Temporal Knowledge Acquisition from Cancer Clinical Trials

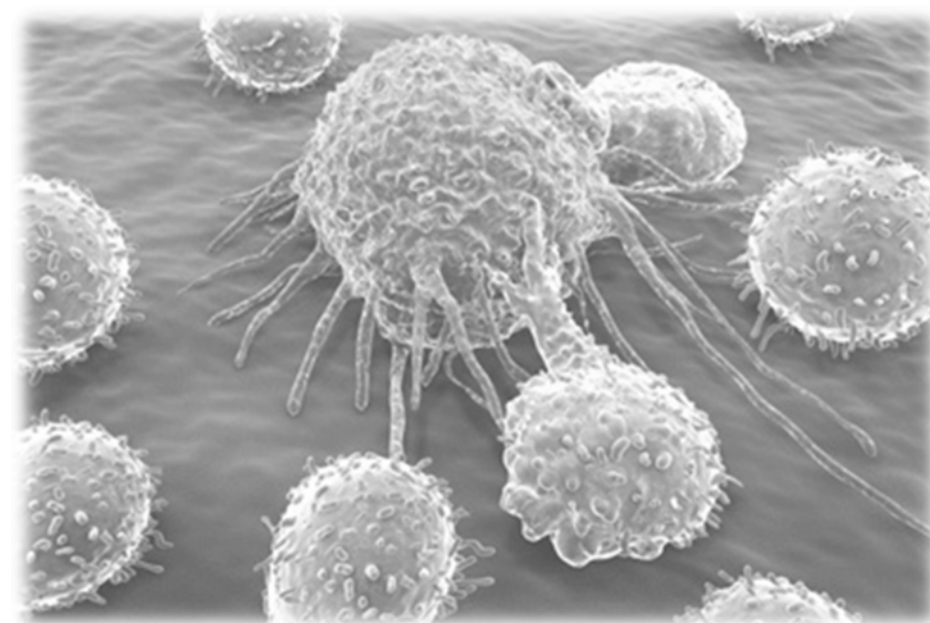Ziran Li and Huanyao Chen
DBMI
Columbia University

Background

➢ **Many review articles have summarized the progress in specific domain of cancer research**

- **Prostate cancer and Breast cancer**
- Diagnostic and therapeutic approaches have improved steadily since late 20ᵗʰ century

- **Pancreatic cancer**
- Limited medication options

Background

Limitations:

1. Only qualitative description of the cancer milestones based on domain experts' experience, which limited the potential of using this temporal knowledge.

e.g. Lack of parallel comparison.

Opportunity:

1. Important progress in treatment of cancer arises from the results of rigorously conducted clinical trials

2. Database of cancer clinical trials enables efficient and scalable knowledge acquisition

3. We may identify cancer treatment progress by analyzing original longitudinal data of clinical trials

Goal

➢ **Identify** anti-cancer medications that trend upwards or downwards

➢ **Show** cancers have similar or different trends of medication interventions in a quantitative way

➢ **Demonstrate** database of clinical trials serves as a valid source for supporting trend analysis and creating temporal knowledge
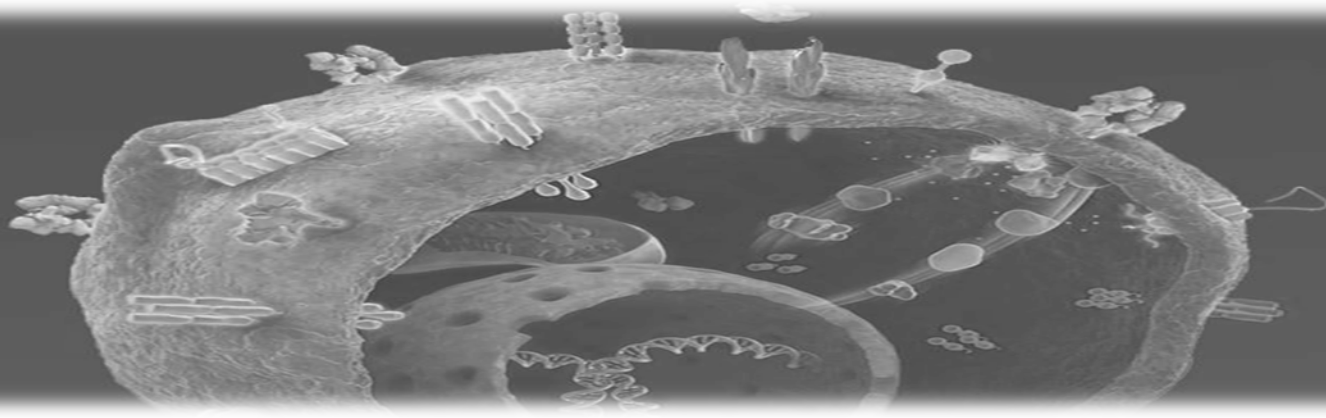
Method

Data Retrieval

Trial information retrieval from ClinicalTrials.gov

Selected "interventional" clinical trials with purpose of "treatment"

Formulate oncology clinical trial dataset by selecting clinical trials with condition concepts under a semantic type of neoplastic process

Method

Terminologies

➢ Conditions were mapped to MeSH terms.

