Quantitative and visualized analysis of 5,097 triple-negative breast cancer (TNBC) publications by using multiple machine learning algorithms: TNBC is an excellent forerunner in cancer research

Kangtao Wang

Heidelberg University <https://orcid.org/0000-0003-0948-0277>

Chanjuan Zheng

Hunan Normal University School of Medicine

Lian Xue

Hunan Normal University School of Medicine

Dexin Deng

Central South University Xiangya School of Medicine

Fengbo Tan

Xiangya Hospital Central South University

Qian Pei

Xiangya Hospital Central South University

Yuqiang Li

Xiangya Hospital Central South University

Dan Wang

Xiangya Hospital Central South University

Liang Zeng

Guangzhou Medical University

Ming Li

Central South University

Xiyun Deng (✉ dengxiyunmed@hunnu.edu.cn)

<https://orcid.org/0000-0003-2203-970X>

Research article

Keywords: triple-negative breast cancer, bibliometric analysis, machine learning, Latent Dirichlet Allocation

Posted Date: August 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-783425/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background:

Triple-negative breast cancer (TNBC) is a subtype of breast cancer proposed at the beginning of this century, which is still the most challenging breast cancer subtype due to its aggressive behavior, including early relapse, metastatic spread, and poor survival. This study explores current research status and deficiencies from a macro perspective on TNBC publications by using machine learning methods.

Methods:

All publications under the MeSH term "Triple Negative Breast Neoplasms" in PubMed were searched and downloaded as of December 2020. R and Python were used to extract MeSH terms, geographic information and other abstracts from metadata. The Latent Dirichlet Allocation algorithm was applied to identify specific research topics. The Louvain algorithm was used to establish a topic network, identifying the relationship between the topics.

Results:

A total of 5,097 publications were identified, with an average annual growth rate of 33.5%. Only 88 countries and regions in the world participated in TNBC research. The publications contain about 25% of clinical trial-related research. The rapid clinical progress will provide an excellent example for diagnosing and treating other tumors, suggesting that TNBC is an excellent forerunner in cancer research. Publications of pathogenic mechanisms and drugs are most studied. Using the Topic model, we found that the publications were mainly focused on three aspects: treatment plan research, new biomarkers research, and the regulation mechanism for TNBC aggressive behavior.

Conclusion:

This study quantitatively summarizes the current status of TNBC research from a macro perspective and will aid in redirecting basic and clinical research toward a better outcome of TNBC.

Introduction

It is well known that breast cancer is the most common malignancy in women. Breast cancer currently accounts for 30% of newly diagnosed malignant tumors in women and causes 15% of women to die from cancer [1]. For a long time, human research on breast cancer has been going on for very long time, Perou et al., for the first time, described the intrinsic molecular subtypes of breast cancer in 2000 using cDNA microarray technology [2]. However, the incidence rate of breast cancer is increasing worldwide due to industrialization and change in lifestyle [3]. Triple-negative breast cancer (TNBC) is the

most aggressive subtype of breast cancer, accounting for about 10–20% of breast cancer cases [4, 5]. Unfortunately, the survival of TNBC is unsatisfactory due to the lack of specific therapeutic targets [6]. In addition, with the rapid progress of the publishing industry, the number of papers published has risen sharply [7]. However, researchers and administrators could hardly understand the TNBC research situation from a macro perspective, rendering it difficult for them to have a global view of the whole picture of research on TNBC. Therefore, if there is no macro understanding and proper design to coordinate research progress, TNBC will continue to bring serious disasters to humankind in the foreseeable future.

Quantitative analysis of the literature could indicate the research interest, public concern and study deficiencies in the past and predict future research. Bibliometrics is a quantitative analysis method of academic publications, which can discover the progress of discipline research from a macro perspective and provide support for future research directions. TNBC related literature information analysis is extremely scarce. Teles et al. conducted a bibliometric study of 1,932 publications in 2018 to study nanomedicine research's global trend on TNBC [8]. However, the inclusion criteria of this study are too broad, and the analysis methods are insufficient to analyze the status quo of the TNBC study. Unfortunately, due to the lack of practical language analysis tools to integrate meta text data, the bibliometric research in all publications on TNBC is still missing.

