The Open Biomarkers Journal

Content list available at: https://openbiomarkerjournal.com

RESEARCH ARTICLE

Bibliometric and Semantic Analysis of the Global Research on Biomarkers in Personalized Medicine

Aida Khakimova[1,*], Fakher Rahim[2] and Oleg Zolotarev[3]

[1]*Department of Development of Scientific and Innovation Activities, Russian New University, Moscow, Russia*
[2]*Health Research Institute, Thalassemia and Hemoglobinopathies Research Center, Ahvaz University of Medical Sciences, Ahvaz, Iran*
[3]*Department of Information Systems in Economics and Management, Russian New University, Moscow, Russia*

Abstract:

Aim:

The aims of the research were to study the citation history of popular articles in the field of biomarkers in personalized medicine, to study the use of terms in the sections of articles, and to consider the key terminology of the most-cited articles and its visualization.

Background:

The article describes approaches to the analysis of publication activity in the field of biomarkers and personalized medicine based on the data from the Web of Science.

Objective:

The aim of this study is a bibliometric and semantic analysis of the investigation field related to the application of biomarkers for the purposes of personalized medicine.

Methods:

The evaluation of a number of publications and its' citations was carried out. The key terms extracted from the most-cited articles were divided into thematic groups. The number of citations of the most popular articles since 2011 was estimated.

Results:

The citation histories of the top ten articles were considered. Analysis of key terms from different parts of the most-cited articles included statistics and thematic ranking. The comparison of key terms from the most-cited article and the citing articles allowed us to show that the key terminology of the cited article extends to the citing articles. We presented the key terms of the most-cited articles as a terminological map.

Conclusion:

The study of citation of the articles in the field of personalized medicine and biomarkers was based on a survey on the Web of Science. Based on the analysis of a number of citations the trends and citation histories were constructed. The statistical and thematic analysis of the use of keywords in different sections of articles was done. We have shown that the citing articles spread the key terms of the cited article to identify trends in knowledge development which could be presented as a terminological map.

Others:

We presented the results in the form of a terminological map of the latest developments in the field of biomarkers in personalized medicine based on proposed principles.

Keywords: Bibliometrics, Research trends, Biomarkers, Personalized medicine, Key terms, Citation.

Article History	Received: April 1, 2022	Revised: July 20, 2022	Accepted: August 23, 2022

1. INTRODUCTION

The current model of modern health care identifies disease with specific symptoms and its treatment, which is based on a method first proposed by Cnidean Medical School about 2,500 years ago [1]. According to this method, millions of people take drugs every day that do not help them. Recently, rising health care costs and relatively poor treatment of diseases have led to a rethinking of medical care and the advent of personalized medicine. Recent genetic discoveries and related developments in genome techniques have led to the commercialization of new diagnostic systems for the study of disease or the measurement of therapeutic outcomes in patients. Personal medicine uses the patient's genetic structure to test methods for diagnosing, treating or preventing a specific disease [2]. Personalized medicine (PM) stems from the fact that individuals have unique characteristics with very few differences in molecular, physiological, environmental, and behavioral fields. Therefore, interventions that are consistent with these unique features may be needed to treat the diseases of these individuals [3]. This belief has been partially confirmed using techniques such as proteomics, DNA sequencing, imaging protocols, and online health monitoring devices that could detect interpersonal changes in pathogenic processes [4, 5]. These techniques have revealed many interpersonal changes in terms of the effects, mechanisms, and factors that contribute to the pathogenic process. The question is what changes should influence the decision of an effective method of treating, monitoring, and preventing disease in a particular individual? In fact, the same heterogeneity in pathogenic processes led to the suggestion that the treatment of a person with a disease or diagnosis and disease prevention should be done personally and according to the unique biochemical, physiological, environmental, and behavioral characteristics of the same person [6, 7]. In other words, PM believes that no two diseases are similar. It should be noted that although the terms “personalized medicine”, “personal medicine”, and “precision” are used interchangeably, people believe that they are different [8].

The development of science is manifested in the growth of its quantitative information parameters. So, the quantitative characteristics of information flows are used as indirect data on the trends and rates of development of specific branches of science. New problem areas differ from an existing branch of science in the rate of growth in the number of publications, the frequency of the emergence of new terms, the names of authors, *etc*. [9]. Bibliometrics is a quantitative and qualitative classification of an issue that can be categorized by articles, authors, and journals. Bibliometrics is mostly used to analyze published articles on scientific topics and literary texts [10]. Hot topics may consist of words or phrases that are taken from the bibliographic data of the article (title, abstract, keywords, *etc*.). These data are considered to contain a detailed picture of the subject of the article and are used to quantify research trends or identify topics [11].

* Address correspondence to this author at the Department of Development of Scientific and Innovation Activities, Russian New University, Moscow, Russia; Tel: +79053729615; E-mail: aida_khatif@mail.ru

The present study is aimed to assess and analyze the global research of biomarkers in PM using bibliometrics and semantic methods.

2. MATERIALS AND METHODS

Data used in this study were retrieved from the Clarivate Analytics Web of Science, by the indexes SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI on 04 September 2021. The database was searched under the combined query: TITLE: (Biomarkers OR Biomarker OR Bio-marker OR Biological markers) + TITLE: (personalized medicament OR personalized medicinal) OR personalised medicine) OR personalised medication) OR personalize medications) OR personalize medications) OR personalizing medical) OR personalized medications) OR personalised medicines) OR personalize medication) OR personalize medical) OR personalized medication) OR personalised medical) OR personalizing medicines) OR personalized medical) OR personalised medical) OR personalized medical) OR personalized medicines) OR personalized medicine) OR personalistic medicine) OR personalised medicine) OR personalising medicine) OR personalize medicine) OR personalizing medicine) OR personalised medicine) OR personality medical) within the timespan for all years.

There were found 209 papers with a number of citations equal to 3095. The average citation per item is 14,81, h-index - 32. The impact factor of a journal was based on the *Journal Citation Report* 2019. The following citation indicators are used:

- The number of citations of a publication from the Web of Science Core Collection in a single year, for example, 2021, was referred to as the C_{2021} [12],
- The total number of citations since publication up to 2021 was referred to as the TC_{2021} [13, 14],
- The percentage of C_{2020} (or another period) in TC_{2021} was referred to as the *PC,*
- The average citation per year (*AC*),
- The relationship number of articles (N),
- The average citations per publication *(ACP)* were used to characterize the publications in PM and biomarkers research field:

$$ACP = TC_{year}/N \quad (1)$$

where *ACP* – the average citations per publication; TC_{year} – the total number of citations since year; *N* – the number of articles.

