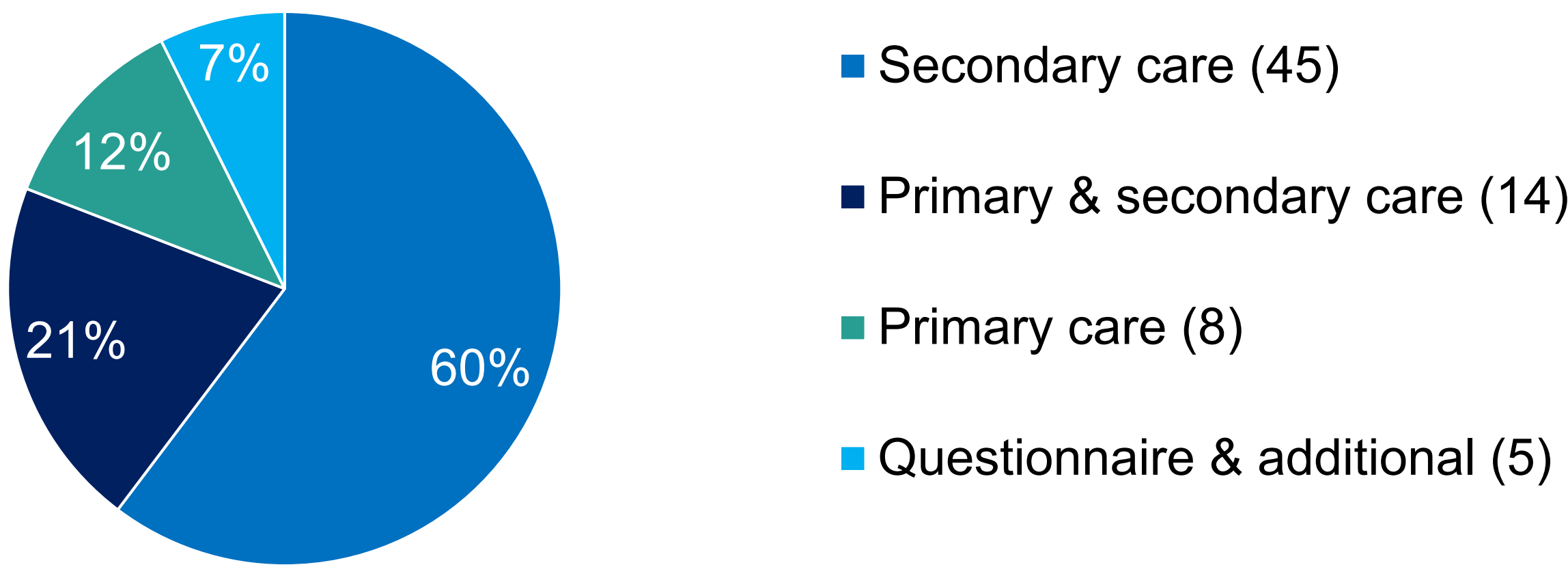
Methods

We conducted a PubMed and Web Of Science search, enhanced by citation tracking, for studies using temporal EHR sequences from structured datasets to identify disease signatures or evaluate disease-specific and disease-agnostic outcomes.