Natural Language Processing (NLP) is a computing technology used to analyze human language, a part of machine learning [9]. This technology has been successfully applied to deal with medical text information [10]. Latent Dirichlet Allocation (LDA) is the most classical topic modeling method in bibliometrics to present many unstructured texts and information [10]. LDA can perform topic analysis on a large number of texts. LDA can discover the semantic analysis behind the text compared with the previous methods and has been widely used in medical and other fields [12]. We recently constructed LDA and NLP methods to analyze more than 23,000 rectal cancer-related publications between 1994 and 2018 [12]. We have successfully found the research deficiencies in the last 25 years and predicted the future research focus. Therefore, through the use of mature LDA methods and machine learning techniques to discover the current research from a macro perspective, at the same time discover the missing research topics in the past, and predict potential research breakthroughs in the future. In the present study, we analyzed all past TNBC publications indexed by PubMed under the medical subject heading (MeSH) term "Triple Negative Breast Neoplasms". We improved our algorithm based on our previous research and conducted a more detailed analysis of all TNBC publications with more visual expression to highlight current hot areas in TNBC, together with research deficiencies and specific areas with future opportunities.

Methods

All publications under the MeSH term "Triple Negative Breast Neoplasms" were downloaded as of December 31, 2020. The complete record of the search results was downloaded in XML format through

R's easyPubMed package. R and Python were used to extract XML data, including the publication year, abstract, research type, geographic information, and other MeSH terminology.

LDA was used to identify more specific research topics in each article. Python was used to model the topics by analyzing the abstracts of all indexed articles in the record. Topics were set at 50. The criteria for selecting the number of topics were based on perplexity, redundancy, and legibility. Based on the algorithmic calculation of topic probability, we finally determined the topic to which each article belongs. Next, we manually checked the names of each glossary based on the abstract. Finally, we used the Louvain algorithm and Gephi to perform cluster analysis to establish a topic network to determine the relationship between topics [14]. In each publication, we identified the two topics with the highest attribution probability, counted the number of simultaneous occurrences of the two topics in each document, and established links between topics.

All the original data were uploaded and publicly available, including all retrieval methods, algorithm codes, and raw literature data in this article. The literature search and download code can be obtained on R by easyPubMed package (<https://cran.r-project.org/web/packages/easyPubMed/index.html>). The R code is publicly available on GitHub (<https://github.com/christopherBelter/pubmedXML>). We have uploaded relevant Python code on GitHub (<https://github.com/yanwen0614/Medicine-Bibliometric-Analysis>). The network visualization in this article is carried out using the software package Gephi (<https://github.com/gephi/gephi>). All studies' original data can be downloaded (<https://pan.baidu.com/s/1nuw6MBvyYDOFMA5BEivtXQ>. code: cgvs). This study used publicly published data and did not need to be approved by the relevant institutional review board or ethics committee.

Results

Overview of the data

Although TNBC was proposed at the beginning of this century, it was not officially included as a MeSH term until 2014. We identified and analyzed all 5,097 publications between 2012 and 2020 (Fig. 1). On average, 141 more publications were published each year than the previous year, with an average annual growth rate of 33.5%. The high growth rate indicates that TNBC is a research hotspot, and the research is progressing rapidly.

Categories of publications on TNBC research

In order to explore the research areas regarding TNBC, we first separated the publications and divided them into nine categories according to the database-provided areas in cancer research (Fig. 2). We found that randomized clinical trials and related studies accounted for roughly 25% of total publications (Fig. 2). This high proportion of clinical trials has rarely been reported in research on other tumor types, indicating that translational research on TNBC has been carried out quickly. The proportion of review and meta-analysis was around 32% and 2.2%, respectively. As high-quality meta-analysis is generally

considered the clinical guiding research, it is reasonable to anticipate that publications on the meta-analysis of TNBC will be increased. The large number of clinical trials of TNBC has already improved and will continue to improve its clinical practice.