Keyword analysis included the following steps:

- The expert extraction of keywords from titles of articles, machine extraction of the semantic core from abstracts and texts of articles;

- The division of terms into thematic groups, calculation of the relative contribution of terms from different thematic groups to the general composition of the key terminology;

- Ranking the contribution of terms from different thematic groups to the formation of titles, abstracts, author's keywords, and the basis of the semantic core of texts.

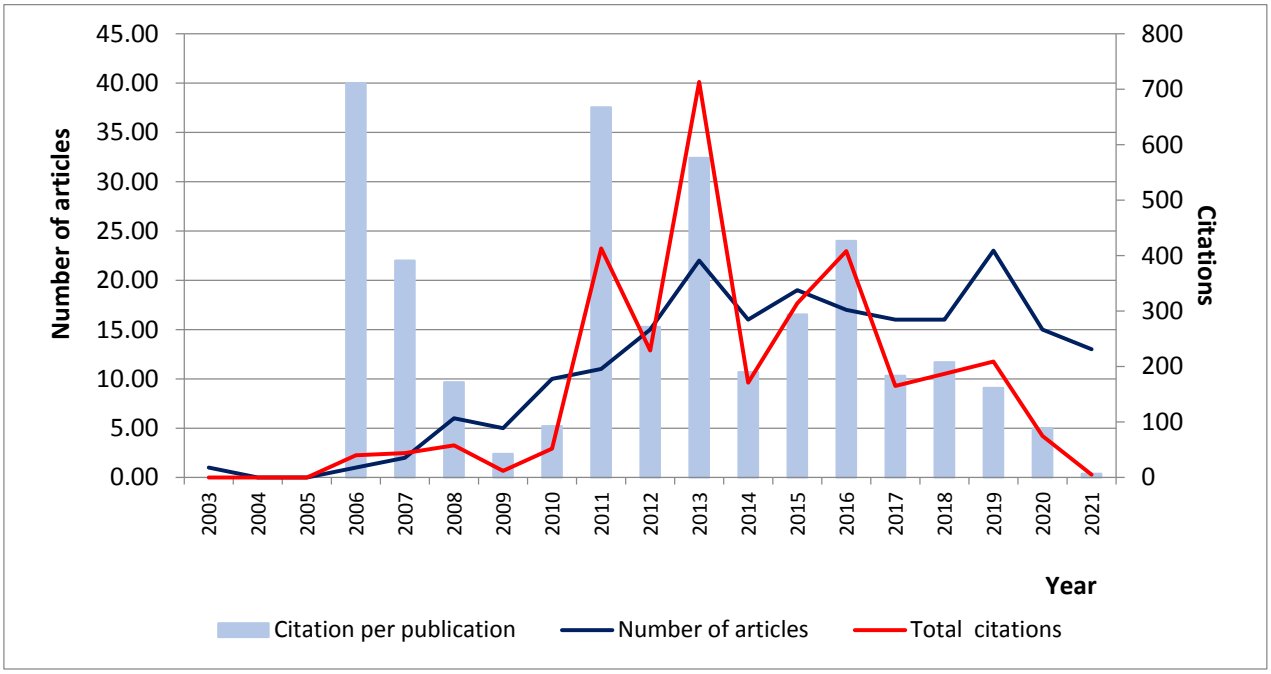


Fig. (1). Dynamics of average citations per publication, total citations of the articles published in the specified year and the total number of articles published in the specified year. All the articles are related to the biomarkers in personalized medicine. The number of citations was referred to September 2021.

The study of the influence of the cited article on the terminology of the citing articles included the selection of the main terms from the cited article and the identification of their presence in the titles and abstracts of all citing articles.

To build a terminological map of the subject area "biomarkers and personalized medicine" the main key terms from the most-cited articles were identified. https://github.com/ol-zolot/Biomarkers-in-Personalized-Medicine

3. RESULTS

3.1. Citation Histories of the Top Articles

The total number of cited articles is one of the main indicators used for comparing scientific performance in research subjects [15]. The total number of citations of an article is probably not sufficient to show the impact it had in the research field [16]. The number of citations in the most recent year (C_{year}) was further considered as an indicator [12]. Table **S1** contains the most frequently cited 26 articles in the field of PM and biomarkers research in Science Citation Index Expanded (SCIE). We used the following parameters to characterize these articles: TC_{2021} – the number of citations (Web of Science Core Collection) from the article publication date to September 2021; C_{2021} – the number of citations in 2021; *PC* – the percentage of C_{2020} in total C_{2021}; *AC* – the average citation per year. The articles were ranked by TC_{2021} and C_{2020}.

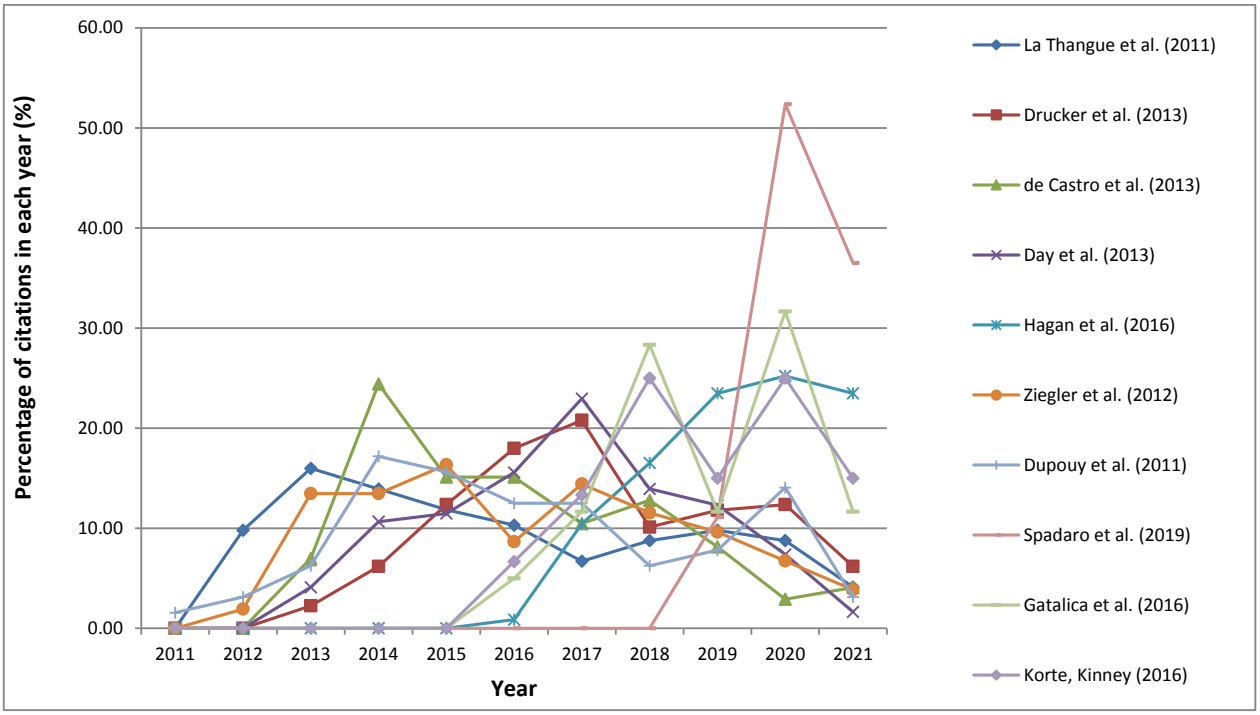
A relationship between the number of articles (*N*) and their citations per publication by years has been studied (formula 1). Fig. (**1**) shows the distribution of the 209 PM and biomarkers articles over the year, their citations, and *ACP*.

