

## СОЦІАЛЬНА МЕДИЦИНА

© Morokhovets H., Lysanets Yu., Ostrovska L., Purdenko T.  
UDC 378.147.227

DOI <https://doi.org/10.31718/mep.2019.23.3-4.11>

### MASTERING THE TOOLS OF CITATION DATABASES IN HIGHER MEDICAL EDUCATION

Morokhovets H., Lysanets Yu., Ostrovska L., Purdenko T.

Ukrainian Medical Stomatological Academy, Poltava, Ukraine

*У сучасному науковому середовищі якість публікацій, публікаційну активність вченого, навчального закладу чи наукової установи оцінюють з допомогою комплексу наукометричних показників. Серед них виділяють наступні: сумарну цитованість, загальну кількість посилань, індекс Хірша, імпакт-фактор, g-індекс, i-10-індекс. Метою статті є висвітлення прийомів роботи зі світовими науково метричними базами наукової літератури, пояснення сутності та алгоритмів розрахунку основних наукометричних показників. У процесі підготовки статті використано бібліо-семантичний метод дослідження для аналізу наукової літератури. У статті висвітлено принципи роботи з міжнародними базами даних Scopus, Web of Science, Google Scholar, бібліографічним менеджером Mendeley, системою ORCID, сутність наукометричних показників, що розраховуються даними ресурсами та сфери їх застосування. Проаналізовано позитивні сторони використання наукометричних показників для визначення впливовості конкретного автора чи установи, а також ряд суб'єктивних недоліків широкого їх впровадження у закладах вищої освіти. Сучасні інструменти роботи з науковою інформацією відіграють провідну роль у науково-дослідній діяльності вчених, закладів вищої освіти, наукових установ. Рівень якості публікацій визначається з допомогою низки наукометричних показників, що розраховуються авторитетними базами наукової літератури. Серед них визнаними є бази Scopus та Web of Science. h-індекси, розраховані на основі публікацій у цих базах, слугують якісними показниками оцінювання наукової діяльності. Scopus та Web of Science, крім власних вбудованих інструментів пошуку, візуалізації, аналізу та відслідковування даних, інтегровані з реєстрами вчених, бібліографічними менеджерами, плагінами, що дозволяють автоматизувати процеси створення цитат, списків літератури, підготовки публікацій за вимогами конкретного видання, роботи з повнотекстовими версіями статей у різних форматах.*

**Ключові слова:** міжнародні наукометричні бази даних, методологія пошуку, Scopus, Web of Science, Google Scholar.

*In the contemporary academic setting, the quality of publications and the publication activity of a scientist, educational or research institution are evaluated using a set of scientific metric indicators. There are the following parameters: total citations, total number of links, the Hirsch index, impact factor, g-index, and i-10 index. The aim of the paper is to cover the methods of working with the world scientific and metric databases of literature, as well as to explain the essence and algorithms for calculating the major scientific and metric indicators. We used the bibliosemantic research method in order to analyze the relevant scientific literature. The article describes the principles of mastering the tools of international databases Scopus, Web of Science, Google Scholar, Mendeley bibliographic manager, ORCID system. The paper focuses on the essence of scientific and metric indicators, calculated by these resources, as well as their scope and possibilities. The authors analyze the positive aspects of using the scientific metrics to determine the influence of a particular author or institution, as well as a number of subjective disadvantages of their widespread adoption in higher education. The modern tools for working with scientific information play a key role in the research activities of scientists, institutions of higher education, research institutions. The quality of publications is determined by a number of scientific metrics, calculated by authoritative scientific literature databases, such as Scopus and Web of Science. The h-indices are calculated in terms of publications in these bases, and serve as qualitative indicators for scientific activity evaluation. Scopus and Web of Science, in addition to their own built-in tools for searching, visualizing, analyzing and tracking data, integrated with the registers of scientists, bibliographic managers, plugins that allow to automate the processes of citation, lists of literature, preparation of publications according to the requirements of a specific edition, work with full-text versions of articles in different formats.*

**Key words:** international scientific metric databases, search methodology, Web of Science, Google Scholar.

Bibliometry as a new direction in the study of science is based on the analysis of bibliographic data of publications and aims at strategic forecasting of the develop-

ment of science (monitoring), rather than obtaining the statistical indicators. The objects of study in the bibliometric research are publications sorted by authors, by

To cite this english version: Morokhovets H., Lysanets Yu., Ostrovska L., Purdenko T. Mastering the tools of citation databases in higher medical education // The Medical and ecological problems. - 2019. - Vol 23, № 3-4. - P. 46-52.

publication sources, by subject sections, by countries and so forth. In the modern scientific world, a number of bibliometric terms are used, including: total number of references, average citation, total citations, impact factor, the Hirsch index (h-index), ranking place, etc. It is important to understand the scope of these terms. Hence, the impact factor is calculated for journals, h-index – for a scientist or an organization, whereas the number of references – for research articles. Let us consider these terms in more detail.