Uneven geographical distribution of global TNBC research

In order to further understand the global TNBC research situation, we analyzed the geographic information. We found that 88 countries or regions worldwide have published publications on TNBC (Fig. 3). The top 10 countries' publications accounted for 74.4%, indicating that the head effect is pronounced. Moreover, more than half of the publications were derived from the United States, China, Italy, and France, accounting for 23.1%, 20.5%, 6.9%, and 6.0% of all publications, respectively (Fig. 4). This phenomenon reminds us that the vast majority of the global population has not participated in TNBC research and even lacks primary data. It also suggests that current TNBC research does not contain enough data on the different genetic backgrounds, and we only have a limited view of the whole picture of TNBC in humanity.

Pathogenic mechanisms and drugs are most studied in TNBC research

A total of 4,299 MeSH terms appeared 133,259 times in all 5,097 publications, indicating that the studies covered multiple aspects. On average, 26 MeSH terms appeared in each publication showing the breadth and depth of studies on TNBC and the intense intersection between the different disciplines. The top 10 cited MeSH terms were listed in Fig. 5, with both metabolism and genetics cited more than 2,000 times so far, suggesting that the past research on TNBC was focused on exploring its molecular pathogenesis. In addition, 5 of the top 10 cited MeSH terms are directly related to medication research. Therefore, we infer that pathogenic mechanism and medication research will continue to be a focus of TNBC research in the foreseeable future.

TNBC research is mainly in three areas

The topic network analyzed by LDA and Louvain algorithm highlights the areas where interrelated topic clusters appear simultaneously and provide remarkable insights into the relationships between the essential topics of interest. We divided publications into 50 topics. As seen in Fig. 6, the circle's size represents the number of papers published, and the thickness of the line represents the degree of relevance between the two research topics. The results of the LDA analysis suggest that all TNBC-related studies are roughly divided into three clusters (Fig. 6).

The therapy plan itself is mainly studied in the TNBC treatment plan research cluster (in orange). However, this cluster also includes some studies on signaling pathways, such as the Wnt/ β -catenin signaling, PTEN, EGFR, BRCA and PARP1. This cluster is particularly close to the other two clusters, indicating that the relationship between the essential clinical integration and the TNBC basic research

was very close. We have also found that basic research can be quickly transformed into clinical practice through clinical trials, improving patient prognosis.

Among the new biomarkers research cluster (in green), receptor study, age and risk study, and immune checkpoint are extensively investigated topics. The relationship among these three topics is very close, and they are closely related to the therapy plan. In the cluster of the regulation mechanism for TNBC aggressive behavior (in purple), growth and regulation factors, apoptosis study, and nanoparticle delivery system are the three most studied topics. Furthermore, we found that magnetic resonance imaging (MRI) and ultrasound were closely related to the hypoxia study, suggesting that imaging changes may be related to hypoxia in the tumor microenvironment.

Discussion

Since the initial proposal of TNBC at the beginning of this century, substantial progress has been made in identifying therapeutic targets and more detailed tumor classification [2]. However, few studies on hospice care, patient perspective, surgical treatment of metastasis, and economics are available. Future research will focus on applying new technologies to understand the complex system of TNBC and finding more specific therapeutic targets. However, the recurrence of TNBC after long-term survival still presents new challenges for its clinical management. TNBC is currently undergoing a large number of clinical trials. The rapid clinical progress of TNBC suggests that TNBC is an excellent pioneering area in tumor research and will provide an excellent example for the diagnosis and treatment of other tumors.