The patterns of citation life cycles of the top-cited articles could provide the characteristics for the top articles [17]. The citation histories of the top-ten articles are shown in Fig. (**2**). The articles with the highest total citations *(TC)* can be considered the most popular PM and biomarkers research articles. Citations of the most-cited articles fluctuate greatly (Fig. **2**).

3.2. Analysis of Key Terms by Titles, Annotations, Texts, and Authors' Keywords

For the analysis of key terminology, the 12 articles were selected based on the following criteria: 1) total citations, 2) citations over the last 5 years with at least 70% of the total citation. The 100 articles met these requirements. However, the absolute number of citations may be small, so we set the minimum citation threshold at 40. The 12 articles met these requirements. Table **S2** shows the selected articles and calculated parameters, such as TC_{2021} (section "Materials and Methods"), $C_{2016\text{-}2021}$ (the number of citations from 2016 to 2021), $PC_{2016\text{-}2021}$ (the percentage of $C_{2016\text{-}2021}$ in total TC_{2021}).

The groups of key terms were extracted from the 12 of the most-cited articles [18 - 29] from: 1) the title of the article, 2) the abstract (semantic core), 3) the full text of the article (semantic core), 4) author's keywords. Key terms from the titles were extracted by experts. The Advego software was used to extract the semantic core from abstracts and texts [30]. The results are presented in Table **S3**.



Figu. (2). The citation histories of the top 10 cited articles by the indexes SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI on 04 September 2021 in percent of the total citation.

When separating the semantic core from abstracts, the number of terms was limited to no more than 15. When isolating the semantic core from text, the limit was the frequency of occurrence in the text of more than 0.25%. Extracted key terms from the previously selected articles (Table **S2**) are shown in Table **S3**.

Since the key terms are quite heterogeneous, we divided them into the following thematic groups: 1) biomolecules, medical materials, and drugs, 2) the field of medicine and health care, 3) diseases; 4) medical technology, 5) medical and biological processes, 6) biological body components, 7) medical vocabulary, 8) general scientific vocabulary and 9) another vocabulary. Table **1** presents the examples of terms for each group (Table **S4** contains all meaningful key terms by subject groups). The total number of key terms by subject groups is presented in Table **1** (keywords from the titles, abstracts, full texts of the articles and author's keywords).

Table 1. The key terms by subject groups from the titles, abstracts, full texts of the articles and author's keywords (the examples).

S. No.	Subject Group	Key Terms from Titles	Key Terms from Abstracts	Author's Keywords	Key Terms from Texts
1	Biomolecules and medical materials, drugs	Biomarker(s), molecular biomarkers, predictive biomarkers, miRNAs, *etc*.	Biomarker, potential biomarkers, miRNA, marker, lncRNA, mRNA, RNA, gene	Biomarker(s), miRNAs; troponin, proadrenomedullin, procalcitonin, *etc*.	Biomarker, protein, gene, cytokine, DNA, allele, aminotransferase, antidepressant, d-dimer, *etc*.
2	The field of medicine and health care	Personaliz(s)ed medicine; predictive and personalised medicine, *etc*.	Personalize, personalized medicine, personalized, predictive, medicine, psychiatric, therapeutic	Personalized medicine, predictive preventive personalised medicine (PPPM), *etc*.	Therapy, chemotherapy, therapeutic, personalize, prognostic
3	Diseases	Acute respiratory distress syndrome, ovarian cancer, cancer, colorectal carcinoma, *etc*.	Cancer, ARDS, attention deficit hyperactivity disorder, colorectal carcinomas, CRC, ocular surface disease, *etc*.	Ovarian cancer, cancer, glaucoma, diabetic retinopathy, neurological disorders, depression, *etc*.	Cancer, periodontitis, lynch syndrome, CRC, carcinoma, colorectal cancers, CRCs, ADHD, ARDS, DED, *etc*.
4	Medical technology	Digital PCR	PCR, EEG, ERP	Digital PCR, quantitative electroencephalography, *etc*.	Digital PCR, EEG, WGCNA, PCR
5	Medical and biological processes	Microsatellite instability, gene co-expression	Gene expression, co-expression, MSI-H	Microsatellite instability, co-expression module	Mutation, expression, invasion, degradation, co-expression, necrosis, MSI-H

(Table 1) contd.....