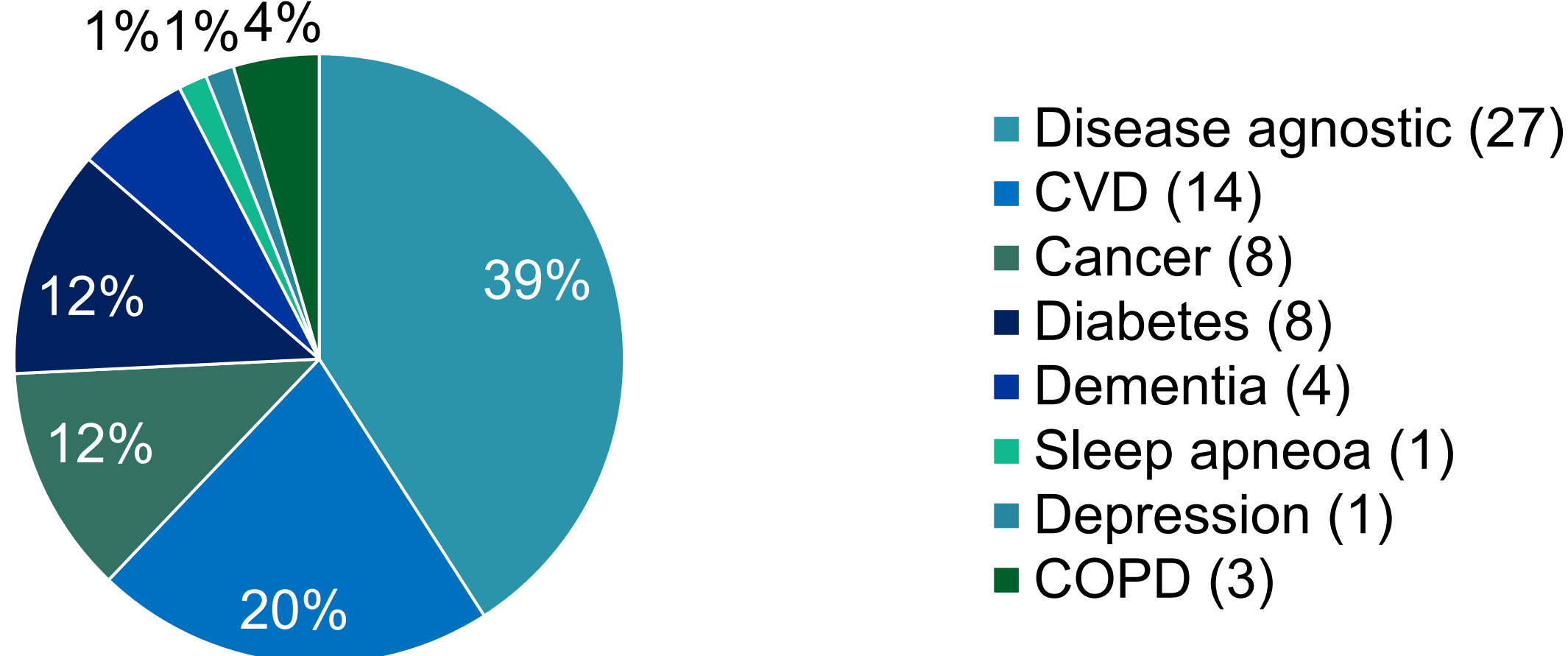
Inclusion criteria for studies were:

- Use of trajectory modelling methods in longitudinal structured EHRs.
- Focus on diagnosis-related outcomes.
- Testing associations between diseases/disease-related factors or predicting disease risk.

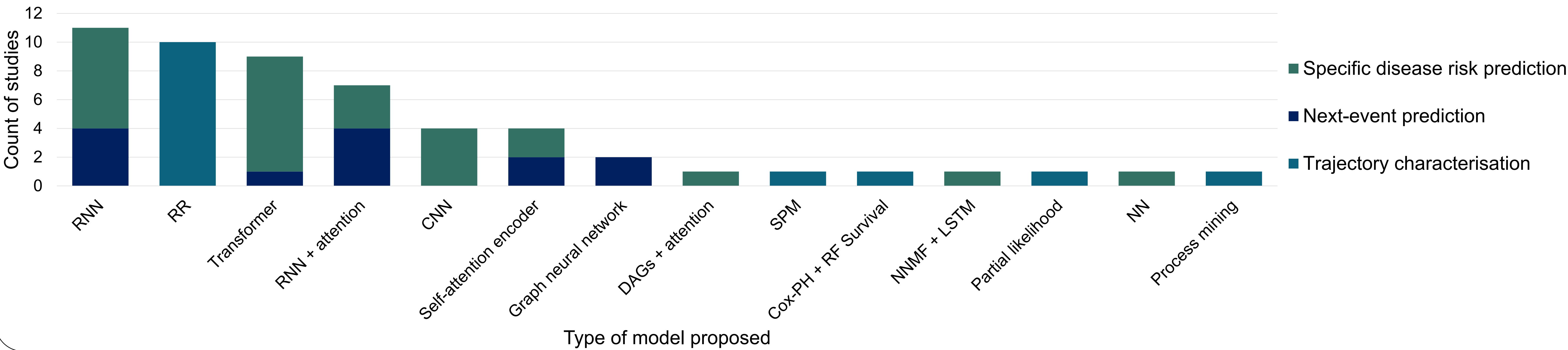
Primary care data sources are underutilised in trajectory studies