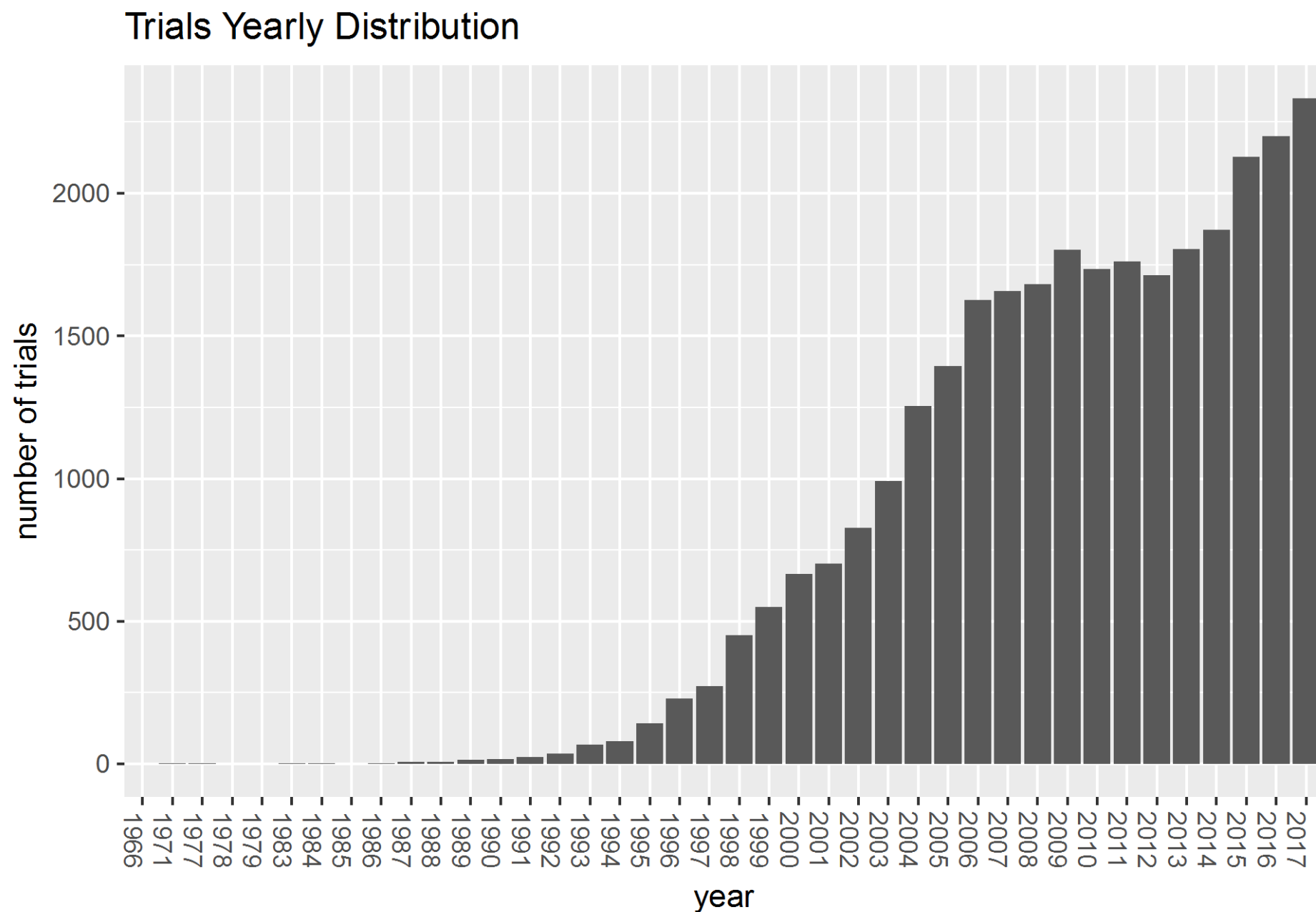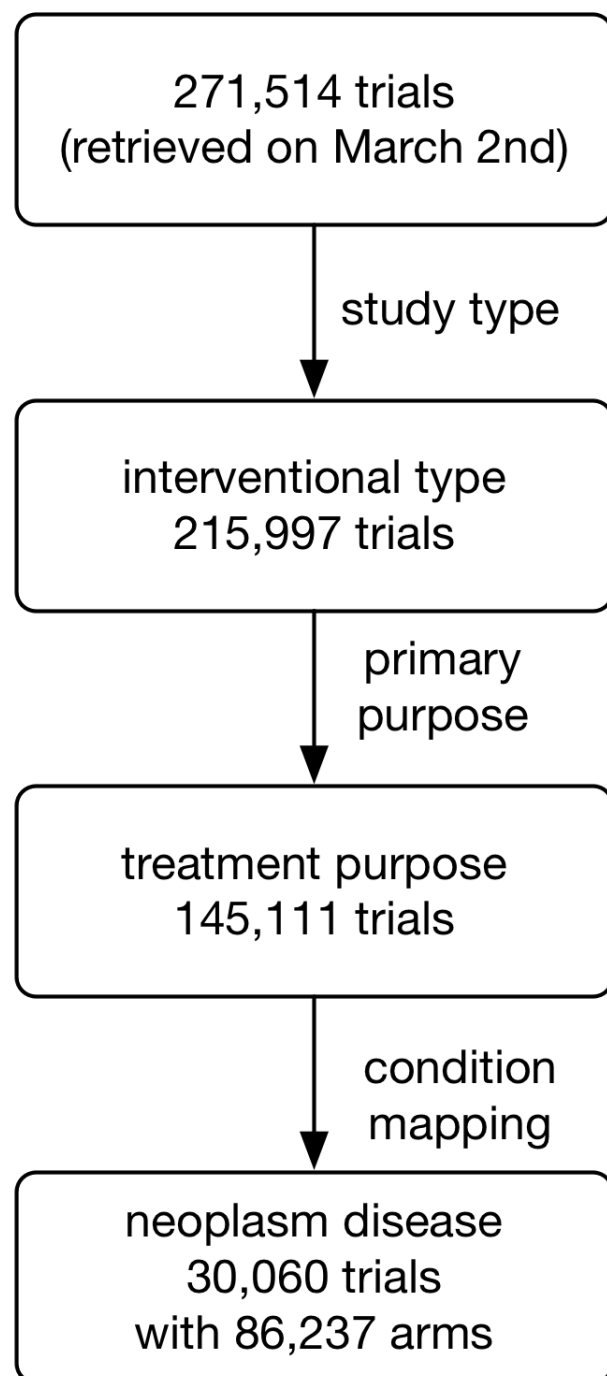➢ Medications were mapped to NCI drug dictionary (NCI thesaurus code)