S. No.	Subject Group	Key Terms from Titles	Key Terms from Abstracts	Author's Keywords	Key Terms from Texts
6	Biological body components	Microbiome, microbiota, tear fluid	Tear fluid, body fluid, blood, microbiome, microbe, microbiota, genome, *etc*.	Tear fluid, Single molecule, flora, microbiome	Microbe, microbial, microbiome, microbiota, blood, gut microbiome, tear fluid
7	Medical vocabulary	Clinical utility, ocular, systemic disease, clinical trait	Patient(s), clinical, care, diagnosis, disease, commensal, *etc*.	Proteome, metabolome, inflammatory, diagnosis, targeted prevention, *etc*.	Patient, clinical, cell, disease, health(y), diagnosis, diagnostic, inflammatory, acute, *etc*.
8	General scientific vocabulary	Prospects, discovery, development, analysis, identification, *etc*.	Potential, identify, field, research, study, advantage, algorithm, *etc*.	Prognostic assessment, regulatory overview, *etc*.	Study, analyse, molecule(ar), level, response, sample, increase, identify, *etc*.
9	Another vocabulary	Hope	Sleep, human, healthy		Help, eye, early, large, age, sleep, human, vigilance
	Total	53	116	64	352

The absolute and relative numbers of terms for each group were calculated. The ratio of key terms by subject groups from titles, abstracts, full texts of the articles and the author's keywords to the total number of key terms in each group are presented in Table **2**.

Among the specific key terms of all categories, terms that relate to biomolecules, medical materials and drugs predominate (Fig. **3**).

Consequently, most of the key terms of 12 of the most recent cited articles are related to the biomolecules, medical materials, and drugs.

Table 2. The ratio of key terms by subject groups from titles, abstracts, full texts of the articles and authors' keywords.

	Subject Group	Key Terms from Titles (%)	Key Terms from Abstracts (%)	Author's Keywords (%)	Key Terms from Texts of the Article (%)	Total by Group* (%)
1	Biomolecules, Medical Materials, Drugs	32,08	12,93	28,13	13,64	16,75
2	The Field of Medicine and Health Care	22,64	7,76	17,19	1,70	6,50
3	Diseases	15,09	7,76	20,31	4,83	8,03
4	Medical Technology	1,89	2,59	7,81	1,14	2,22
5	Medical and Biological Processes	3,77	2,59	3,13	2,56	2,74
6	Biological Body Components	3,77	6,90	6,25	1,99	3,59
7	Medical Vocabulary	7,55	23,28	10,94	30,68	24,96
8	General Scientific Vocabulary	11,32	33,62	6,25	41,19	33,16
9	Another Vocabulary	1,89	2,59	0,00	2,27	2,05
	Total	100,00	100,00	100,00	100,00	100,00

Note: * The total by group is calculated as the relation of key terms amount by each group to the total number of key terms.

For each category, the subject groups were ranked by a number of key terms from max (rank = 9) to min (rank = 1). The number of ranks of the groups is presented in Table **3**.

Table 3. The comparison ranks of subject groups of key terms.

S. No.	Subject Group	Key Terms from Title	Key Terms from Abstract	Author's Keywords	Key Terms from the Body of Article	The Number of the Ranks
		Rank	Rank	Rank	Rank	
1	Biomolecules and medical materials, drugs	9	7	9	7	32
2	The field of medicine and health care	8	6	7	2	23
3	Diseases	7	6	8	6	27
4	Medical technology	2	4	5	1	12
5	Medical and biological processes	4	4	2	5	15
6	Biological body components	4	4	4	3	15
7	Medical vocabulary	5	8	6	8	27
8	General scientific vocabulary	6	9	4	9	28
9	Another vocabulary	1	4	1	4	10

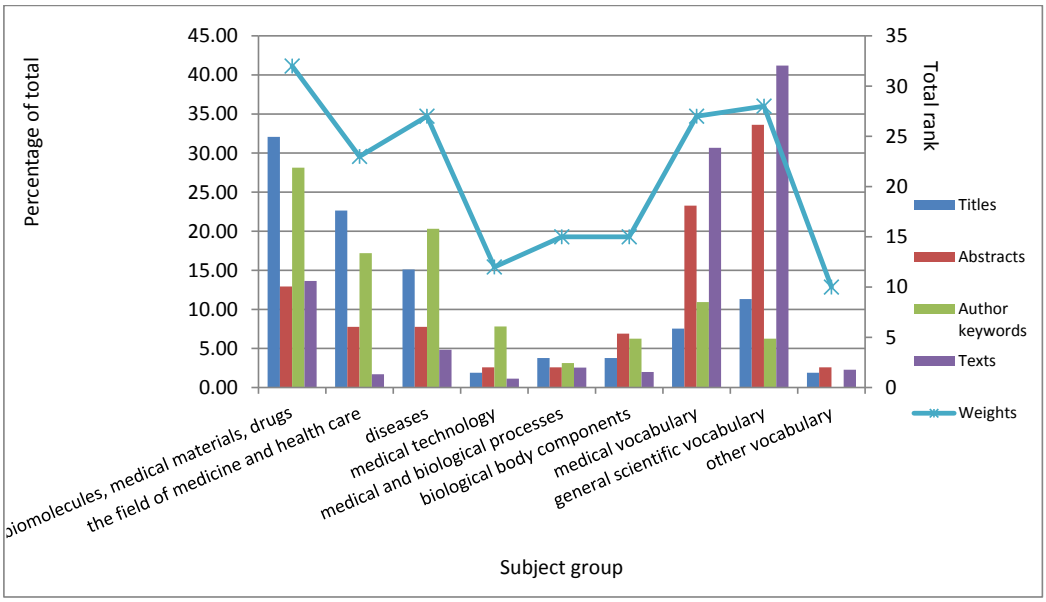


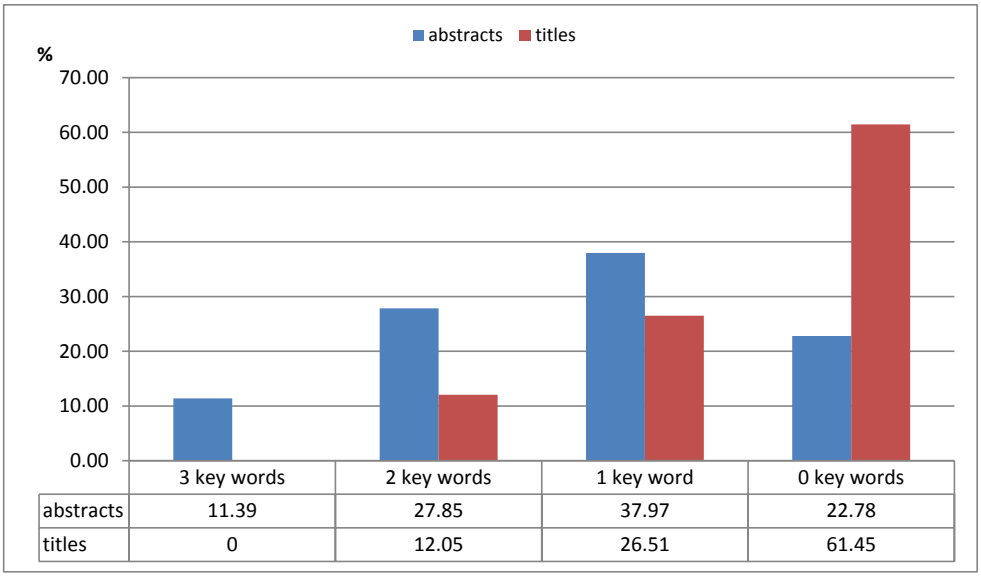
Fig. (3). The percentage of the key terms of the four categories (of the total number of key terms), presented by subject groups (total rank=weight presented in Table **3**.

