

Supplementary Material

A. Data appendix

| Journal | #papers |
|---|---------------|
| Acta Oncologia | 5,846 |
| British Journal of Cancer | 306 |
| radiation oncology journal | 2,810 |
| Journal of Clinical Oncology | 522 |
| Journal of Hematology and Oncology | 1,103 |
| Lancet | 1,110 |
| Nature Reviews Clinical Oncology | 195 |
| Radiation Oncology | 775 |
| Radiologica Medica | 1,873 |
| Strahlentherapie and Onkologie | 2,985 |
| Technology in Cancer and Research | 1,886 |
| Technology in Cancer Research and Treatment | 388 |
| Total | 19,799 |

Table 1: Statistics of text corpus for oncology-domain

The biomedical KG defines 18 entity types in the medical domain, including anatomy, neoplastic process, microorganism, eukaryote, physiology, chemical or drug, diagnostic procedure, laboratory procedure, research activity or technique, therapeutic or preventive procedure, medical device, research device, pathology, disease or syndrome, anatomical abnormality, mental or behavioral dysfunction, injury or poisoning and sign, symptom or finding.

Meanwhile, 19 relationship types are defined in the biomedical KG including “is_a”, “reverse_is_a”, “is_part_of”, “reverse_is_part_of”, “may_treat”, “reverse_may_treat”, “found_in”, “reverse_found_in”, “may_cause”, “reverse_may_cause”, “expressed_in”, “is_expression_of”, “encodes”, “encoded_by”, “significant_drug_interaction”, “involved_in_biological_process”, “biological_process_involves”, “is_active_ingredient_in”, “has_active_ingredient”.

Links of coarse-domain KG and pre-trained language models (PLMs):

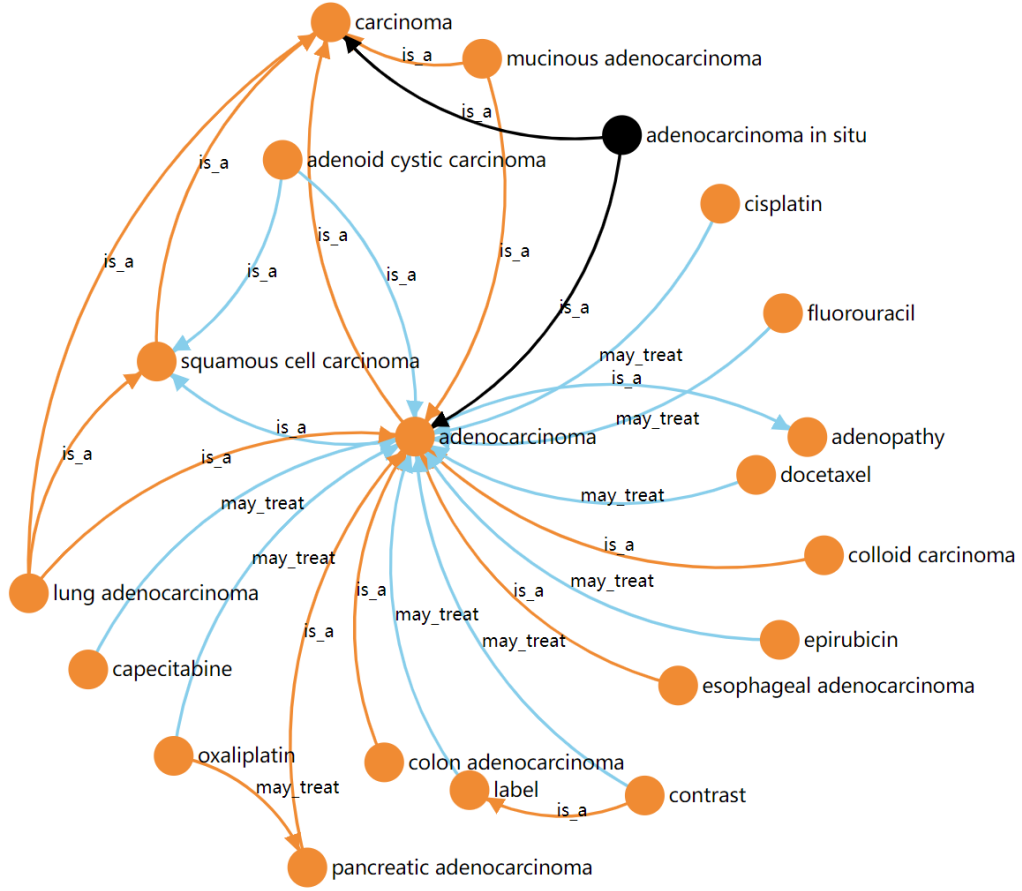
coarse-domain KG (BIOS 2022v1(beta)): <https://bios.idea.edu.cn/Download>

emilyalsentzer-Bio_ClinicalBERT: https://huggingface.co/emilyalsentzer/Bio_ClinicalBERT

allenai-biomed_roberta_base: https://huggingface.co/allenai/biomed_roberta_base

bert-base-uncased: <https://huggingface.co/bert-base-uncased>

B. Knowledge graph display



Knowledge graph of the *adenocarcinoma*. The orange edges represent the overlapping triples \mathbb{T}_o , the blue edges denote the triples of new relations \mathbb{T}_r , and the black edges denote the triples of new entities \mathbb{T}_E , while the black node represents the new entity. For visual presentation, we only show 20 related entities which relationship type is *is_a* or *may_treat*.