Method

Data Preprocessing

➢ Classify medication into 4 groups based on the semantic type of ingredients (conventional chemical drugs, hormone, immunologic factors and bioactive substance)

➢ Normalized Frequency(NF): in a period of time, the number of trial arms of interested drug(or drug combinations) divided by the total number of trial arms.
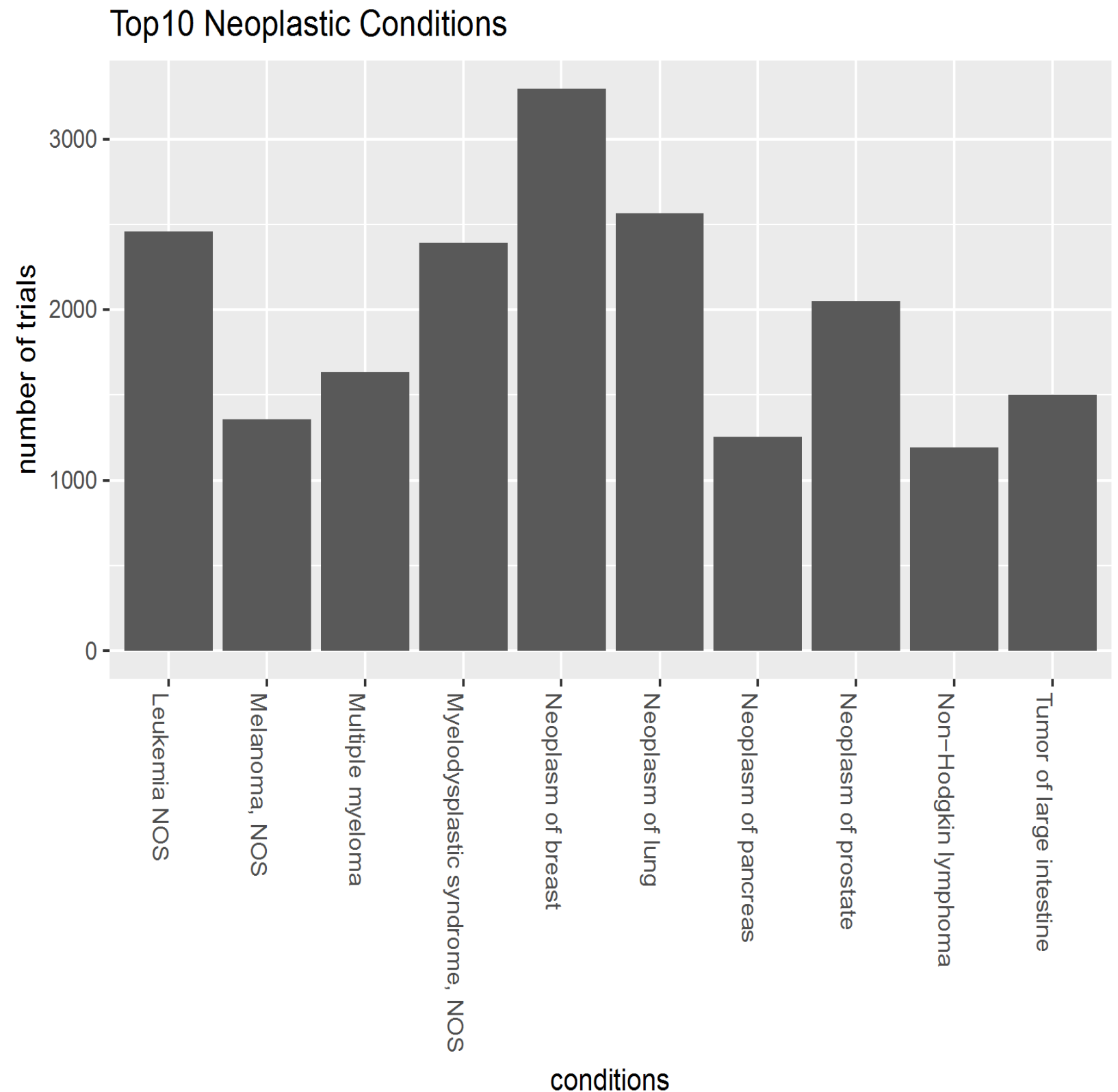
Result

