



Original Article

How Should Biobanks Collect Biosamples for Clinical Application? A 20-year Biomarker-related Publication and Patent Trend Analysis

Jae-Eun Lee*

Division of Biobank for Health Sciences, Center for Genome Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Cheongju, Korea

ABSTRACT

Article history:

Received: May 9, 2018
Revised: May 23, 2018
Accepted: May 28, 2018

Keywords:

biomarkers,
patent, publication,
trends

Objectives: This study was designed to analyze biomarker-related publications and patent trends which biobanks could consider in planning biosample collections for biomarker research.

Methods: Publications and patents containing the term “biomarker” in the title published between 1998 to 2017 were retrieved using Scopus database and Google Patents search engine.

Results: Over the last 20 years there has been a steady increase in biomarker-related publications and patents; however this has slowed for patents over the last few years. Publications in 2017 that contained blood, serum, and plasma search terms in the abstract accounted for 50%, and serum as a search term in the title and abstract was more numerous than those containing blood, plasma, tissue, or urine. Blood-related patents were the most common patent in the last 10 years, and accounted for 110 patents in 2017. Biomarker-related publications since 2010 containing RNA and protein search terms in the title and abstract, were more numerous than those containing DNA and metabolite search terms. More than 27% of biomarker-related publications in 2017 and 21% of biomarker-related patents were associated with cancer.

Conclusion: The results of this study will help biobanks establish a biosample collection strategy for clinical application.

<https://doi.org/10.24171/j.phrp.2018.9.3.04>
pISSN 2210-9099 eISSN 2233-6052

©2018 Korea Centers for Disease Control and Prevention. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Recently, medicine has been moving toward a more personalized, precision medicine, based on a patient's genetic, environmental, and clinical characteristics. For the successful implementation of personalized, precision medicine it is imperative that new diagnostic and prognostic biomarkers specific to disease subtypes and individual patients are continually discovered [1]. Current and prospective biomarkers of disease include DNA, RNA, protein, and metabolite biomarkers [2-6] which could be used for early detection, prognosis, prediction of treatment response, or monitoring of disease recurrence [2,7]. Protein assays are the most widely

used test for biomarkers of disease that are routinely used in clinical analysis [8], but at the forefront of this research is the challenge to discover new DNA, RNA, and metabolite biomarkers as well as new protein biomarkers for various diseases.

Biobanks need to secure biosamples (such as human blood, serum, plasma, tissue, and urine) suitable for biomarker discovery and clinical application and to do this, an understanding of the trends in biomarker-related publications and patents would be highly relevant.

Geyer et al reported the trends in the literature associated with plasma biomarker studies and mass spectrometry-based proteomics [8]. In this study, the trends in biomarker-related

*Corresponding author: Jae-Eun Lee

Division of Biobank for Health Sciences, Center for Genome Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Cheongju, Korea
E-mail: eunwin21@korea.kr

©2018 Korea Centers for Disease Control and Prevention. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

publications and patents in the last 20 years depending on the type of biosample, type of analyte, or biomarker discovery purposes continually increase. The Scopus database and Google Patents search engine were used for literature and patent analyses, respectively. The Scopus database is 1 of the most common bibliographic databases and has more than 59,000 references containing the term “biomarker” in the title. This study will provide information that could help biobanks collect biosamples suitable for biomarker discovery and clinical application.

Materials and Methods

1. Publication trend analysis

Scientific publication trends related to biomarker were analyzed using Elsevier’s Scopus database (www.scopus.com) in March 2018. Publications containing the term “biomarker” in the title which were published from 1998 to 2017 were retrieved. For the literature published in 2017, some pertinent information such as the subject area, document type, source title, author’s country, and source type was investigated. All publications belonged to 1 or more subject areas. Publications containing search terms for various biosamples (blood, serum, plasma, tissue, and urine) and analytes (DNA, RNA, microRNA, protein, and metabolite) in the title or the abstract were extracted from the retrieved biomarker-related publications, respectively. Furthermore, publications containing search terms related to biomarker discovery purposes (cancer, prognosis, therapy, diagnosis, and predictive) in the title or abstract were extracted from the retrieved biomarker-related publications. The search terms were queried in the database of publication titles and abstracts, and asterisks were used to retrieve publications including derivations of the search strings. The document types of the searched publications included article, article in press, abstract report, book, book chapter, conference paper, editorial, letter, note, review, and short survey.