The Hirsch index (h-index) is a scientific metric indicator, based on the calculation of the number of researcher publications and the number of citations of these publications. The h-index is a quantitative characteristic of a scientist's productivity for the whole period of their activity. The indicator was suggested by the American physicist Jorge Hirsch, Professor at the University of California, San Diego, in 2005. The algorithm for calculating the h-index is as follows: if a scientist has published  $N$  articles,  $n$  of which have been cited at least  $n$  times, then the scientist's h-index will be  $n$ . For example, a scholar's array of articles comprises 80 papers, 7 of which have been cited at least 7 times each. The h-index of this scientist is 7.

The h-index is automatically calculated by resources, publicly available on the Internet, including Google Scholar, Scopus, Web of Science databases.

The g-index is a bibliometric indicator, calculated on the basis of the scientist's citation distribution. The algorithm for calculating the g-index is as follows: for a set of articles, sorted in the descending order of citation, the g-index is the largest number, such that  $g$  of the most cited articles has received (in total) at least  $g^2$  citations.

The i-10 index is a bibliometric indicator, used by Google Scholar to dynamically evaluate a scientist's or organization's publishing activity and is calculated on the basis of bibliometric indicators. The i-10 index calculation algorithm is as follows: the index shows the number of publications that have at least 10 bibliographic references.

The h-index and i-10 index are calculated by scientific and metric databases automatically as search engines (crawlers) track new links to relevant articles on the Internet, posted in open access. The g-index, unlike the previous two, is not calculated automatically, it must be determined manually.

The above mentioned scientometric indexes are used to determine the influence of a particular author, their publishing activity. These indices are qualitative parameters of scientific activity of a scientist. In addition to their undeniable benefit, they have a number of subjective shortcomings: the author of a paper that has been quoted for 100 times and the author who wrote 10 articles, each cited once, have the h-index equaling to one; scientific and metric indices do not reveal the promising directions of young scientists, whose articles are still little cited; the h-index, calculated for organizations and institutions does not take into account the average number of scientists in institutions, etc.

Impact factor (IF) is a numerical scientometric indicator of the impact of a scientific journal, calculated as the ratio of the number of references received by a journal in the current year to articles published in this journal in the previous two years, to the number of articles published in this journal in these two previous years. Thus, the impact factor is the measure that determines the frequency at which the average article is cited. As a criterion for

evaluating a journal, the impact factor indicates that a journal that publishes a significant number of articles that other scholars actively refer to is sufficiently influential.

The impact factor of scientific journals is calculated by the Institute for Scientific Information (<https://isindexing.com/isi/>) and published by *Journal Citation Report* (<https://clarivate.com/products/journal-citation-reports/>). The Institute for Scientific Information does not embrace reports, editorials and other types of publications for their calculation, and for new journals IF is calculated for a two-year period.

Bibliometric assessment of science is also carried out to provide society with a holistic view of the research environment at the state level. In Ukraine, there is a single register of bibliometric profiles of scientists and organizations – "Bibliometrics of Ukrainian Science" (<http://www.nbuv.gov.ua/bpnu/>). The system provides comprehensive statistical information on the sectoral, departmental and regional structure of science in Ukraine, and it is a national segment of the international Ranking of Google Scholar Profiles. The system processes bibliographic profiles of scientists in Google Scholar and Scopus.

The Scopus database contains abstracts and information about peer-reviewed scientific literature. In addition, the database has built-in tools for finding, analyzing, visualizing and tracking data. Scopus presents a wide array of scientific information in compliance with quality control standards. Scopus data quality is supported by the leading research and educational institutions around the world and is therefore used by renowned rating agencies to compile world university rankings. The database includes publications such as open access, books, conference materials, patents, specialized publications and the like. The database displays the full volume of publication metadata, digital indicators and article IDs, author and organization profiles. With Scopus tools, one can compare journals by bibliometric indicators (CiteScore, SNIP, SJR). The organization profile presented in Scopus has detailed information to evaluate its scientific activities.