271,514 trials
(retrieved on March 2nd)

study type

interventional type
215,997 trials

primary
purpose

treatment purpose
145,111 trials

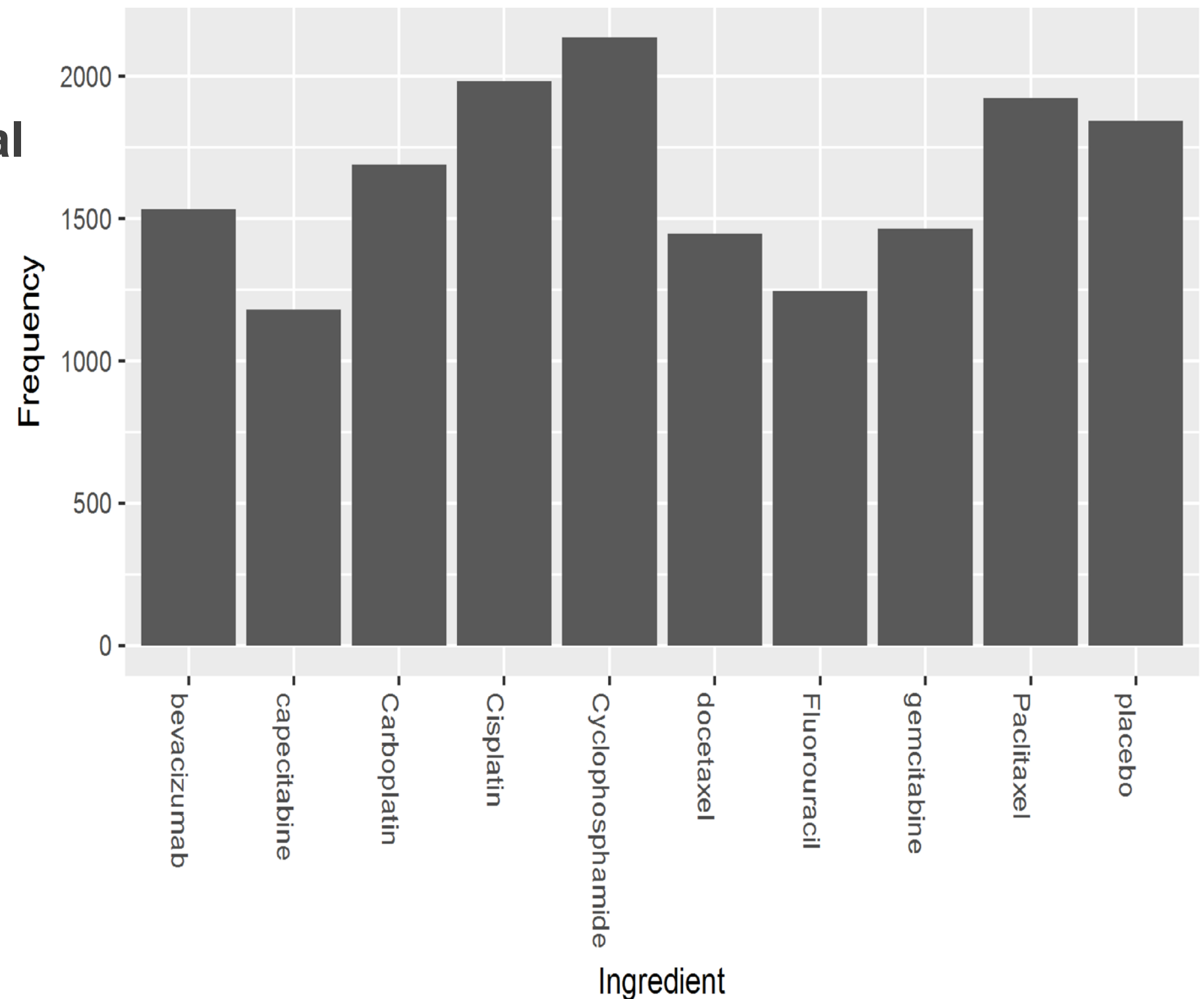condition
mapping

neoplasm disease
30,060 trials
with 86,237 arms

Result

Identified 260 condition concepts under the semantic type of neoplastic process after manually review