3.3. The Relationship between the Key Terms of the Cited Article and the Citing Articles

We have attempted to assess whether the key terminology of the cited article is extended to the citing articles. Such an evaluation will be helpful in disseminating knowledge trends through the quoting terms.

We compared key terms from different parts of the article by Spadaro *et al*. (2019) - title, abstract, text of the article and author's keywords (Table **S3**). All the parts include the terms "biomarker" and "acute respiratory distress syndrome (ARDS)". The three parts (apart from the title) contain the term "inflammatory". The abstract and the text contain the terms "clinical", "diagnosis", and "gene". For further analysis, we chose the terms "biomarker", "acute respiratory distress syndrome (ARDS)", and "inflammatory". The terms "clinical", "diagnosis", "gene" are general medical and do not reflect the specifics of the study.

Next, we examined the occurrence of chosen terms in the titles and abstracts of all the 83 articles citing the article by Spadaro *et al*. (2019). We determined the number of articles that contain from one to three terms (or not contain). Then, we expressed the number of articles in percent of the total number of articles. The results are shown in Fig. (**4**).



	3 key words	2 key words	1 key word	0 key words
abstracts	11.39	27.85	37.97	22.78
titles	0	12.05	26.51	61.45

Fig. (4). The ratio of the number of articles citing an article by Spadaro *et al*. (2019), in the titles and abstracts of which there are from 0 to 3 key terms from the cited article.

Table 4. Key terms of the 12 of the most-cited articles in recent years.

References	Main Key Terms
Drucker *et al*. (2013) [18]	Biomarker, predictive medicine
Hagan *et al*. (2016) [19]	Biomarker, predictive, Tear fluid
Day *et al*. (2013) [20]	Biomarker, Digital PCR, PCR, personalized medicine
Spadaro *et al*. (2019) [21]	Biomarkers, Acute Respiratory Distress syndrome, inflammatory
Gatalica *et al*. (2016) [22]	Colorectal cancer, colorectal carcinoma, CRC, microsatellite instability (MSI-H), Lynch syndrome, PD-1
Korte, Kinney (2016) [23]	inflammatory, matrix metalloproteinase, periodontal disease
Goretti *et al*. (2014) [24]	biomarker, miRNA, myocardial infarction, personalized medicine
Olbrich *et al*. (2015) [25]	Attention Deficit Hyperactivity Disorder, Biomarker, EEG, electroencephalography, Event-related potentials, ERP, Major Depressive Disorder, MDD, Marker, Personalized Medicine
Willis *et al*. (2015) [26]	Immune biomarkers
Schuetz *et al*. (2015) [27]	Biomarker, d-dimer, emergency medicine, personalised
Li, Zhan (2019) [28]	Biomarker, Cancer, co-expression, invasion, lncRNA, mRNA, OCs, Ovarian cancer, personalized medicine, RNA, WGCNA
Rajpoot *et al*. (2018) [29]	Cancer, Gut, Microbe, Microbial, Microbiome; Microbiota, personalized medicine,

4. DISCUSSION

Citation analysis showed that the top-ten articles were published by 42 authors from the UK, Austria, Spain, France, Italy, Bosnia and Herzegovina, USA and Germany. We used the following parameters to characterize the most-cited articles: the number of citations from Web of Science Core Collection from the date of publication to September 2021 (TC_{2021}); the number of citations in 2021 (C_{2021}); the percentage of C_{2020} in total C_{2021} (*PC)*; the average citation per year. The article "Predictive biomarkers: a paradigm shift towards personalized cancer medicine" (2011) [31, 32] was ranked first by TC_{2021}. The article "Making Personalized Cancer Medicine a Reality: Challenges and Opportunities in the Development of Biomarkers and Companion Diagnostics" (2012) [33] was ranked first by C_{2020}.

The top-ten most-cited articles were published in high journal impact factor with IF_{2019} from 2.375 ("Familial Cancer") to 53.28 ("Nature Reviews Clinical Oncology").

The distribution of the 209 articles related to PM and biomarkers over the year and their *ACP* were considered. It is noticeable that the year 2019 had the most annual articles with 23. In 2006, 2011 and 2013 these articles had the highest *ACP* (more than 30), (Fig. **1**).

The study of citation history showed that the article by Spadaro *et al*. (2019) [21] had extremely high citations in comparison with other highly cited articles already in the first year after publication (52,38%) (Fig. **2**). The article by Hagan *et al*. (2016) [19] has quite consistently cited in recent years.

The calculation of keywords' ratio by subject groups from the titles, abstracts, full texts of the articles and author's keywords was completed. The analysis showed that the total distribution of key terms by all groups is as follows in comparison rank:

General scientific vocabulary > medical vocabulary > biomolecules, medical materials, drugs > diseases > the field of medicine and health care > biological body components > medical and biological processes > medical technology.

The distribution of key terms by groups in titles is as follows:

Biomolecules, medical materials, drugs > the field of medicine and health care > diseases > general scientific vocabulary > medical vocabulary > biological body components = medical and biological processes > medical technology.

The distribution of key terms by groups in abstracts is as follows:

General scientific vocabulary > medical vocabulary > biomolecules, medical materials, drugs > the field of medicine and health care = diseases > biological body components > another vocabulary = medical and biological processes = medical technology.

The distribution of key terms by groups in the author's keywords is as follows:

Biomolecules, medical materials, drugs > diseases > the field of medicine and health care > medical vocabulary > medical technology > general scientific vocabulary = biological body components > medical and biological processes.