It is well-known that TNBC is substantially heterogeneous. Therefore, it is necessary to classify TNBC into multiple subtypes further [15, 16]. The successful subtyping provides a solid theoretical basis for the precision therapy of TNBC [17]. Gene sequencing technology allows us to fully understand the mutation rate of TNBC, which is about 1.68bp/Mb [18]. Mutations occur in genes in multiple key signaling pathways such as PI3K/Akt/mTOR pathway, RAS/RAF/MEK pathway, JAK/STAT pathway, DNA repair pathway, and cell cycle checkpoint [19–21]. Therefore, a variety of drugs targeting the signal pathways are currently undergoing clinical trials. Some inhibitors have been used as potential medications for TNBC treatment, including PI3K, MEK, PARP, EGFR, VEGF, and AR inhibitors [22]. On the other hand, studies on operations and radiotherapy, especially for re-operations related to local-regional recurrence risk or distant metastasis, were rarely reported. Actually, many studies suggest that surgery has an essential role in treating distant metastases of cancers, such as colorectal cancer [23]. In addition, many studies on other kinds of cancers, including pancreatic cancer and colorectal cancer, demonstrated that the tumor microenvironment, especially the extracellular matrix, has been found to play an essential role in cancer metastasis, local recurrence, and chemotherapeutic drug resistance [23]. Many potential drugs are used due to their ability to target the extracellular matrix, such as PEGPH20 (an enzyme that targets matrix hyaluronic acid), pegilodecakin (a PEGylated IL-10) [25, 26]. However, so far, the study on extracellular matrix in TNBC is insufficient. It is hoped that the study on TNBC extracellular matrix is likely to play an essential role in the future.

In the present study, we quantitatively analyzed all 5,097 publications on TNBC through multiple machine learning algorithms for the first time. We found that the research on TNBC mainly focuses on three clusters: TNBC treatment plan research, the new biomarkers research and the regulation mechanism for TNBC aggressive behavior. In the cluster of regulation mechanisms for TNBC aggressive behavior, we found an exciting part of a close connection between hypoxia and imaging. Hypoxia refers to malignant cells' inability to receive enough oxygen as the tumor expands [27]. Low oxygen levels stabilize hypoxia-inducible factors (HIFs), which in turn regulates the transcriptional activation of a group of genes, allowing cells to survive under hypoxic conditions [28]. On the one hand, we speculate that these manifestations will be evident in imaging, which will improve the early diagnosis of TNBC and the detection of possible distant metastases in the future. On the other hand, the hypoxic properties of TNBC bring new hopes for treatment with the aid of imaging such as MRI and ultrasound, which are highly sensitive to the hypoxic microenvironment of TNBC. Studies have found that carbonic anhydrase-IX-directed albumin nanoparticles have a deadly effect on hypoxia-mediated T lymphocyte-negative breast cancer cells and can be visualized in a real-time fashion [29]. At present, new hypoxia-inducible factor-2 α antagonists such as PT2385 have shown encouraging results in phase I clinical trials of previously treated advanced clear cell renal cell carcinoma [30]. More than two-thirds of patients reported clinical benefits with acceptable toxicity profiles. PT2977, a more effective second-generation hypoxia-inducible factor-2 α antagonist, is the subject of several ongoing solid tumor clinical trials (NCT02974738, NCT03634540, NCT03401788) [31]. With the in-depth study on TNBC hypoxia aided by its imaging, developing new treatment strategies for cancers is possible.

Although the research on TNBC has made significant progress in many aspects, the present research also found some research deficiencies on TNBC. First of all, there are considerable differences in geographic-related research. It has been found that individuals with different genetic backgrounds have different responses to TNBC, survival rates, and mortality rates [32, 33]. However, the global distribution of TNBC research is uneven, and the research on populations with different genetic backgrounds is insufficient. For example, only 61 publications in the PubMed database are related to African descent. Strengthening TNBC research on individuals with different genetic backgrounds could facilitate our understanding of TNBC from different perspectives [34]. Secondly, there is a lack of research on TNBC from patient's perspectives, e.g., health economics, hospice care. Although, at present, the 5-year overall survival rate of most tumors has been greatly improved, helping tumor patients with psychological issues re-enter society will become a new important research topic [35]. Compared with other breast cancer subtypes, TNBC patients are more likely to relapse and metastasize, resulting in more significant pressure, both mentally and economically, on patients and their families. Studies on patients with more prolonged survival can better understand TNBC and even other long-term survival tumors [36]. We will face more challenges for patients with a long survival period of 5–10 years in the future [37].