Top10 Neoplastic Conditions

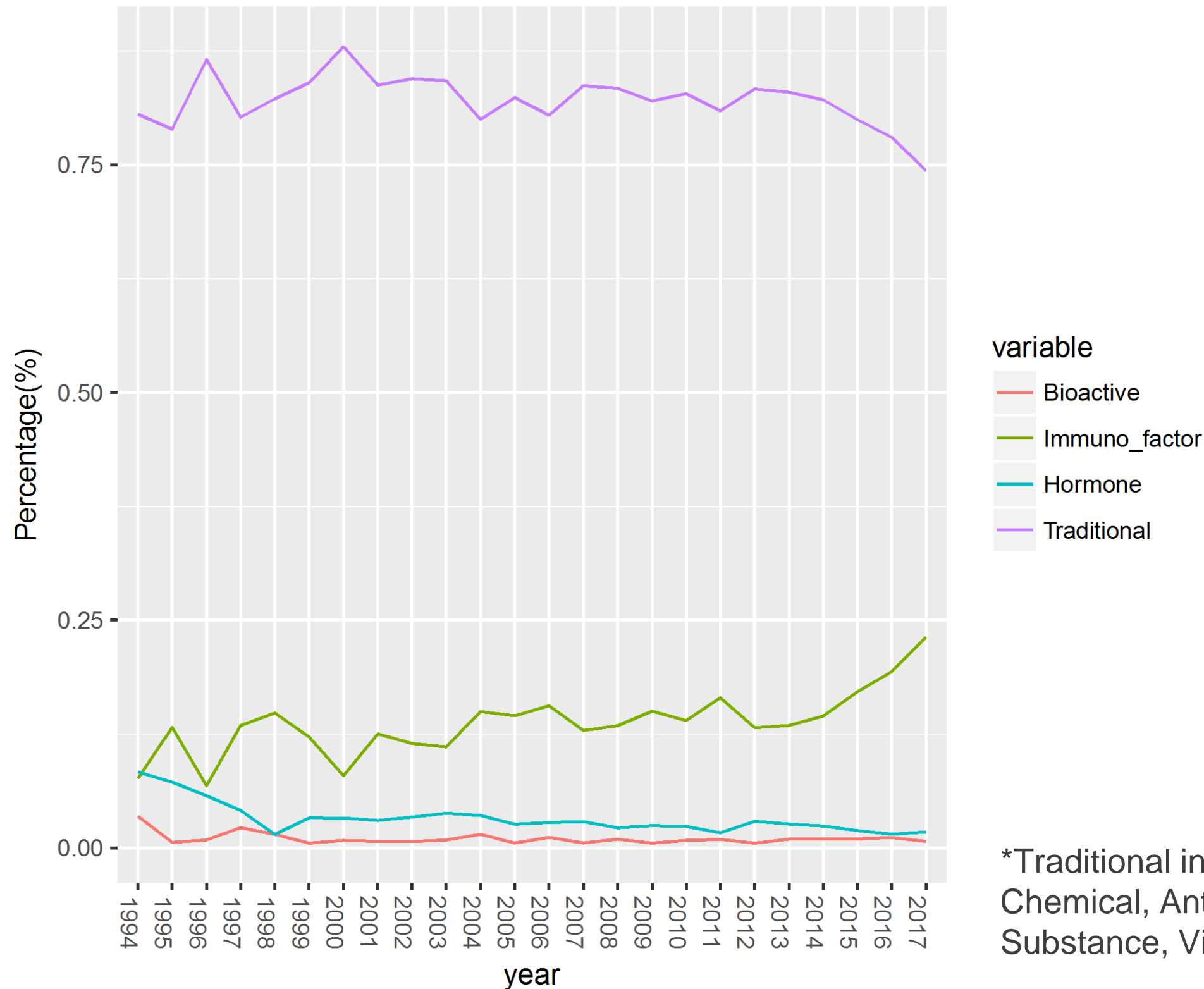# Result

Identified 1035 medications in total

Top10 Ingredients in cancer clinical trials

Result

Anti-cancer medications trend

Drug Type Trends

Top10 immunologic factor medications:

1. Bevacizumab
2. Pembrolizumab
3. Trastuzumab
4. Rituximab
5. Cetuximab
6. Nivolumab
7. Filgrastim
8. Ipilimumab
9. Atezolizumab
10. Aldesleukin