The Scopus peer-reviewed database is updated twice a year and presented as a spreadsheet with a set of filters that allow to select journals for publication by various indicators, including impact factor, country of publication, area of science, specialty, etc.

Scopus author profile is a rather informative parameter that shows the total number of publications and citations, the h-index, the citation schedule of papers, co-authors, links to the affiliated organization profile and so on.

Scopus built-in features allow to find the most cited works of the author of the organization, to check the impact factors of journals in real time, to find partner organizations, authors working on similar scientific topics. The advanced Scopus features and products, integrated with it, are aimed at automating routine processes when creating publications and submitting them to editorial boards. These include the Mendeley bibliographic manager and the ORCID international registry of scientists. Further, let us focus on these tools in more detail.

The recommended list of references is approved by the Order of the Ministry of Education and Science of Ukraine No. 40 "On Approval of Requirements for Dissertation" as of 12.01.2017. The list contains a recommended set of styles for designing the references to publications, including MLA (Modern Language Association)

style, APA (American Psychological Association) style, Chicago / Turabian style, Harvard style, ACS (American Chemical Society) style, AIP (American Institute of Physics) style, IEEE (Institute of Electrical and Electronics Engineers) style, Vancouver style, OSCOLA, APS (American Physics Society) style, Springer MathPhys Style. The choice of reference style depends on many conditions, including the area of expertise within which the research is performed. Preparing an article for the requirements of a particular journal, including citation styles, is a very laborious process. For this purpose, special software (bibliographic manager) is used to simplify the work with scientific information and to automate the processes of citation and the list of references. Bibliographic manager Mendeley is an on- and off-line application, designed for individual and shared use. Mendeley is compatible with PDF versions of articles.

Scopus is integrated with the ORCID international registry of scientists. Open Researcher and Contributor ID (ORCID) is a registry of author IDs, designed to present transparently the research activities and freely ac-

cess its results. The main advantage of ORCID, in our opinion, is that it can be linked to the Scopus Author ID and the Web of Science Researcher ID. Thus, the identifier binds all the unique author IDs and all possible spellings, allowing to export citations from other databases, to present the results of research. In addition, ORCID brings together academics from around the world, including research organizations, publishing houses, grant organizations.

Google Scholar is a freely accessible search engine that indexes the full texts of scientific publications in all fields of knowledge and formats. The database includes most on-line publications from European and US publishing houses. To start working with the database, one must follow the link <https://scholar.google.com>.

The database allows you to search for publications by keywords, authors' names (Fig. 1). By clicking the icon ★, the selected publication is added to your own library. Other additional possibilities of Google Scholar are presented at Figures 2-4.

The screenshot shows the Google Scholar interface. At the top, the search bar contains the text "біоінформатика" and a magnifying glass icon. Below the search bar, it indicates "Статті" (Articles) and "Приблизна кількість результатів: 17 500 (0,10 сек.)". On the left side, there are filters for "Будь-коли" (Any time) with a date range from 2019 to 2015, and checkboxes for "включаючи патенти" (include patents) and "включаючи цитування" (include citations). The main search results are listed as follows:

- [HTML] Bloconductor: open software development for computational biology and bioinformatics** [HTML] [biomedcentral.com](http://biomedcentral.com)  
RC Gentleman, VJ Carey... - Genome ... , 2004 - genomebiology.biomedcentral.com  
The Bloconductor project is an initiative for the collaborative creation of extensible software for computational biology and bioinformatics. The goals of the project include: fostering collaborative development and widespread use of innovative software, reducing barriers to ...  
☆ 99 Цитовано в 10766 джерелах Пов'язані статті Кількість версій: 29 Web of Science: 7396 00
- [ЦИТАТА] Bioinformatics: sequence and genome analysis**  
DW Mount, DW Mount - 2001 - Cold spring harbor laboratory press ...  
☆ 99 Цитовано в 1807 джерелах Пов'язані статті Кількість версій: 10
- Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources** [PDF] [psu.edu](http://psu.edu)  
DW Huang, BT Sherman, RA Lempicki - Nature protocols, 2009 - nature.com  
DAVID bioinformatics resources consists of an integrated biological knowledgebase and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists. This protocol explains how to use DAVID, a high-throughput and integrated data ...  
☆ 99 Цитовано в 19125 джерелах Пов'язані статті Кількість версій: 14 Web of Science: 14445
- [КНИГА] Bioinformatics: the machine learning approach**  
P Baldi, S Brunak, F Bach - 2001 - books.google.com  
A guide to machine learning approaches and their application to the analysis of biological data. An unprecedented wealth of data is being generated by genome sequencing projects and other experimental efforts to determine the structure and function of biological ...  
☆ 99 Цитовано в 1823 джерелах Пов'язані статті Кількість версій: 13

