

Analyzing the Landscape of Anti-Cancer Drug Research in Pancreatic Cancer

Yoo Kyung Jeong
Dept. of LIS
Yonsei University,
Seoul, South Korea
yk.jeong
@yonsei.ac.kr

Go Eun Heo
Dept. of LIS
Yonsei University,
Seoul, South Korea
goeun.heo
@yonsei.ac.kr

Keun Young Kang
Dept. of LIS
Yonsei University,
Seoul, South Korea
ky.kang
@yonsei.ac.kr

Dong Sup Yoon
Pancreatobiliary
Cancer Clinic, Dept. of
Surgery, Gangnam
Severance Hospital,
Yonsei University
College of Medicine
yds6110@yuhs.ac

Min Song
Dept. of LIS
Yonsei University,
Seoul, South Korea
min.song
@yonsei.ac.kr

Categories and Subject Descriptors

I.7.1 [Document and Text Processing]: Document and Text Editing – Document management

General Terms

Experimentation.

Keywords

Pancreatic cancer; Text mining; Named entity recognition; Anti-cancer drug analysis; Topic modeling

1. ABSTRACT

Due to the importance of anti-cancer drug development, there has been a surge of publications pertinent to the field of pancreatic cancer. Therefore, a need for analysis of drug research trend is recognized to minimize risk of new drug testing and understand the anti-cancer drug research area. The purpose of this paper is to identify the landscape of pancreatic cancer drug research by analyzing the anti-cancer drugs extracted from PubMed records and clinical trials datasets. We conduct drug cluster analysis by using the topic modeling technique and network analysis of drug clusters. Comparison of two distinct datasets, scientific publications vs. clinical trials, gives us a new insight into identifying the different portrait of anti-cancer drug research in pure research and clinical settings.

The results show that two distinct research trends are observed by ClinicalTrials dataset and PubMed records. It is the major difference with other bibliometric studies in cancer research. Using two different data sources, we can identify different drug research patterns. First, the assumption that drug research published in PubMed is preceded by clinical trials is statistically confirmed in our study. Second, a research trend of new drug testing with various targets is observed in clinical data. On the other hand, we identify that diverse chemicals (e.g. sodium, phosphatidylinositols, and calcium) together with standard therapeutic agent are studied in scientific publications in PubMed.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
DTMBIO '15, October 23, 2015, Melbourne, Australia.
ACM 978-1-4503-3787-8/15/10.
DOI: <http://dx.doi.org/10.1145/2665970.2665982>

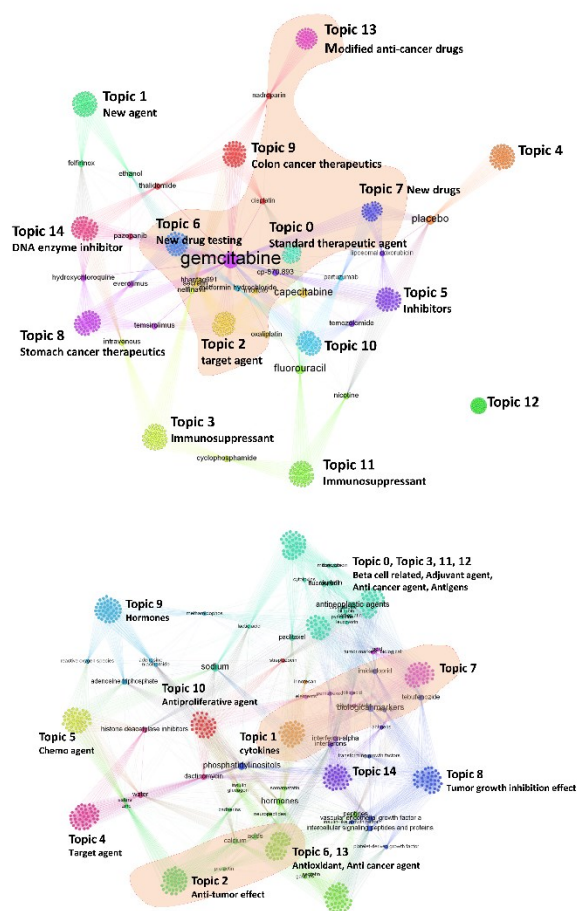


Figure 1. Drug cluster network of ClinicalTrials.gov (Top) and PubMed (Bottom)

2. ACKNOWLEDGMENTS

This work was supported by the Bio-Synergy Research Project (NRF-2013M3A9C4078138) of the Ministry of Science, ICT and Future Planning through the National Research Foundation.