The search strings were as follows: biomarker (“biomarker*”), blood (“blood” OR “whole blood”), serum (“serum” OR “sera”), plasma (“plasma”), tissue (“tissue*”), urine (“urine” OR “urinary”), DNA (“DNA” OR “SNP” OR “genome*” OR “genomic*” OR “single nucleotide polymorphism*” OR “copy number variation*”), RNA (“RNA” OR “mRNA*” OR “miRNA*” OR “microRNA*” OR “gene expression*” OR “transcript*” OR “transcriptome*”), microRNA (“miRNA*” OR “microRNA*”), protein (“protein*” OR “proteome*” OR “proteomic*”), metabolite (“metabolite*” OR “metabolome*” OR “metabolomic*”), cancer (“cancer” OR “carcinoma”), prognosis (“prognosis” OR “prognostic”), therapy (“therapy” OR “therapeutic” OR “treatment”), diagnosis (“diagnosis” OR

“diagnostic”), and predictive (“predictive” OR “predict*”).

2. Patent trend analysis

Biomarker-related patent trends were analyzed using the Google Patents search engine (Google Inc., Mountain View, California, USA) in March 2018. Patents containing the term “biomarker” in the title which was published between 1998 to 2017 were retrieved. Patents containing search terms for various biosamples (blood, serum, plasma, tissue, and urine) in the abstract were extracted from the retrieved biomarker-related patents. The search strings, which were queried in the database of patent titles or abstracts, were the same as those used for the publication trend analysis.

3. Statistical analysis

Linear regression analysis was performed to assess annual publication numbers and patent trends associated with biomarkers, using SPSS, version 18.0 (SPSS, Chicago, IL, USA). In this analysis, the number of publications and patents published each year was used as a dependent variable and the year of publication was analyzed as an independent variable. Statistical significance was reached when $p < 0.05$.

Results

1. Publications related to biomarker

Publications containing the term “biomarker” in the title which were published between 1998 and 2017 were retrieved using the Scopus database. The number of biomarker-related publications has increased steadily over the past 20 years (Figure 1) with a rapid increase since 2004 from 777 to 7,312 in 2017.

Classifying the retrieved publications by subject area, showed that “medicine” was the major subject area (4,712; 64.4%), followed by “biochemistry, genetics, and molecular biology” (2,496; 34.1%) (Table 1). The retrieved publications were mostly articles (5,456; 74.6%) and generally published by researchers in the United States (2,006; 27.4%) and China (1,318; 18.0%).

2. Biomarker-related publications by the type of biosample

Publications containing search words for various biosamples such as blood, serum, plasma, tissue, and urine that were contained in either the title or the abstract were extracted from all the retrieved biomarker-related publications (Figure 2). The number of blood, serum, plasma, tissue, and urine-related publications (except for publications containing the term “tissue” in the title) showed increasing trends over the past 20 years, respectively. Amongst them, serum-related publications

Table 1. Top10 of the subject area, document type, source title, author's country, and source type of biomarker-related publications in 2017 (n = 7,312).

Subject Area	Document Type	Source Title	Country	Source Type					
Medicine	4,712	Article	5,079	Oncotarget	257	United States of America	2,006	Journals	7,078
Biochemistry, Genetics and Molecular Biology	2,496	Review	1,097	Scientific Reports	188	China	1,318	Book Series	102
Pharmacology, Toxicology and Pharmaceutics	555	Article in Press	377	Plos One	159	United Kingdom	643	Conference Proceedings	72
Neuroscience	541	Editorial	196	Biomarkers In Medicine	59	Germany	504	Books	58
Chemistry	466	Letter	168	Journal Of Alzheimer's Disease	54	Italy	498	Trade Publications	2
Agricultural and Biological Sciences	449	Book Chapter	123	International Journal Of Molecular Sciences	50	Spain	364		
Immunology and Microbiology	393	Conference Paper	123	International Journal Of Clinical And Experimental Pathology	36	France	342		
Environmental Science	367	Note	104	Methods In Molecular Biology	36	Canada	323		
Nursing	221	Short Survey	39	Oncology Letters	33	Japan	303		
Multidisciplinary	218	Book	5	Biomed Research International	32	Netherlands	269		