Figure 1. Search by authors' names

✕

## Послатися

MLA Gentleman, Robert, et al., eds. "Bioinformatics and computational biology solutions using R and Bioconductor." (2008).

APA Gentleman, R., Carey, V., Huber, W., Irizarry, R., & Dudoit, S. (Eds.). (2008). Bioinformatics and computational biology solutions using R and Bioconductor.

ISO 690 GENTLEMAN, Robert, et al. (ed.). Bioinformatics and computational biology solutions using R and Bioconductor. 2008.

[BibTeX](#)   [EndNote](#)   [RefMan](#)   [RefWorks](#)

Figure 2. Generating the references

The screenshot shows the Google Scholar library interface. On the left, there is a navigation menu with options like 'Моя бібліотека', 'Кошик', and 'Керувати мітками...'. The main area displays search results for 'bioinformatics'. The first result is 'ExpASY: SIB bioinformatics resource portal' from oup.com. The second result is 'Bioinformatics and computational biology solutions using R and Bioconductor' from psu.edu, which is the book by Gentleman et al. mentioned in Figure 2. The third result is 'Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists' from oup.com. The fourth result is 'Bioinformatics: the machine learning approach' from books.google.com. The fifth result is 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources' from psu.edu. Each result includes a checkbox, a citation icon, and citation statistics.

