HBP SURGERY WEEK 2023
MARCH 23 THU - 25 SAT, 2023 | BEXCO, BUSAN, KOREA
www.khbps.org
& The 58th Annual Congress of the Korean Association of HBP Surgery
The Korean Association of Hepato-Biliary-Pancreatic Surgery

ABST-0521

Systematic Review Of Research Progress On Hepatocellular Carcinoma Over 30 Years: A Machine-learning Bibliometric Analysis

Ki-Seong LEE[1], **Ji Woong HWANG***[2], Hee Ju SOHN[2], Sanggyun SUH[2], Sun-Whe KIM[2]

[1]Da Vinci College Of General Education, Chung-Ang University College Of Medicine, REPUBLIC OF KOREA
[2]Department Of Surgery, Chung-Ang University Gwangmyeong Hospital, REPUBLIC OF KOREA

Background : The Latent Dirichlet allocation (LDA) is one of the most classical topic modeling techniques and usually used to analyze scientific publications comprehensively by identifying research topics. The aim of this study was to evaluate the research progress and development on HCC performing bibliometric analysis by machine learning and suggest future direction of studies.

Methods : Comprehensive research was performed in the PubMed indexed under MeSH term “hepatocellular carcinoma” between 1991 and 2020. The publication year, the type of research, the abstract, and the MeSH terms of each article were extracted and analyzed. Only Mesh term with 100 times or more of appearance were enrolled in this analysis. Using the text of the abstract, LDA analyses provided further information about hot topics and their relationships by constructing a topic network. The network visualizations were constructed by Gephi software.

Results : A total of 62,856 documents related to HCC from the past 30 years were searched and finally included in the bibliometric analysis. 980 MeSH terms, appeared 100 time or more, were included in this analysis, with a total of 945,754 times of occurrence. Among the diagnosis-related terms, “liver cirrhosis” was the most studied. “hepatectomy” was the most studied in the treatment-related MeSH terms, but recently studies related to “antineoplastic agents” outpaced this. LDA analyses provided hot topics and their relationships by constructing a topic network. The network was divided into three clusters including “basic research”, “diagnosis and treatment research”, and “epidemiology research”. The basic research cluster and diagnosis and treatment research cluster showed a poor connection with other clusters, while epidemiology cluster showed many connections to other clusters comparatively. The “gene & protein expression” of the basic cluster showed strong connections to “survival analysis, prognosis” and “pathology” belonging to other clusters.

Conclusions : This study, for the first time, analyzed more than 60,000 publications concerning HCC through machine learning during the past three decades, which will help researchers understand the current research trends and future research directions. Although much effort has being put into basic research, the connection with clinical practice is still lacking.

Corresponding Author : **Ji Woong HWANG** (dattoree@naver.com)