There are some limitations in the present study. Firstly, besides PubMed, several other databases, including Scopus, Web of Science, and Embase, could be used for bibliometric research. Although PubMed contains the highest quality peer-reviewed research and excludes irrelevant, non-peer-reviewed papers, the literature will provide detailed and comprehensive knowledge if other databases are explored

simultaneously. Secondly, we considered that all publications tend to publish more positive research results. Negative results and clinical participants' perspectives are naturally more difficult to be published.

In summary, with the further expansion of the publishing industry and the explosion of the number of information dimensions, the vast array of documents will make the understanding of TNBC more difficult. Researchers are lost in the information as it is not easy to obtain the information they want. With the development of complete medical record texts, publication databases, and improved algorithms, it is reasonable for machine learning to play a more active auxiliary role in future clinical practice. Machine learning and natural language processing may be an extremely effective new tool for scientists, who will extract objective and comprehensive clues from large amounts of data. The data presented in this study will hopefully help scientists understand the current status of TNBC research and design more relevant basic and clinical research projects.

Concluding Remarks

We analyzed 5,097 TNBC publications through the NLP method. We summarized the current research status from a macro perspective and found that the current research is mainly concentrated in three clusters. TNBC research direction shows some insufficiencies, especially in long-term survival-related research and lack of research from patients' perspectives. TNBC is a good forerunner in cancer research.

Declarations

Ethics approval and consent to participate

This study used publicly published data and did not need to be approved by the relevant institutional review board or ethics committee.

Consent for publication

Not applicable

Availability of data and materials

The original contributions presented in the study are included in the article. Further, inquires can be directed to the corresponding authors.

Competing interests

The authors declare no potential conflicts of interest.

Funding

This work was supported by funds from China Scholarship Council in the form of a scholarship to K. Wang(202006370023), Guangzhou Institute of Pediatrics/Guangzhou Women and Children's Medical

Center (4001013-04, 5001-4001008), and National Natural Scientific Foundation of China (30771122, 81472496, 81872167).

Authors' contributions

K.W., L.Z., M.L., X.D.: Concept and design;

K.W., C.Z., L.X., D.D., F.T., Q.P., Y.L., D.W.: Development of methodology;

K.W., C.Z., L.X., D.D., F.T., Q.P.: Acquisition of data;

K.W., Y.L., D.W.: Analysis and interpretation of data;

K.W., L.Z., M.L., X.D.: Writing, review and/or revision of the manuscript

Acknowledgements

We would like to express our gratitude to Mr. Wen Yan, who helped us with the programming and coding.

Authors' information (optional)

Kangtao Wang, MD student, Email: kangtao.wang@stud.uni-heidelberg.de

Chanjuan Zheng, PhD student, Email: zhengchanjuan@hunnu.edu.cn

Lian Xue, MSB student, Email: xuelian@hunnu.edu.cn

Dexin Deng, MD student: 571790201@qq.com

Fengbo Tan, MD, Email: fengbotan@csu.edu.cn

Qian Pei, MD, Email: peiqian1221@163.com

Yujuan Li, MD, Email: liyuqiang22@yeah.net

Dan Wang, MD student, Email: wangdan19931006@outlook.com

Liang Zeng, PhD Scientist, Email: zlxx03@126.com

Ming Li, PhD, Scientist, Email: liming@csu.edu.cn

Xiyun Deng, Group Leader, Email: dengxiyunmed@hunnu.edu.cn

References

1. Banerjee S, Tian T, Wei Z, Shih N, Feldman MD, Peck KN, et al. Distinct Microbial Signatures Associated With Different Breast Cancer Types. FRONT MICROBIOL. 2018;9:951.