*Traditional includes Inorganic/Organic Chemical, Antibiotic, Pharmacologic Substance, Vitamin.
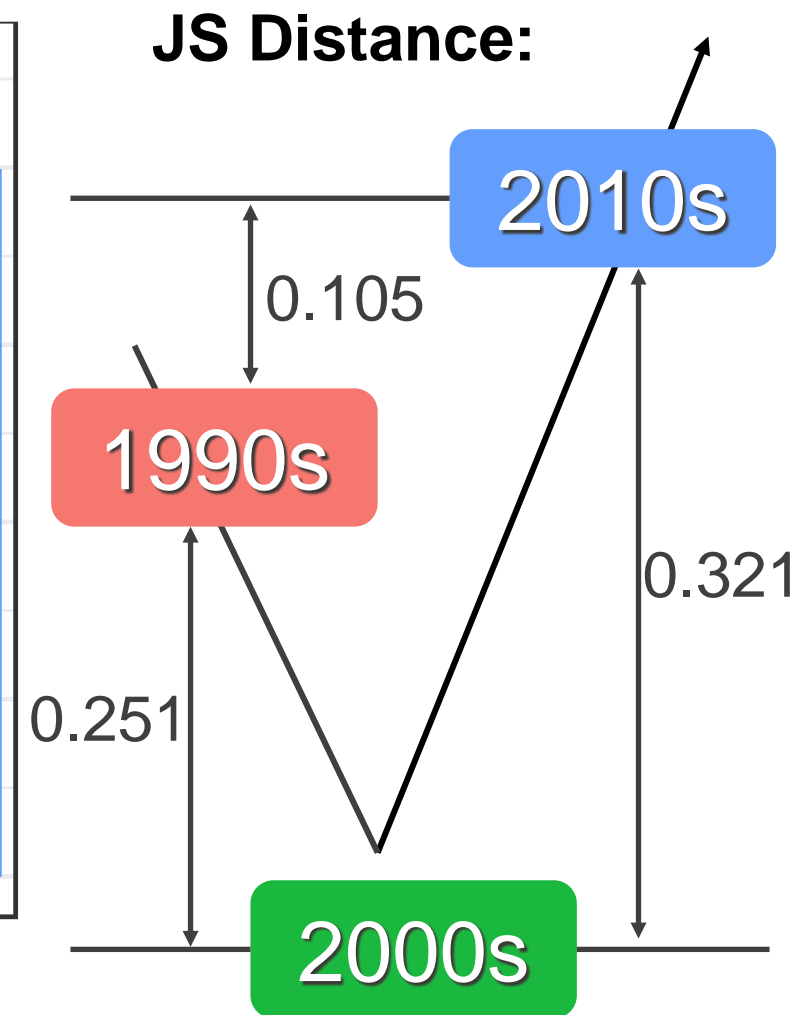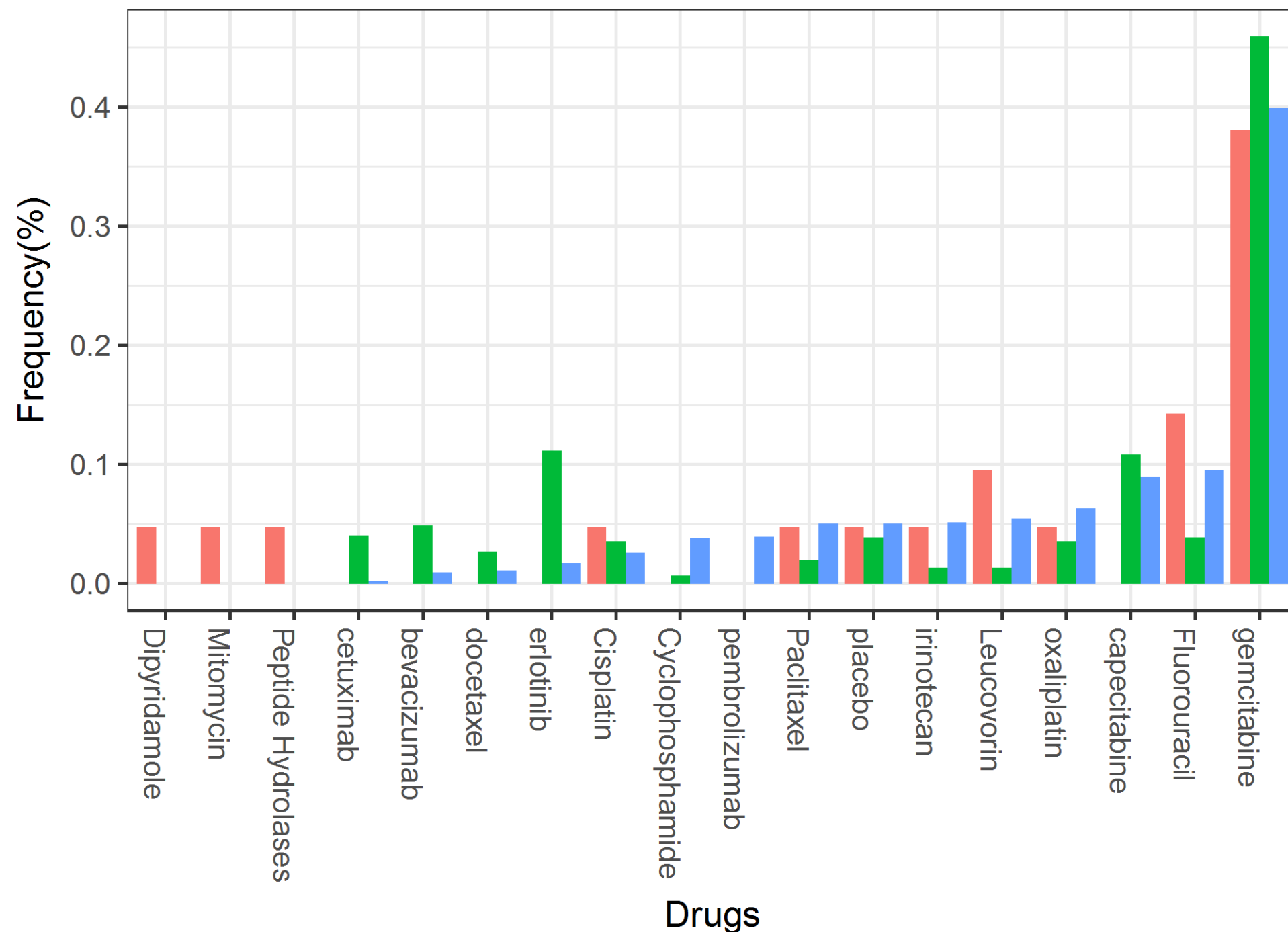
Result