Predicted disease outcome proportions in trajectory studies



The distribution of model types across trajectory studies

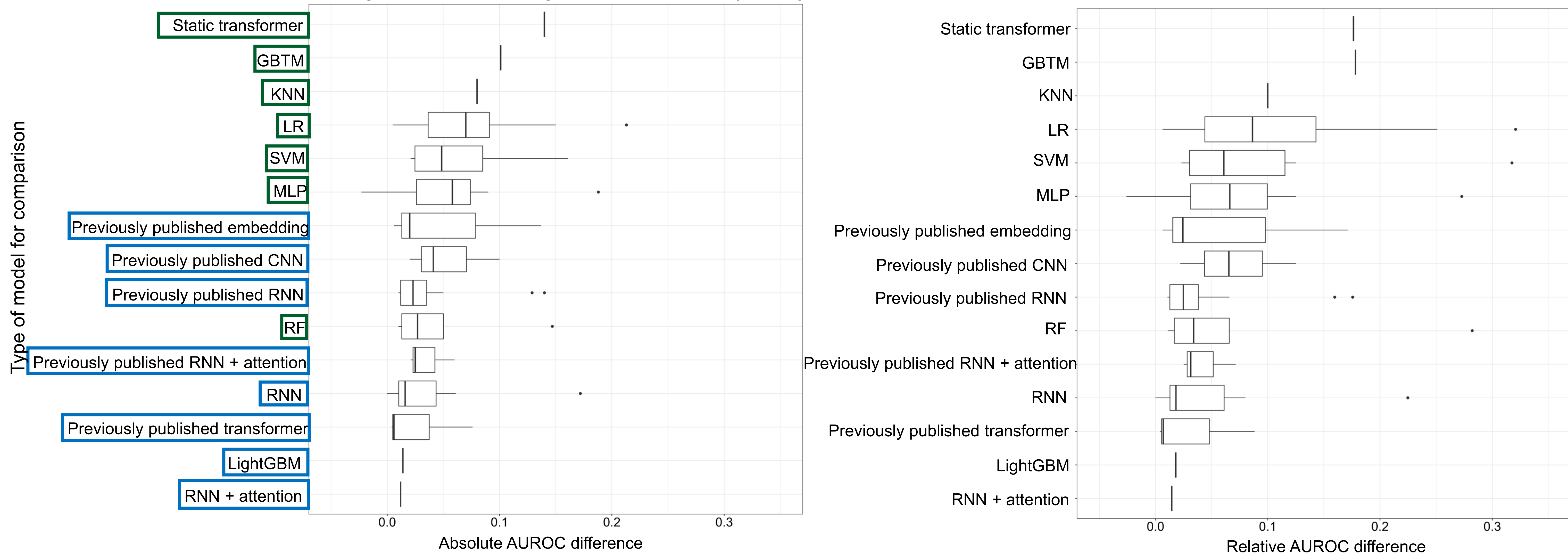


Results

- 54 studies from databases, 8 via citation tracking.
- Trajectory models involved statistical and deep learning methods.
- 48 risk prediction studies, 6 associations studies.
- 60% used secondary care data.

- 50% of risk prediction trajectory studies compared to static, and 77% compared to temporal models.
- Temporal models achieved relative AUC gains (Median: +5.7%).
- Model performance decreased when excluding data close to diagnosis.
- Performance metric use varied; 72% reported 2+.

Larger performance gain is seen for trajectory models in comparison with static vs. temporal baselines



Boxplots displaying the distribution of incremental gain in absolute and relative AUC performance across trajectory-based models in comparison to the authors' chosen baselines, with the baseline categories displayed on the Y axis. The green border around a baseline model type indicates a static model type and blue a temporal model type.

Conclusions

- Primary care datasets are lacking, as picked up by previous review articles.
- There is a general lack of application of trajectory methods to cancer outcomes.

- For outcomes with known risk predictors, improvements could be limited.
- For conditions without known predictors, trajectory analysis may reveal factors of interest, improve patient stratification and personalised predictions.
- Continued research needed to determine if this growing field can deliver meaningful clinical benefits.



Publication available at:
JAMIA
A SCIENTIFIC JOURNAL OF INFORMATION IN HEALTH AND BIOMEDICINE