The distribution of key terms by groups in full texts of articles is as follows:

General scientific vocabulary > medical vocabulary > biomolecules, medical materials, drugs > diseases > medical and biological processes > another vocabulary > biological body components > the field of medicine and health care > medical technology.

Calculation with the assignment of ranks to categories showed that among the keywords of all sources terms that relate to biomolecules, medical materials and drugs predominate.

We can assume that the terms listed in Table **4** have been especially widespread in recent years. We obtained them from 12 of the most-cited articles in recent years using the criterion of occurrence in at least two sections of articles. In addition, these terms should not refer to general medical and general scientific terms.

We presented the results in the form of a terminological map of the latest developments in the field of biomarkers in PM (Fig. **5**).

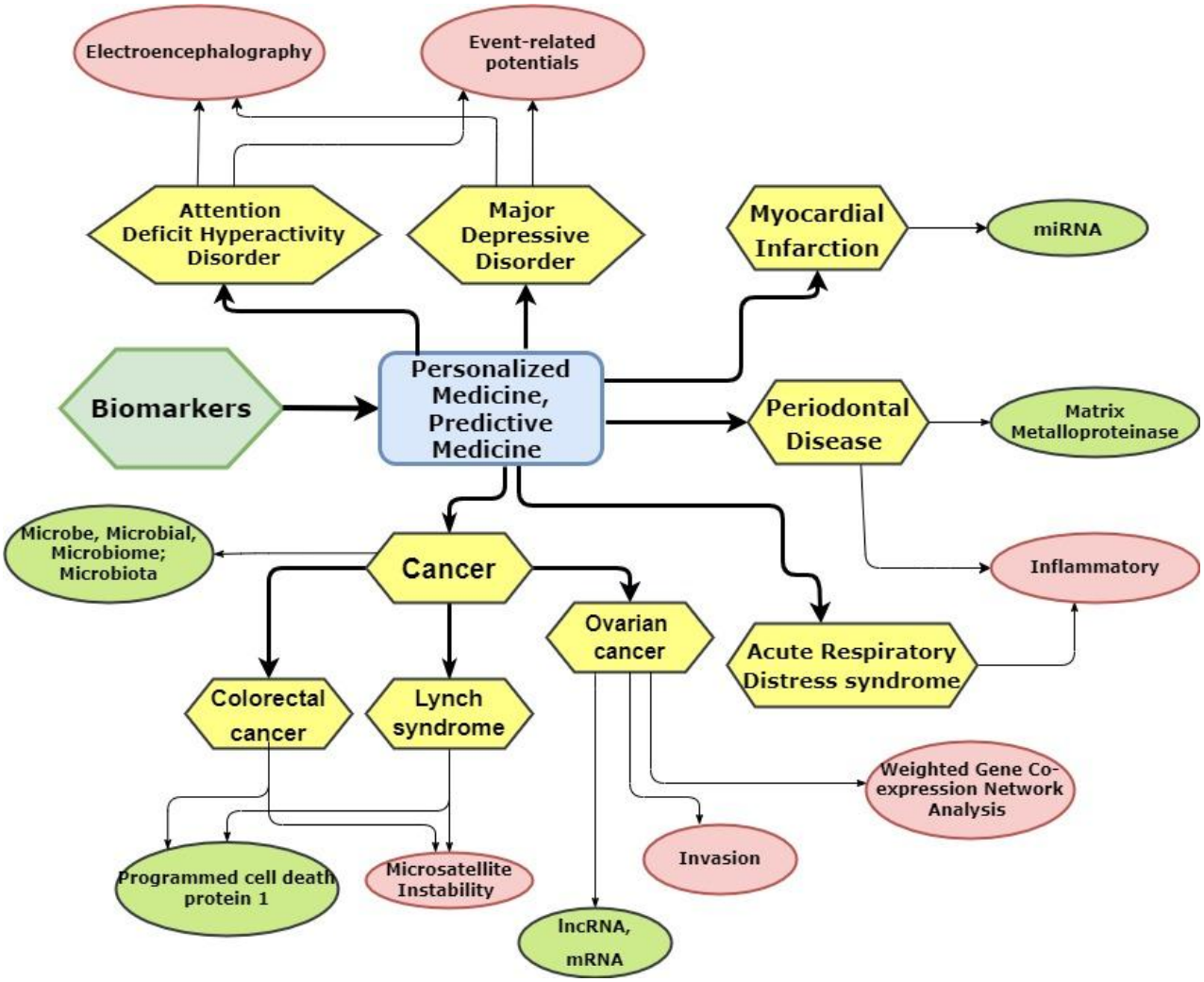


Fig. (5). Terminological map of the latest developments in the field of biomarkers in personalized medicine.

When executing the terminological map, we applied the following rules. The first level of detailing is the object of study (discussion). In this case, it is the type of disease. PM includes the treatment and diagnosis of diseases: cancer (colorectal cancer, colorectal carcinoma, Ovarian cancer, *etc*.), mood disorders (Major Depressive Disorder), neurodevelopmental disorders (Attention Deficit Hyperactivity Disorder), periodontal disease and myocardial infarction. The next level of detailing is the properties of the object. In our case, these are symptoms/characteristics/signs of diseases (inflammatory, invasion, microsatellite instability (MSI-H)); biomolecules involved in the process (matrix metalloproteinase, PD-1, d-dimer, miRNA, lncRNA, mRNA, RNA). The third level of detailing is the methods of influencing the object (diagnosis, treatment). In our case, these are Digital PCR, PCR, electroencephalography (EEG), Event-related potentials (ERP), and weighted gene co-expression network analysis (WGCNA).

CONCLUSION

This paper presents a study of the articles' citations related to personalized medicine and biomarkers based on the survey of the Web of Science text resource corpse, as well as an analysis of the intensity of using key terms in the most-cited articles. The most-cited articles in the specified field of medicine were selected. Based on the analysis of the citation depending on the time of publications, the trends and citation history were constructed.

Dynamics of average citations per publication and the number of articles related to the biomarkers in PM show that from 2003 to 2013 the annual number of articles in SCI-EXPANDED increased (Fig. **1**). The growth of publication activity was observed in 2013. These articles are actively cited up to 2021. A decline in activity was observed in 2009 and 2014, which may be explained by the problems in the economy of the developed countries.

The article by Spadaro *et al*. (2019) had extremely high citations already in the first year after publication in SCIE (52.38%). According to Semantic Scholar, as of 09/06/2021, the article received 83 citations [31].