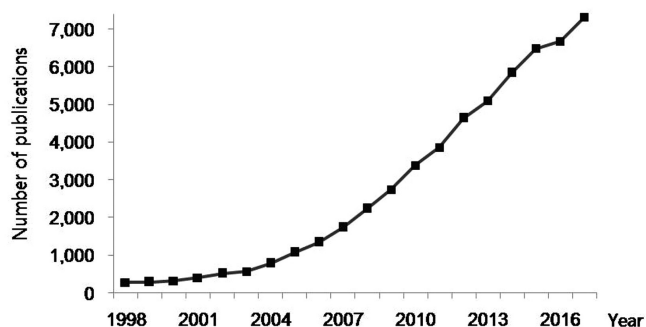


Figure 1. Annual number of publications containing the term "biomarker" in the title which published from 1998 to 2017.

were consistently the most common search terms in the title since 2005 and the abstract since 2009.

The number of publications in 2017 containing the term "serum" in the title and abstract was 484 and 1,474, respectively; Of the 7,312 biomarker-related publications, the number of serum-related publications accounted for 6.6% in the title and 20.2% in the abstract, and was about 2 times

more than the number of plasma-related publications. In title searches in 2017, urine-related publications were the second highest followed by plasma, blood, and tissue. In case of abstract searches, blood-related publications were the second most common, followed by tissue, plasma, and urine.

3. Biomarker-related publications by the type of analyte

Publications containing search words for various analytes (such as DNA, RNA, protein, and metabolite) in the title or the abstract were extracted from all the retrieved biomarker-related references, respectively (Figure 3). Publications with RNA or protein search terms were the most common in both title and abstract searches since 2010.

The number of publications that contained RNA search terms in the title has increased sharply since 2010 and has been higher than other publications since 2015. The number of RNA-related publications in 2017 was 599 and was the most common. The number of publications in 2017 that contained RNA search terms in the abstract was 1,311. Of the 7,312 biomarker-related publications, RNA-related publications

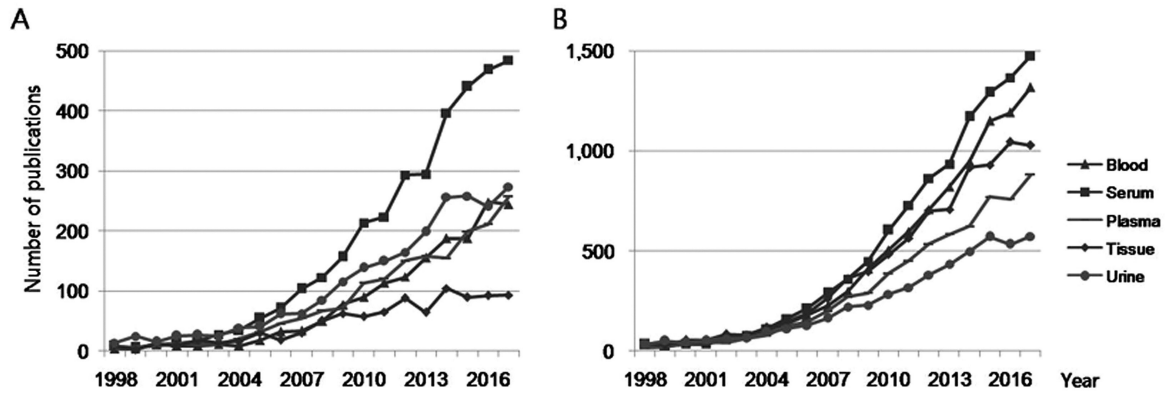


Figure 2. Publications containing various biosample search terms in (A) the title, (B) the abstract. These were extracted from the biomarker-related publications.