2. Perou CM, Sørli T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *NATURE*. 2000;406(6797):747–52.
3. Fahad UM. Breast Cancer: Current Perspectives on the Disease Status. *ADV EXP MED BIOL*. 2019;1152:51–64.
4. Yao H, He G, Yan S, Chen C, Song L, Rosol TJ, *et al*: **Triple-negative breast cancer: is there a treatment on the horizon?** *Oncotarget* 2017, **8**(1):1913–1924.
5. Pareja F, Reis-Filho JS. Triple-negative breast cancers — a panoply of cancer types. *NAT REV CLIN ONCOL*. 2018;15(6):347–8.
6. Lyons TG. Targeted Therapies for Triple-Negative Breast Cancer. *CURR TREAT OPTION ON*. 2019;20(11):82.
7. Belter CW. Bibliometric indicators: opportunities and limits. *Journal of the Medical Library Association: JMLA*. 2015;103(4):219–21.
8. Teles RHG, Moralles HF, Cominetti MR. Global trends in nanomedicine research on triple negative breast cancer: a bibliometric analysis. *Int J Nanomedicine*. 2018;**Volume** 13:2321–36.
9. Buchlak QD, Esmaili N, Leveque J, Farrokhi F, Bennett C, Piccardi M, et al. Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. *NEUROSURG REV*. 2020;43(5):1235–53.
10. Feng C, Wu Y, Gao L, Guo X, Wang Z, Xing B: **Publication Landscape Analysis on Gliomas: How Much Has Been Done in the Past 25 Years?** *FRONT ONCOL* 2020, **9**:1463.
11. Sommeria Klein G, Zinger L, Coissac E, Iribar A, Schimann H, Taberlet P, et al. Latent Dirichlet Allocation reveals spatial and taxonomic structure in a DNA-based census of soil biodiversity from a tropical forest. *MOL ECOL RESOUR*. 2020;20(2):371–86.
12. Tran BX, Latkin CA, Sharafeldin N, Nguyen K, Vu GT, Tam WWS, et al. Characterizing Artificial Intelligence Applications in Cancer Research: A Latent Dirichlet Allocation Analysis. *JMIR Medical Informatics*. 2019;7(4):e14401.
13. Wang K, Feng C, Li M, Pei Q, Li Y, Zhu H, et al. A bibliometric analysis of 23,492 publications on rectal cancer by machine learning: basic medical research is needed. *Therap Adv Gastroenterol*. 2020;13:320843838.
14. Traag VA. Faster unfolding of communities: Speeding up the Louvain algorithm. *PHYS REV E*. 2015;92(3):32801.
15. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J CLIN INVEST*. 2011;121(7):2750–67.
16. Deng X, Faqing T, Rosol TJ. Triple-Negative Breast Cancer. In: *Triple-Negative Breast Cancer*. World Scientific; 2020. pp. 21–70.
17. Lee Y, Oh MH, Go J, Han K, Choi S. Molecular subtypes of triple-negative breast cancer: understanding of subtype categories and clinical implication. *GENES GENOM*. 2020;42(12):1381–7.