The analysis of the use of key terms in article titles,

abstracts, texts, and the keywords of the authors was made. The groups of key terms for the most-cited articles from the Web of Science were selected as the semantic core of the text.

Key terms were divided into thematic groups. For the most-cited articles from the Web of Science the 9 ranked groups of key terms were distinguished. The occurrence of key terms in the titles, abstracts, and texts (and author keywords) was calculated, and the percentage ratio of key terms was also calculated by thematic groups.

The statistical and thematic analysis of key terms by titles, abstracts, the bodies of articles, and author's keywords allowed us to divide all terms into 9 categories: 1) biomolecules, medical materials, drugs; 2) the field of medicine and health care; 3) diseases; 4) medical technology; 5) medical and biological processes; 6) biological body components; 7) medical vocabulary; 8) general scientific vocabulary; 9) another vocabulary.

An analysis of the distribution of key terms by categories in different sections of the 12 of the most-cited articles showed that the authors try to indicate specific names of biomolecules, medical materials and drugs in the titles of articles and author's keywords. Perhaps this fact contributes to the more frequent appearance of these articles in search queries, which, in turn, can lead to increased citation of them. We used a similar approach in RFBR and NSFC project Nº21-57-53018 during the study of articles related to breast cancer.

We assumed that the citing articles propagate the keywords of the cited article. We have considered the use of key terms of the cited article (Spadaro *et al*., 2019 [21]) in different sections of the citing articles. It turned out that these terms appear in titles (from 1 to 3) in 38.55% of cases, which is not surprising due to the brevity of the titles. There are even complete coincidences of three terms (11.39%) in abstracts. In total, these terms are found in 77.22% of the citing articles. This suggests our hypothesis that the citing articles are actively disseminating the key terms of the cited article. This result permits to identify the trends in knowledge development.

Key terms were selected from the most-cited articles and a terminological map of the development of the subject area of the application of biomarkers in PM was created. The first level of detailing is the object of study (disease). The next level of detailing is the symptoms/characteristics/signs of diseases and biomolecules involved in the process. The third level of detailing is the methods of diagnosis and treatment. We used a similar approach to create a terminological map representing the views of opponents of vaccination against coronavirus in the RFBR project Nº 20-04-60185 and other works [34 - 36].

Study limitations are related to a specific search query in the Web of Science, the period under study and the number of articles analyzed. In the future, we plan to develop the principles of extracting new terms and forming their trends to discover the hot spots of research using statistical methods and machine learning.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

The reported study was funded by RFBR and NSFC, project number 21-57-53018 and by RFBR, project number 20-04-60185.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIALS

Table **S1**: Keywords from the most-cited in recent years' articles in Science Citation Index Expanded; Table **S2**: Parameters of the selected articles for keyword analysis; Table **S3**: Keywords from the most-cited in recent years articles in Science Citation Index Expanded; Table **S4**: The keywords by subject groups from the title, annotations, author keywords and the full text of the article (in brackets is the number of articles in the semantic core of which this term occurs). The supported data:
https://github.com/ol-zolot/Biomarkers-in-Personalized-Medicine

REFERENCES

[1] Sugeir S, Naylor S. Critical Care and Personalized or Precision Medicine: Who needs whom? J Crit Care 2018; 43: 401-5. [http://dx.doi.org/10.1016/j.jcrc.2017.11.026] [PMID: 29174462]

[2] Guttmacher AE, McGuire AL, Ponder B, Stefánsson K. Personalized genomic information: preparing for the future of genetic medicine. Nat Rev Genet 2010; 11(2): 161-5. [http://dx.doi.org/10.1038/nrg2735] [PMID: 20065954]

[3] Allison M. Is personalized medicine finally arriving? Nat Biotechnol 2008; 26(5): 509-17. [http://dx.doi.org/10.1038/nbt0508-509] [PMID: 18464779]

[4] Dias D, Paulo Silva Cunha J. Wearable Health Devices—Vital Sign Monitoring, Systems and Technologies. Sensors (Basel) 2018; 18(8): 2414. [http://dx.doi.org/10.3390/s18082414] [PMID: 30044415]

[5] Miller RR, Montoya V, Gardy JL, Patrick DM, Tang P. Metagenomics for pathogen detection in public health. Genome Med 2013; 5(9): 81. [http://dx.doi.org/10.1186/gm485] [PMID: 24050114]

[6] Becker F, van El CG, Ibarreta D, *et al.* Genetic testing and common disorders in a public health framework: how to assess relevance and possibilities. Background Document to the ESHG recommendations on genetic testing and common disorders. European journal of human genetics 2011; 19 Suppl 1(Suppl 1): S6-44. [http://dx.doi.org/10.1038/ejhg.2010.249]

[7] van El CG, Cornel MC. Genetic testing and common disorders in a public health framework. Eur J Hum Genet 2011; 19(4): 377-81. [http://dx.doi.org/10.1038/ejhg.2010.176] [PMID: 21267009]

[8] Ong S, Ling J, Ballantyne A, Lysaght T, Xafis V. Perceptions of 'Precision' and 'Personalised' Medicine in Singapore and Associated Ethical Issues. Asian Bioeth Rev 2021; 13(2): 179-94. [http://dx.doi.org/10.1007/s41649-021-00165-3] [PMID: 33959200]

[9] Hicks D, Wouters P, Waltman L, de Rijcke S, Rafols I. Bibliometrics: The Leiden Manifesto for research metrics. Nature 2015; 520(7548):