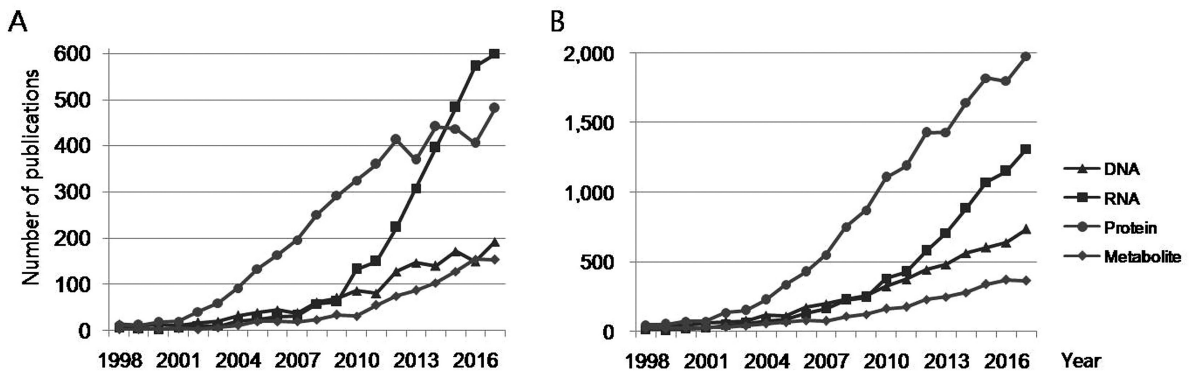


Figure 3. Literature containing various analyte search terms in the title (A) and abstract (B). These were extracted from the biomarker-related publications.

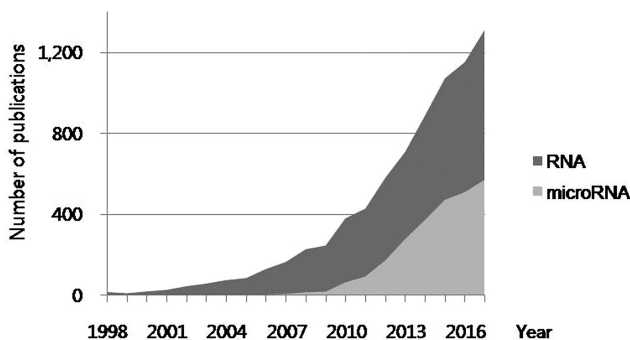


Figure 4. Publications containing RNA and microRNA search terms in the abstract. These were extracted from the biomarker-related publications.

accounted for 8.2% in the title and 17.9% in the abstract. The number of publications in 2017 that contained protein search terms in the title and the abstract was 481 and 1,979, respectively. Of the biomarker-related publications, protein-related publications accounted for 6.6% in the title and 27.1% in the abstract. The number of RNA and protein-related publications were double that of other publications.

In abstract searches, microRNA-related publications started to appear in 2006 and have dramatically increased every year. Since 2014, microRNA-related publications are in excess of 40% of RNA-related publications (Figure 4).

4. Biomarker-related publications by biomarker discovery purposes

Publications containing search terms related to biomarker discovery purposes, in the title or the abstract were extracted

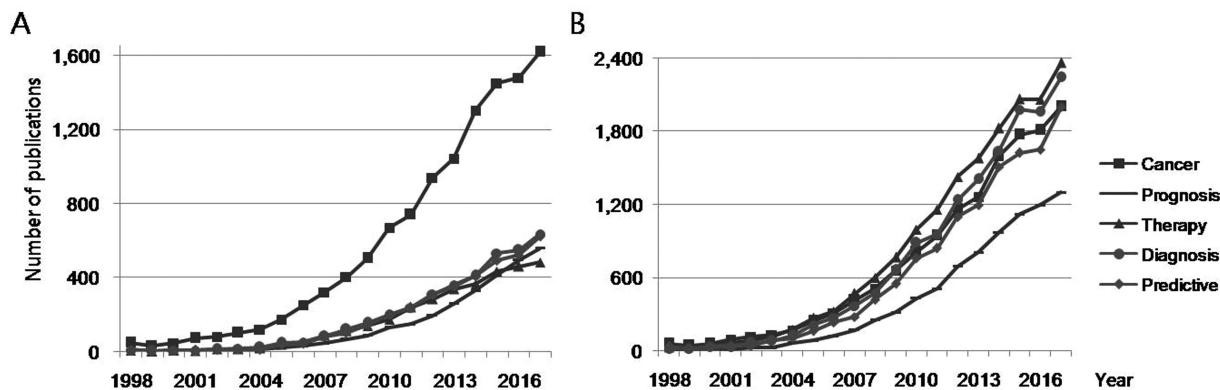


Figure 5. Literature containing search terms related to biomarker discovery purposes in the title (A) and abstract (B). These were extracted from the biomarker-related publications.