Figure 3. Adding the selected publication to one's own library

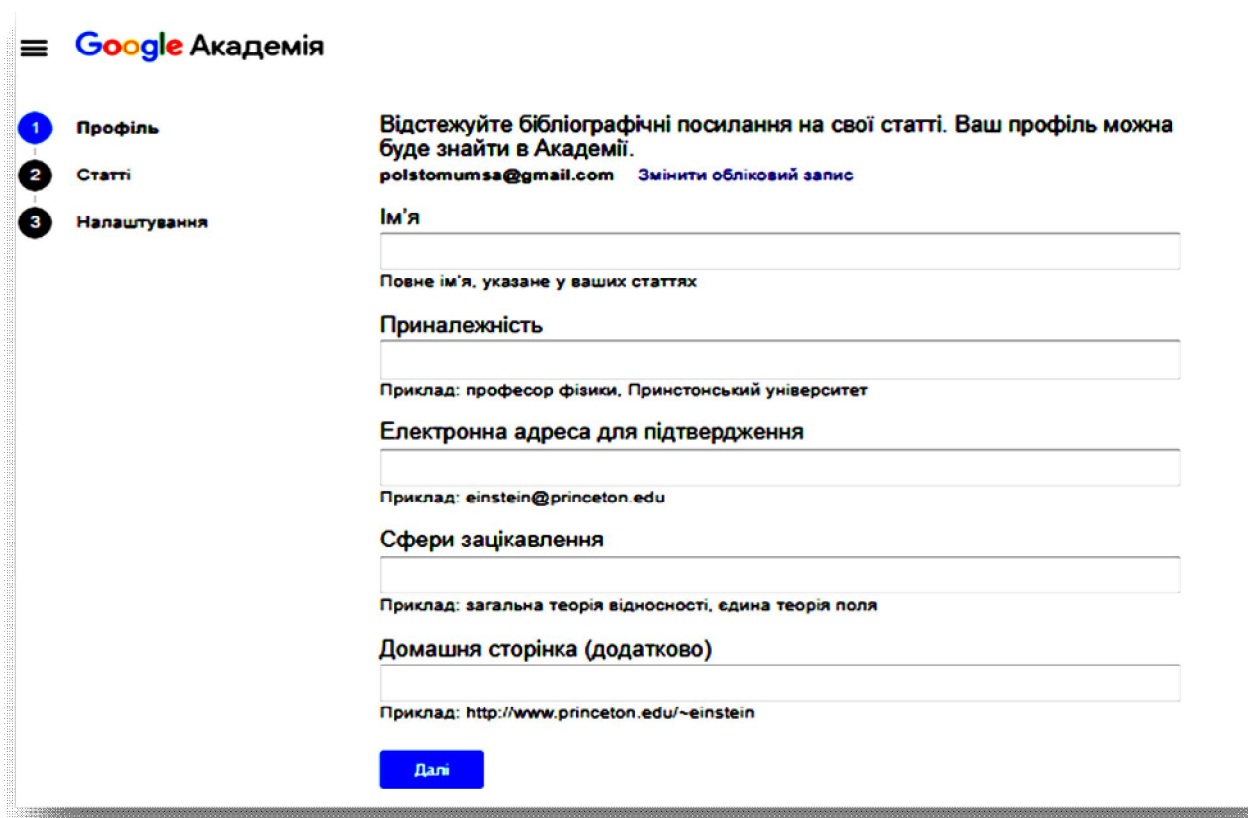


Figure 4. Tracing the citations by using Google Scholar

Web of Science is a scientific metric multidisciplinary database, which includes 15 databases, over 34 thousand journals and over 150 million documents (Fig. 5).

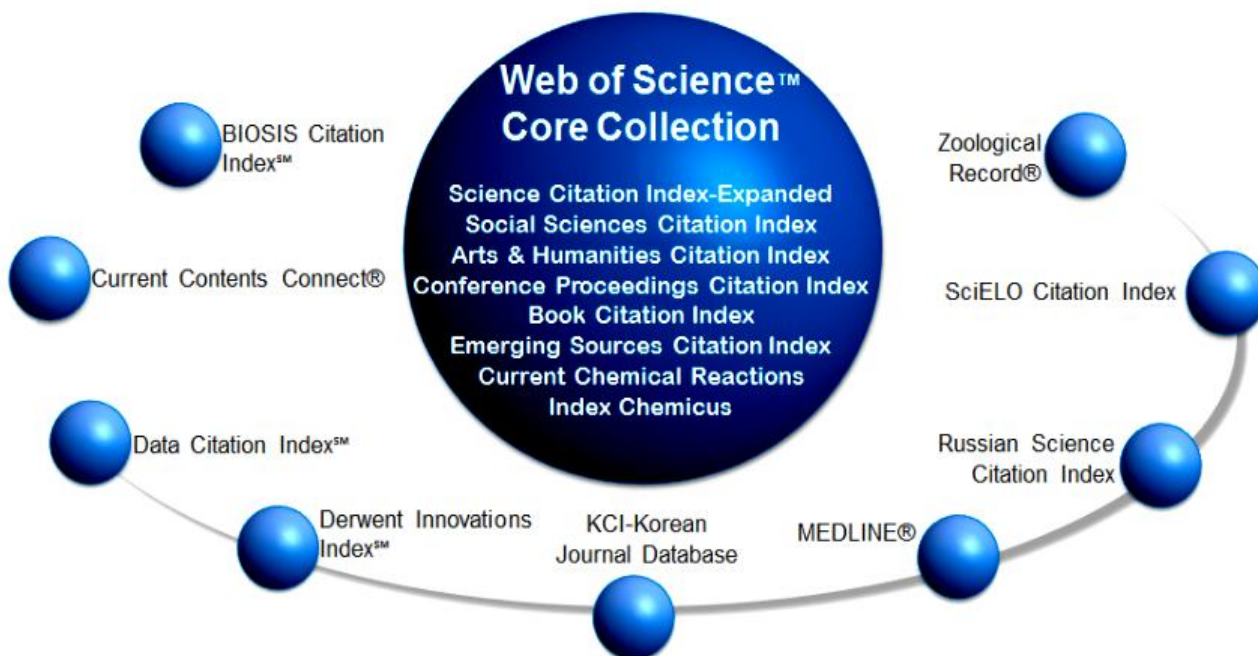


Figure 5. The structure of Web of Science

The database allows us to search by 254 subject categories. To start working with the database, one must follow the link <https://apps.webofknowledge.com> (Fig. 6).