429-31.
[http://dx.doi.org/10.1038/520429a] [PMID: 25903611]
[10] Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? Scientometrics 2015; 105(3): 1809-31.
[http://dx.doi.org/10.1007/s11192-015-1645-z] [PMID: 26594073]
[11] Zhang J, Yu Q, Zheng F, Long C, Lu Z, Duan Z. Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. J Assoc Inf Sci Technol 2016; 67(4): 967-72.
[http://dx.doi.org/10.1002/asi.23437]
[12] Ho YS. Top-cited articles in chemical engineering in Science Citation Index Expanded: A bibliometric analysis. Chin J Chem Eng 2012; 20(3): 478-88.
[http://dx.doi.org/10.1016/S1004-9541(11)60209-7]
[13] Wang MH, Fu HZ, Ho YS. Comparison of universities' scientific performance using bibliometric indicators. Malays J Libr Inf Sci 2011; 16(2): 1-19.
[14] Chuang KY, Wang MH, Ho YS. High-impact papers presented in the subject category of water resources in the essential science indicators database of the institute for scientific information. Scientometrics 2011; 87(3): 551-62.
[http://dx.doi.org/10.1007/s11192-011-0365-2]
[15] Terajima K, Åneman A. Citation classics in anaesthesia and pain journals: a literature review in the era of the internet. Acta Anaesthesiol Scand 2003; 47(6): 655-63.
[http://dx.doi.org/10.1034/j.1399-6576.2003.00137.x] [PMID: 12803581]
[16] Ho YS. Classic articles on social work field in Social Science Citation Index: a bibliometric analysis. Scientometrics 2014; 98(1): 137-55. b
[http://dx.doi.org/10.1007/s11192-013-1014-8]
[17] Aksnes DW. Characteristics of highly cited papers. Res Eval 2003; 12(3): 159-70.
[http://dx.doi.org/10.3152/147154403781776645]
[18] Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. EPMA J 2013; 4(1): 7.
[http://dx.doi.org/10.1186/1878-5085-4-7] [PMID: 23442211]
[19] Hagan S, Martin E, Enríquez-de-Salamanca A. Tear fluid biomarkers in ocular and systemic disease: potential use for predictive, preventive and personalised medicine. EPMA J 2016; 7(1): 15.
[http://dx.doi.org/10.1186/s13167-016-0065-3] [PMID: 27413414]
[20] Day E, Dear PH, McCaughan F. Digital PCR strategies in the development and analysis of molecular biomarkers for personalized medicine. Methods 2013; 59(1): 101-7.
[http://dx.doi.org/10.1016/j.ymeth.2012.08.001] [PMID: 22926236]
[21] Spadaro S, Park M, Turrini C, *et al.* Biomarkers for Acute Respiratory Distress syndrome and prospects for personalised medicine. J Inflamm (Lond) 2019; 16(1): 1.
[http://dx.doi.org/10.1186/s12950-018-0202-y] [PMID: 30675131]
[22] Gatalica Z, Vranic S, Xiu J, Swensen J, Reddy S. High microsatellite instability (MSI-H) colorectal carcinoma: a brief review of predictive biomarkers in the era of personalized medicine. Fam Cancer 2016; 15(3): 405-12.
[http://dx.doi.org/10.1007/s10689-016-9884-6] [PMID: 26875156]
[23] Korte DL, Kinney J. Personalized medicine: an update of salivary biomarkers for periodontal diseases. Periodontol 2000 2016; 70(1): 26-37.
[http://dx.doi.org/10.1111/prd.12103] [PMID: 26662480]
[24] Goretti E, Wagner DR, Devaux Y. miRNAs as biomarkers of myocardial infarction: a step forward towards personalized medicine? Trends Mol Med 2014; 20(12): 716-25.
[http://dx.doi.org/10.1016/j.molmed.2014.10.006] [PMID: 25457620]
[25] Olbrich S, van Dinteren R, Arns M. Personalized Medicine: Review and Perspectives of Promising Baseline EEG Biomarkers in Major Depressive Disorder and Attention Deficit Hyperactivity Disorder. Neuropsychobiology 2015; 72(3-4): 229-40.
[http://dx.doi.org/10.1159/000437435] [PMID: 26901357]
[26] Willis JCD, Lord GM. Immune biomarkers: the promises and pitfalls of personalized medicine. Nat Rev Immunol 2015; 15(5): 323-9.
[http://dx.doi.org/10.1038/nri3820] [PMID: 25814400]
[27] Schuetz P, Aujesky D, Müller C, Müller B. Biomarker-guided personalised emergency medicine for all – hope for another hype? Swiss Med Wkly 2015; 145w14079
[http://dx.doi.org/10.4414/smw.2015.14079] [PMID: 25695147]
[28] Li N, Zhan X. Identification of clinical trait–related lncRNA and mRNA biomarkers with weighted gene co-expression network analysis as useful tool for personalized medicine in ovarian cancer. EPMA J 2019; 10(3): 273-90.
[http://dx.doi.org/10.1007/s13167-019-00175-0] [PMID: 31462944]
[29] Rajpoot M, Sharma AK, Sharma A, Gupta GK. Understanding the microbiome: Emerging biomarkers for exploiting the microbiota for personalized medicine against cancer. Semin Cancer Biol 2018; 52(Pt 1): 1-8.
[http://dx.doi.org/10.1016/j.semcancer.2018.02.003] [PMID: 29425888]
[30] Available from: https://advego.com/text/seo/
[31] Available from: https://www.semanticscholar.org/paper/Biomarkers-for-Acute-Respiratory-Distress-syndrome-Spadaro-Park/b2b2602266cd735aae693b599f4c7591b3a8be27?sort=relevance&page=9
[32] La Thangue NB, Kerr DJ. Predictive biomarkers: a paradigm shift towards personalized cancer medicine. Nat Rev Clin Oncol 2011; 8(10): 587-96.
[http://dx.doi.org/10.1038/nrclinonc.2011.121] [PMID: 21862978]
[33] Parkinson DR, Johnson BE, Sledge GW. Making personalized cancer medicine a reality: challenges and opportunities in the development of biomarkers and companion diagnostics. Clin Cancer Res 2012; 18(3): 619-24.
[http://dx.doi.org/10.1158/1078-0432.CCR-11-2017] [PMID: 22298894]
[34] Fakher Rahim, Aida Khakimova, Ammar Ebrahimi, Oleg Zolotarev, Fatemeh Rafiei Nasab. Global scientific research on SARS-CoV-2 vaccines: a bibliometric analysis. Cell J 2021; 23(5): 523-31.
[http://dx.doi.org/10.22074/cellj.2021.7794]
[35] Khakimova AK, Zolotarev OV, Berberova MA. Coronavirus infection study: bibliometric analysis of publications on COVID-19 using PubMed and Dimensions databases. Sci Vis 2020; 12(5): 112-29.
[http://dx.doi.org/10.26583/sv.12.5.10]
[36] Khakimova AKh, Zolotarev OV, Berberova MA. Visualization of bibliometric networks of scientific publications on the study of the human factor in the operation of nuclear power plants based on the bibliographic database Dimensions. Scientific Visualization 2020; 12(2): 127-38.
[http://dx.doi.org/10.26583/sv.12.2.10]