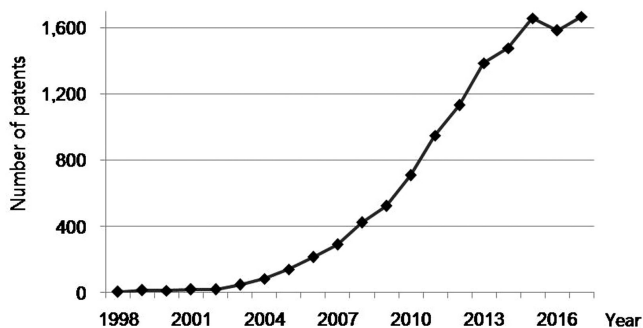


Figure 6. Annual number of patents containing the term “biomarker” in the title which published from 1998 to 2017.

from all the retrieved biomarker-related publications (Figure 5). Publications over the last 20 years containing “cancer” as a search term in the title were about 3 times those containing other search terms (prognosis, therapy, diagnosis, and predictive). Conversely, the number of publications containing “cancer” as a search word in the abstract was generally similar to the number of other publications (except for publications on “prognosis” as a search word). Publications in 2017 containing “cancer” as a search term in the title accounted for 1,617 publications, and in the abstract accounted for 2,002 publications (22.1% and 27.4% of biomarker-related publications), respectively.

5. Patents related to biomarker

A search for patents containing the term “biomarker” in the title published between 1998 and 2017 was carried out using Google Patents search engine (<https://patents.google.com>). Biomarker-related patents showed an increasing trend

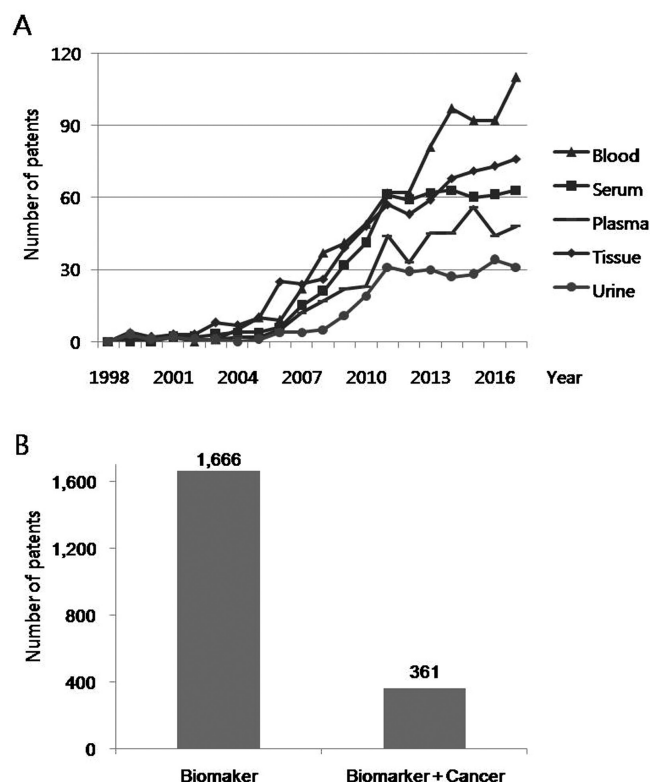


Figure 7. Annual number of patents related to blood, serum, plasma, tissue, and urine (A) and cancer-related patents published in 2017 (B). These were extracted from the biomarker-related patents.

over the past 20 years, but has slowed in recent years (Figure 6). There were 1,666 biomarker-related patents published in 2017. Patents containing search terms for blood, serum, plasma, tissue, and urine in the abstract were extracted from the retrieved biomarker-related patents (Figure 7). In general, blood-related patents were the most common and urine-

related patents were the least. Blood, tissue, serum, plasma, and urine-related patents were 110, 76, 63, 48, and 31 in 2017, respectively. Amongst biomarker-related patents that were published in 2017, 361 (21.7%) patents contained “cancer” as a search term in the abstract (Figure 7).