Medication Evolution for Breast Cancer

Drug Distribution for Breast Cancer

JS Distance:

2010s

0.221

2000s

0.436

0.351

1990s
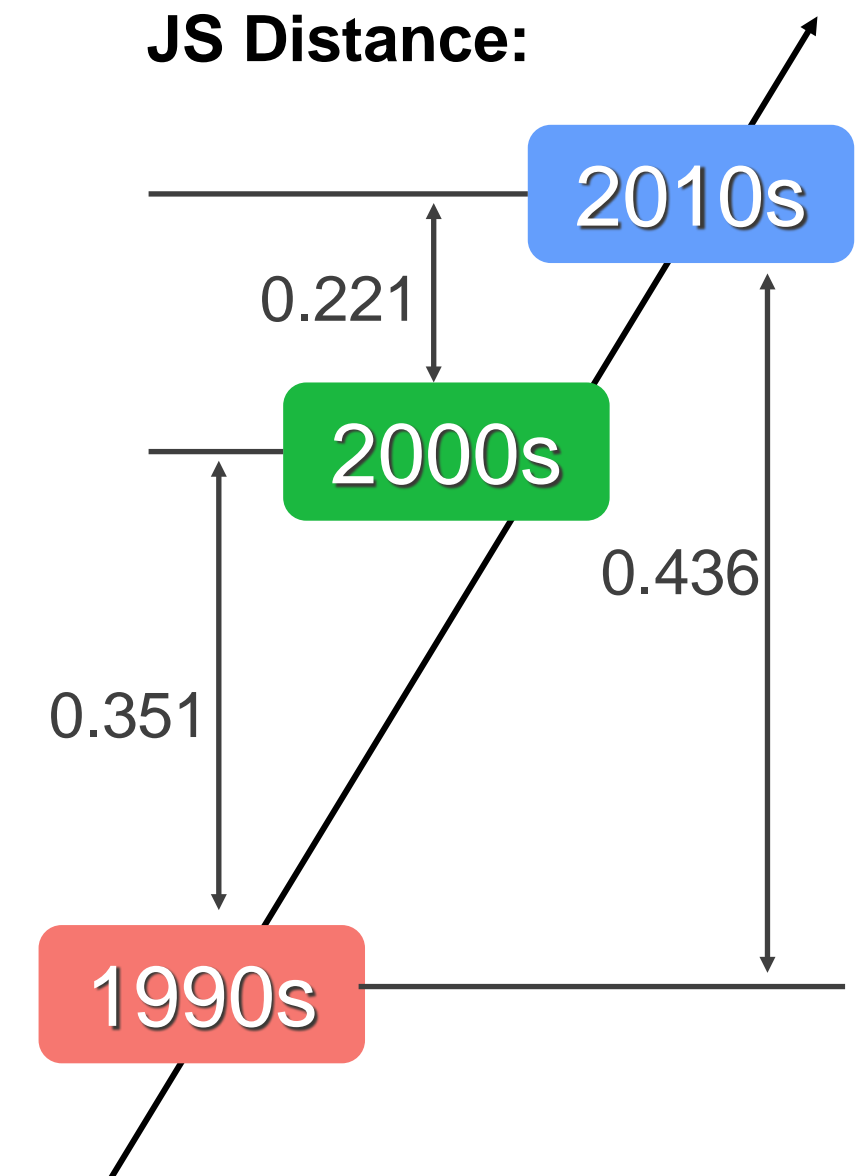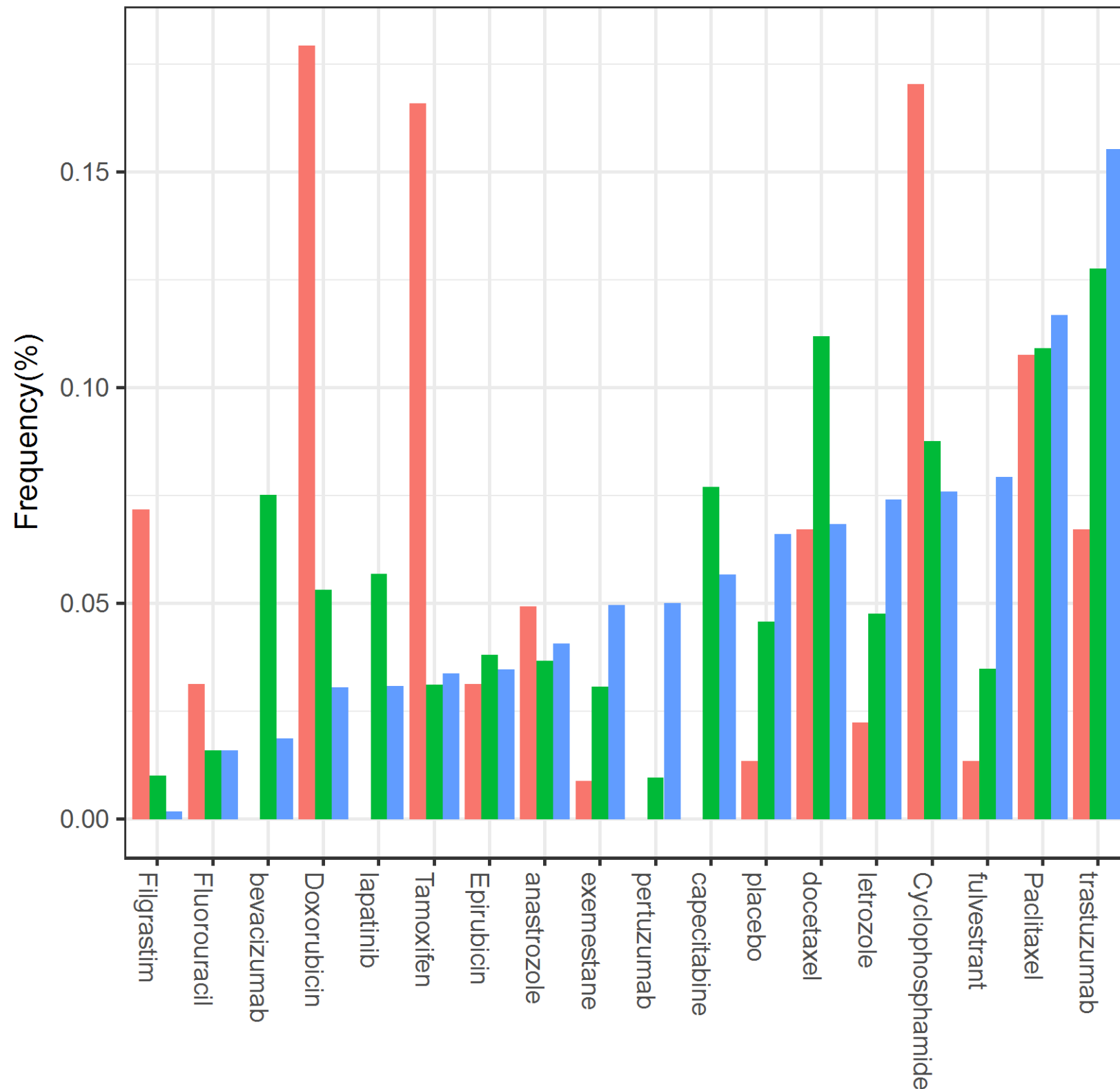
Result

Medication Evolution
for Pancreatic Cancer

Drug Distribution for Pancreatic Cancer

Discussion

➢ Focus on NF as our unit. The NF can be interpreted as a reflective of research popularity over a specific time period, the change of NF reflects the evolution and future trends.

➢ The medication development evolution and trends can be represented in a quantitative manner and compared using the distance between distributions.

Working on …

➢ Make this analysis more scalable.

  ➢ For any condition/sub-type of condition in any period of time

  ➢ To discover new drug and drug combinations tested in clinical trials since any time.

  ➢ …

➢ Standardize the query process and build a knowledge base.

Future plan …

- ➢ Compare the evolution and trends identified in clinicaltrial.gov with the patients' data in OHDSI

  - ➢ If the success achieved in clinical trial can be applied into the real world?

  - ➢ How long does it take to introduce an intervention from clinical trial to clinical practice?

  - ➢ Is there any difference of speed of this process among different diseases?

  - ➢ ……

Any feedback and suggestion will be greatly appreciated.