18. Mittendorf EA, Philips AV, Meric-Bernstam F, Qiao N, Wu Y, Harrington S, et al. PD-L1 Expression in Triple-Negative Breast Cancer. *CANCER IMMUNOL RES*. 2014;2(4):361–70.
19. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *NATURE*. 2012;486(7403):395–9.
20. Vanhaesebroeck B, Guillermet-Guibert J, Graupera M, Bilanges B. The emerging mechanisms of isoform-specific PI3K signalling. *NAT REV MOL CELL BIO*. 2010;11(5):329–41.
21. Sun X, Wang M, Wang M, Yu X, Guo J, Sun T, et al. Metabolic Reprogramming in Triple-Negative Breast Cancer. *FRONT ONCOL*. 2020;10:428.
22. Islam R, Lam KW. Recent progress in small molecule agents for the targeted therapy of triple-negative breast cancer. *EUR J MED CHEM*. 2020;207:112812.
23. Wang K, Feng C, Li M, Pei Q, Li Y, Zhu H, et al. A bibliometric analysis of 23,492 publications on rectal cancer by machine learning: basic medical research is needed. *THER ADV GASTROENTER*. 2020;13:1320574573.
24. Kim DJ, Yang J, Seo H, Lee WH, Ho Lee D, Kym S, et al. Colorectal cancer diagnostic model utilizing metagenomic and metabolomic data of stool microbial extracellular vesicles. *SCI REP-UK*. 2020;10(1):2860.
25. Gourd E. PEGPH20 for metastatic pancreatic ductal adenocarcinoma. *The Lancet Oncology*. 2018;19(2):e81.
26. Hecht JR, Lonardi S, Bendell J, Sim H, Macarulla T, Lopez CD, et al: **Randomized Phase III Study of FOLFOX Alone or With Pegilodexakin as Second-Line Therapy in Patients With Metastatic Pancreatic Cancer That Progressed After Gemcitabine (SEQUOIA)**. *J CLIN ONCOL* 2021:20–2232.
27. Wang J, Wang Y, Xing P, Liu Q, Zhang C, Sui Y, et al. Development and validation of a hypoxia-related prognostic signature for breast cancer. *ONCOL LETT*. 2020;20(2):1906–14.
28. Chaturvedi P, Gilkes DM, Takano N, Semenza GL: **Hypoxia-inducible factor-dependent signaling between triple-negative breast cancer cells and mesenchymal stem cells promotes macrophage recruitment**. *Proceedings of the National Academy of Sciences* 2014, **111**(20):E2120-E2129.
29. Goggins E, Kakkad S, Mironchik Y, Jacob D, Wildes F, Krishnamachary B, et al: **Hypoxia Inducible Factors Modify Collagen I Fibers in MDA-MB-231 Triple Negative Breast Cancer Xenografts**. *NEOPLASIA* 2018, **20**(2):131–139.
30. Courtney KD, Ma Y, Diaz De Leon A, Christie A, Xie Z, Woolford L, et al. HIF-2 Complex Dissociation, Target Inhibition, and Acquired Resistance with PT2385, a First-in-Class HIF-2 Inhibitor, in Patients with Clear Cell Renal Cell Carcinoma. *CLIN CANCER RES*. 2020;26(4):793–803.
31. Xu R, Wang K, Rizzi JP, Huang H, Grina JA, Schlachter ST, et al: **3-[(1S,2S,3R)-2,3-Difluoro-1-hydroxy-7-methylsulfonylindan-4-yl]oxy-5-fluorobenzonitrile (PT2977), a Hypoxia-Inducible Factor 2 α (HIF-2 α) Inhibitor for the Treatment of Clear Cell Renal Cell Carcinoma**. *J MED CHEM* 2019, **62**(15):6876–6893.
32. Philipovskiy A, Dwivedi AK, Gamez R, McCallum R, Mukherjee D, Nahleh Z, et al. Association between tumor mutation profile and clinical outcomes among Hispanic Latina women with triple-negative

breast cancer. PLOS ONE. 2020;15(9):e238262.

33. Thyagarajan S, Zhang Y, Thapa S, Allen MS, Phillips N, Chaudhary P, et al. Comparative analysis of racial differences in breast tumor microbiome. *Sci Rep*. 2020;10(1):14116.
34. Hercules SM, Hercules JC, Ansari A, Date S, Skeete D, Smith CS, et al. High triple-negative breast cancer prevalence and aggressive prognostic factors in Barbadian women with breast cancer. *CANCER-AM CANCER SOC*. 2020;126(10):2217–24.
35. Watkins CC, Kamara KI, Hamilton JB, Kozachik SL, Gaston-Johansson F. Differences in Coping Among African American Women With Breast Cancer and Triple-Negative Breast Cancer. *ONCOL NURS FORUM*. 2017;44(6):689–702.
36. Mediratta K, El-Sahli S, D'Costa V, Wang L. **Current Progresses and Challenges of Immunotherapy in Triple-Negative Breast Cancer.** *Cancers (Basel)* 2020, 12(12).
37. Ertas G, Basal FB, Ucer AR, Benzer E, Altundag MB, Demirci U, et al. Clinical features of metaplastic breast carcinoma: A single-center experience. *J CANCER RES THER*. 2020;16(6):1229–34.

Figures

Figure 1

Publications numbers per year citing the MeSH term: Triple Negative Breast Neoplasms. Due to the publication delay, the 2020 data were estimated based on the trend in previous years.

Figure 3

Global distribution of publications citing TNBC as the MeSH term in PubMed.

Figure 4

Top 10 countries in the world with the most significant number of publications on TNBC. The 2012 data were not shown due to insufficient number of publications available.

Figure 5

Top 10 other MeSH terms most frequently cited in the publications on TNBC.