Discussion

The aim of this study was to analyze literature and patent trends; the results of which may be useful to biobanks when considering biosample acquisition and biomarker research. Biomarker-related publications retrieved using the Scopus database, showed a clear upward trend every year as did biomarker-related patents (retrieved using the Google Patents search engine). Of the biomarker-related publications in 2017, the proportion of publications containing blood-based biofluids (blood, serum, and plasma) search terms in the abstract was 50.2%. Clinical testing of blood is the most common diagnostic procedure [8] and is relatively simple. It is generally assumed that 70% of diagnoses are made based on the results from clinical testing of blood [8]. Biomarker-related publication trends when categorized into type of biosample, might reflect the clinical setting. A large number of biomarker-related references published in 2017 were from researchers in the United States (27.4%) and China (18.0%) suggesting these 2 countries are at the forefront of biomarker research.

Of the retrieved biomarker-related publications the number of serum-related publications was recently greater than blood, plasma, tissue, and urine-related publications. In title searches in 2017, there were twice as many serum-related publications as blood and tissue-related publications. In abstract searches in 2017, there were twice as many serum-related publications as urine-related publications. Of the retrieved biomarker-related patents, the number of serum-related patents was higher than plasma and urine-related patents. These publication and patent trends should be considered when determining type and quantity of biosample collection.

Analysis of the retrieved biomarker-related publications by the type of analyte, publications containing RNA and protein search strings in the title and abstract, were much more than those containing DNA and metabolite search strings and have been since 2010. RNA and protein profiles can vary depending on pre-analytical (collection, processing, transport, and storage conditions of biosamples) and analytical variables (type of analyte and method of analysis) [9-12]. Therefore, appropriate conditions for the collection and handling of the biosample may be different [10]. Biobanks could secure biosamples from several hospitals or medical facilities according to the researcher's detail requests (participant's information, type and amount of biosamples, conditions for collecting, processing,

transporting, and storing biosamples, and the biosample collection period), and provide biosamples suitable for specific research. After a researcher's application to access biosamples is complete, biobanks could start to collect the requested biosamples creating a “custom-made biobanking system.” This would be an economical use of storage facilities and equipment at biobanks.

The biobank should secure information on the history of the biosample collection and handling [10, 13]; information on serum and plasma should be inclusive of timestamps (date and time) of blood collection (phlebotomy), transport to the laboratory, separation of serum and plasma, and freezer storage. Based on this information, the researcher can select and use biosamples suitable for their research purpose.

In routine clinical analysis, protein assays are the most widely used (42% of total analysis), followed by cellular and molecular analysis [8] however, clinical application of microRNA analysis has recently been receiving a lot of attention. In abstract searches, microRNA-related publications have exceeded 40% of RNA-related publications since 2014 and for the first time, in 2008 the presence of microRNAs in serum and plasma was reported [14-16]. MicroRNA profiles in serum or plasma may change during disease development and progression [14, 16, 17]. Serum and plasma microRNAs are stable against endogenous RNase activity [16-18] but may be affected by pre-analytical variables such as dietary conditions, clotting time, and hemolysis [11, 19-21] and so for these reasons standardization is crucial. Standard operating procedures (SOPs) needs to be applied for the collection and handling of the biosamples. Marzi et al developed microRNA SOPs for serum sampling and analysis for the early detection of lung cancer [11]. One of the SOPs included a protocol that discards the first 3 mL of blood for serum sampling to avoid contamination with skin cells. The blood is collected in tubes containing clot activator, is left for 2-3 hours at room temperature and the serum is isolated immediately after centrifugation. Serum aliquots were immediately snap-frozen on dry ice. The biobank may need to develop SOPs for clinical application depending on the type of biosample and specific intended uses.

In 2017, more than 27% of biomarker-related publications and 21% of biomarker-related patents were associated with cancer. Cancer is one of the leading causes of death worldwide. In the United States of America in 2014, deaths due to cancer represented 22.5% of all deaths [22] and in Korea in the same year 28.6% of all deaths were due to cancer [23]. Biobanks should consider research and development trends, pre-analytical and analytical variables by the type of biosample, health-related lifestyle and environmental conditions, and incidence and death rate by the type of disease, to establish biobanking strategies for future biomedical research. This

could be defined as “evidence-based biobanking.”