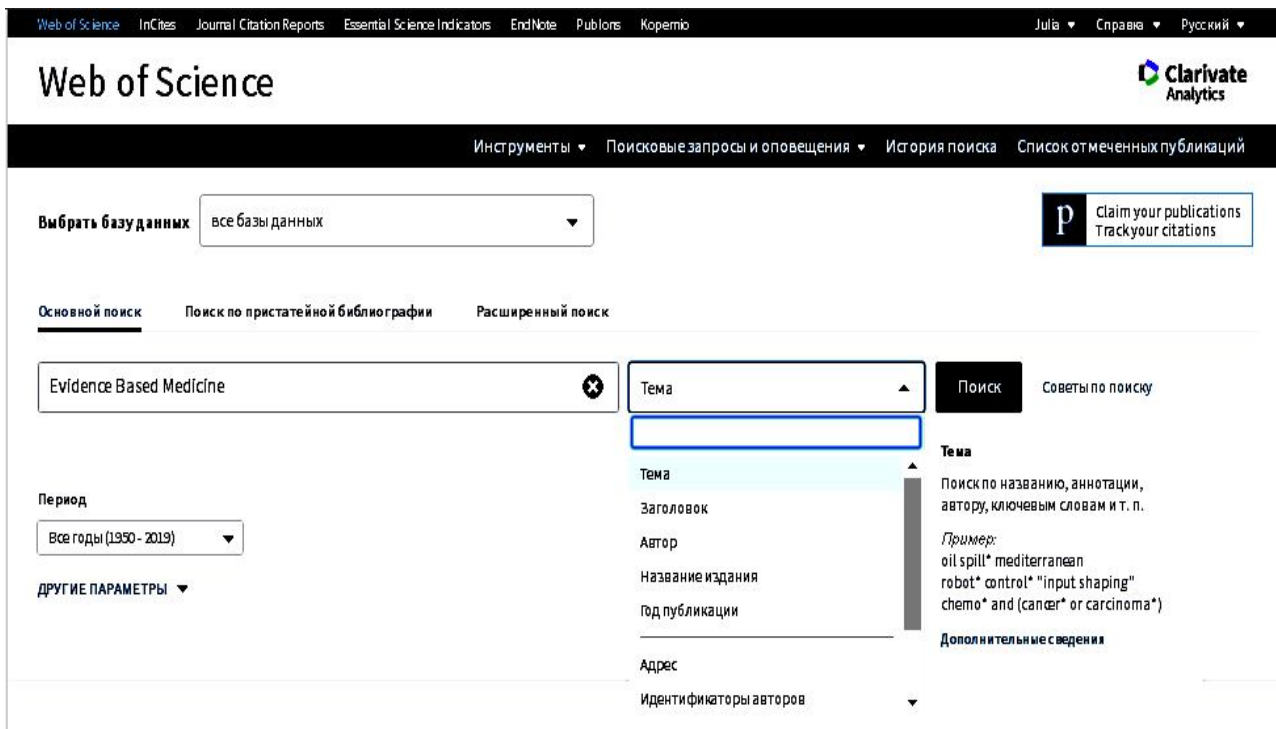


Figure 6. The welcome page of Web of Science

As a search query, one can choose the name of publication, keywords, etc. The search query is detailed by the title, author, and author's identifier (Fig. 9). The search result is shown in Fig. 9. As can be seen from Fig.

9, the search results contain the terms of search query in any order. To filter the results and reduce their number, the search query must be entered in the quotation marks (Fig. 7).