This study shows trends in biomarker-related publications and patents by the type of biosample, analyte, or biomarker for discovery purposes and this information could help biobanks determine the quantity of biosample collection by the type of biosample. The observations in this study could also be considered in establishing SOPs, and biobanking system to secure biosamples for clinical application.

Conflicts of Interest

The author declares that they have no conflicts of interest.

References

- [1] Niklinski J, Kretowski A, Moniuszko M, et al. Systematic biobanking, novel imaging techniques, and advanced molecular analysis for precise tumor diagnosis and therapy: The Polish MOBIT project. *Adv Med Sci* 2017;62(2):405-13.
- [2] Armand-Labit V, Pradines A. Circulating cell-free microRNAs as clinical cancer biomarkers. *Biomol Concepts* 2017;8(2):61-81.
- [3] Cheng X, Zhang L, Chen Y, et al. Circulating cell-free DNA and circulating tumor cells, the “liquid biopsies” in ovarian cancer. *J Ovarian Res* 2017;10(1):75.
- [4] Havelund JF, Heegaard NHH, Færgeman NJK, et al. Biomarker Research in Parkinson's Disease Using Metabolite Profiling. *Metabolites* 2017;7(3). pii: E42.
- [5] Zamay TN, Zamay GS, Kolovskaya OS, et al. Current and Prospective Protein Biomarkers of Lung Cancer. *Cancers (Basel)* 2017;9(11). pii: E155.
- [6] Zhao Y, Gao Q, Pei L et al. Current status and future prospects of biomarkers in the diagnosis of hepatocellular carcinoma. *Int J Biol Markers* 2017;32(4):e361-9.
- [7] Tang Y, Qiao G, Xu E, et al. Biomarkers for early diagnosis, prognosis, prediction, and recurrence monitoring of non-small cell lung cancer. *Oncotargets Ther* 2017;10:4527-34.
- [8] Geyer PE, Holdt LM, Teupser D, et al. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol* 2017;13(9):942.
- [9] Betsou F, Bulla A, Cho SY, et al. Assays for Qualification and Quality Stratification of Clinical Biospecimens Used in Research: A Technical Report from the ISBER Biospecimen Science Working Group. *Biopreserv Biobank* 2016;14(5):398-409.
- [10] Lee JE, Kim YY. Impact of Preanalytical Variations in Blood-Derived Biospecimens on Omics Studies: Toward Precision Biobanking? *OMICS* 2017;21(9):499-508.
- [11] Marzi MJ, Montani F, Carletti RM, et al. Optimization and Standardization of Circulating MicroRNA Detection for Clinical Application: The miR-Test Case. *Clin Chem* 2016;62(5):743-54.
- [12] Zimmerman LJ, Li M, Yarbrough WG, et al. Global stability of plasma proteomes for mass spectrometry-based analyses. *Mol Cell Proteomics* 2012;11(6):M111.014340.
- [13] Lee JE, Kim JW, Han BG, et al. Impact of Whole-Blood Processing Conditions on Plasma and Serum Concentrations of Cytokines. *Biopreserv Biobank* 2016;14(1):51-5.
- [14] Chen X, Ba Y, Ma L, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 2008;18(10):997-1006.
- [15] Lawrie CH, Gal S, Dunlop HM, et al. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol* 2008;141(5):672-5.
- [16] Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 2008;105(30):10513-8.
- [17] Skog J, Würdinger T, van Rijn S, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol* 2008;10(12):1470-6.
- [18] Arroyo JD, Chevillet JR, Kroh EM, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci U S A* 2011;108(12):5003-8.
- [19] Kirschner MB, Edelman JJ, Kao SC, et al. The Impact of Hemolysis on Cell-Free microRNA Biomarkers. *Front Genet* 2013;4:94.
- [20] McDonald JS, Milosevic D, Reddi HV, et al. Analysis of circulating microRNA: preanalytical and analytical challenges. *Clin Chem* 2011;57(6):833-40.
- [21] Pritchard CC, Kroh E, Wood B, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila)* 2012;5(3):492-7.
- [22] Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin* 2017;67(1):7-30.
- [23] Jung KW, Won YJ, Oh CM, et al. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2014. *Cancer Res Treat* 2017;49(2):292-305.