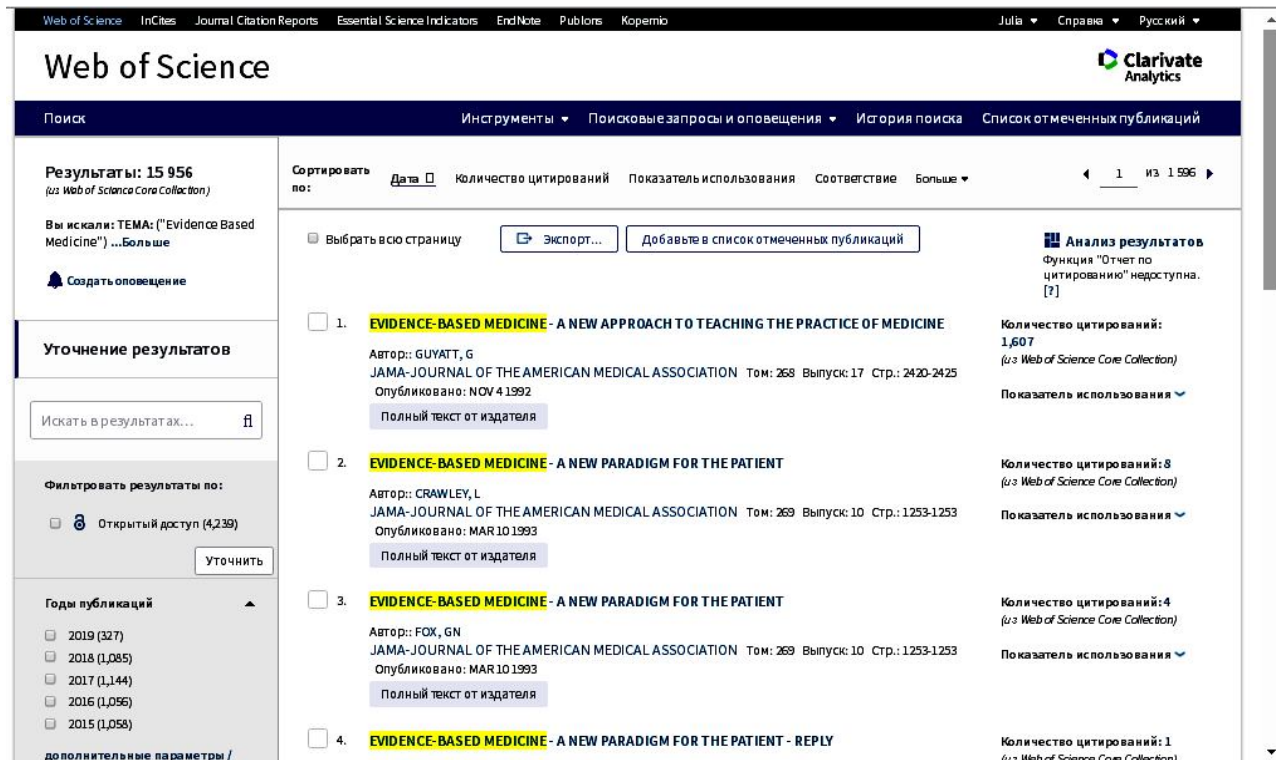


Figure 7. The search queries in Web of Science

The results refinement panel allows us to limit the years of publication (1970-2019), by categories, types of

documents (articles, reviews, and conference materials), organization profiles, type of access (Fig. 8).

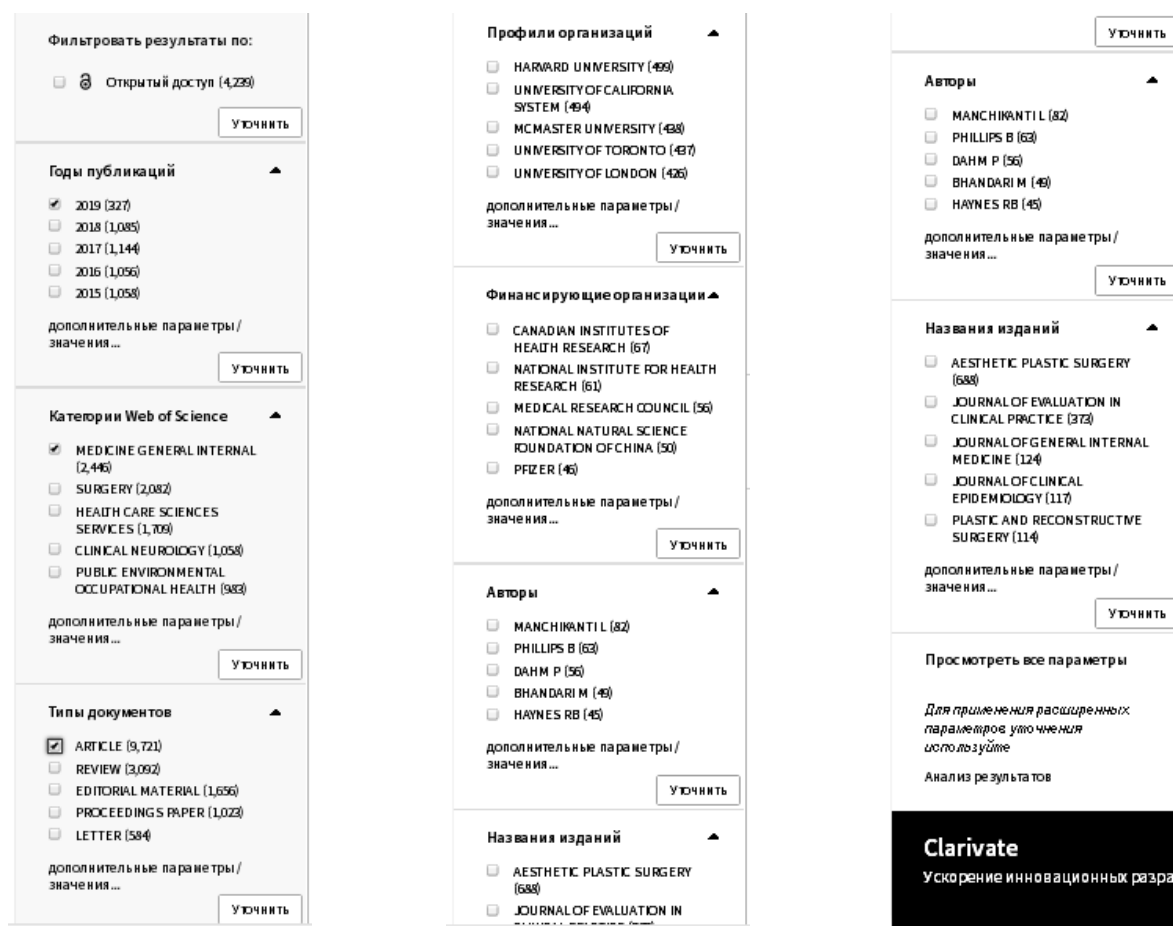


Figure 8. The results refinement panel

Thus, modern higher medical education requires a targeted and methodically thought-out algorithm of mastering the tools of citation databases, such as Google Scholar and Web of Science [8; 9; 10]. Training of Ph.D. students is accompanied by learning a large array of scientific information: articles, monographs, patents, etc. In connection with this, it is necessary to develop the mechanisms for search, sorting and analysis of scientific information, writing and registration of one's own research works. Google Scholar and Web of Science provide such mechanisms in the modern scientific environment including bibliographic descriptions of periodicals, expanded possibilities of searching for necessary information, medical information from directories, popular scientific literature, information about methods of treatment, information about certain medical institutions and individual doctors, results from recent clinical trials and new treatments. Adequately selected algorithms of mastering the tools of citation databases can significantly enhance and update the training course according to the modern requirements. The prospects for study consist in further in-depth research of the potential of Google Scholar, Web of Science and other databases in training medical and dental Ph.D. students.

### References

1. On Higher Education: The Law of Ukraine as of 27.07.2017 No. 1556-VIII.
2. Order of the Ministry of Education and Science "On the publication of the results of dissertations for obtaining scientific degrees of Doctor and Candidate of Sciences" as of December 17, 2012, No. 1112.

3. On Approval of the Procedure for Training Graduates of Higher Education in the Degree of Doctor of Philosophy and Doctor of Science at Higher Educational Institutions (Scientific Establishments): approved by the Resolution of the Cabinet of Ministers of Ukraine as of March 23, 2016 No. 261.
4. Order of the Ministry of Education and Science "On Approval of the Procedure for the Assignment of Scientific Rankings to Scientific and Educational Workers" as of January 14, 2016 No. 13.
5. On conducting experiment on awarding the degree of Doctor of Philosophy: approved by the decision of the Cabinet of Ministers of Ukraine as of March 6, 2019
6. Morokhovets HYu, Lysanets' YuV Developing the Professional Competence of Future Doctors in the Instructional Setting of Higher Medical Educational Institutions .Wiadomości Lekarskie. 2017;70(1):101-104.
7. Morokhovets H.Y., Uvarkina O.V., Bieliaieva O.M., Lysanets Yu.V., Senkevych H.A., Stetsenko S.A. Development of Motivation Towards Education in medical Students. Wiadomosci lekarskie. 2019;72(1):7-11.
8. Morokhovets HYu, Lysanets YuV Methods of effective teaching bioinformatics in medical academic setting. Medical and Ecological Problems. 2018;22(3-4):18-21.
9. Morokhovets HYu, Lysanets YuV, Silkova OV, Ostrovskaya LY, Purdenko TY Evidence-based medicine as a modern methodology in the field of public health. Bulletin of Biology and Medicine Issues. 2018;Issue 4, 1(146):160-166.
10. Morokhovets H.Yu. Discipline "Methodology of scientific research" in the system of training of doctors of philosophy of the field of knowledge "Healthcare". Ukrainian Professional Education. 2018;3:46-51.

Матеріал надійшов до редакції 09.04.2019