

# Immunoinformatics: A Basic Bibliometric Analysis

Elaheh Rahimpour<sup>1</sup>, Taha Samad-Soltani<sup>2\*</sup>

<sup>1</sup>Pharmaceutical Analysis Research Center and Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran

<sup>2</sup>Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran

## ARTICLE INFO

### Article History:

Received: 9 April 2021

Accepted: 17 April 2021

ePublished: 17 April 2021

### Keywords:

Immunoinformatics

Bibliometric

Knowledge body

Immunology

Informatics

## Abstract

**Background:** Traditionally, immunologists have been using experimental methods for a long time, which have generated functional, clinical and epidemiological data. So, it is inevitable to develop and apply novel computational approaches to collect, store, analyze, and report the latent knowledge behind these data. Immunoinformatics as a new-emerging field uses computational tools that cover multi dimensions of bioinformatics such as databases management and development definition of both structural and functional signatures and the modeling of predictive algorithms. Recently, with the dramatic growth in research documents publishing, the need for new methodologies to structure and visualize knowledge emerged.

**Methods:** Capability to analyze huge amounts of published items with a detailed level of granularity and report statistical, mathematical and visual knowledge within it, introduced an approach called bibliometrics analysis. According to the best of our knowledge, there is no bibliometrics analyses based on immunoinformatics keyword. So, the aim of current study was to perform a brief and basic bibliometrics analysis on the immunoinformatics publications to report the most influencers on this field by statistical and mathematical methods to identify key authors, titles, keywords and other parameters that extracted from publication's metadata.

**Results:** Results were summarized in statistical tables and diagrams. It was revealed that the immunoinformatics is a relatively novel field and the time trends showed that the strategic research plans can be directed to expand this field.

**Conclusion:** We can conclude that the future is the age of immunoinformatics for immunologists and related specialties.

## Introduction

In the viewpoint of informatics experts, an immune system may be considered as a complex network of connected molecules, which leads to many intertwined responses. The structure and function of these networks vary in different persons or organisms.<sup>1</sup> Complex networks generate a huge volume of data related to immune systems. Big Data is a common phrase to explain such networks. Traditionally, immunologists have been using experimental methods for a long time, which have generated functional, clinical and epidemiological data. So, it is inevitable to develop and apply novel computational approaches to collect, store, analyze, and report the latent knowledge behind these data.<sup>2</sup> In the recent years, informatics tools and software are appearing in the immunology context which help in understanding the structure and behavior of the whole immune system. This has given rise to a new keyword, called immunoinformatics.<sup>3</sup>

Immunoinformatics as a new-emerging field uses

computational tools that cover multi dimensions of bioinformatics such as databases management and development definition of both structural and functional signatures and the modeling of predictive algorithms. Immunoinformatics give us a better perception of the immune system of both man and animals to the effective fight against some less predictable pathogenesis.<sup>4</sup> Despite the recent growing interest in immunoinformatics, the current body of bibliometric knowledge regarding it is scarce. Although some studies have been conducted to review the state of immunoinformatics applications and methods, these have been limited to specific clinical and technical questions using topics and tables that displayed a summary of goals, tools, and outcomes of the original works.<sup>5,6</sup>

Recently, with the dramatic growth in the research documents publishing, the need for new methodologies to structure and visualize knowledge has also emerged. Capability to analyze huge amounts of published items

\*Corresponding Author: Taha Samad-Soltani, Email: [samadsoltani@tbzmed.ac.ir](mailto:samadsoltani@tbzmed.ac.ir)

© 2021 The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

with a detailed level of granularity and report statistical, mathematical and visual knowledge within it, introduced an approach called bibliometrics analysis. This approach summarized publication's metadata to an aggregated and intelligible representation.<sup>7</sup> It can be seen that the bibliometric analysis has been successfully applied in the different scopes of medicine.<sup>8,9</sup> But, according to the best of our knowledge, there isn't any bibliometrics analysis based on immunoinformatics keyword. So, the aim of current study was to perform a brief and basic bibliometrics analysis on the immunoinformatics publications to report the most influencers on this field by statistical and mathematical methods to identify key authors, titles, keywords and other parameters that extracted from publication's metadata. Our research allows researchers and other stakeholders to gain a more in-depth understanding of a field of immunoinformatics while also encouraging interdisciplinary collaboration.

## Methods

The current study is retrospective in nature and applies a bibliometric approach of metadata. This type of analysis generates useful information for researchers evaluating scientific activity. Document information was extracted from Scopus Database by searching titles in article tab. This database is multidisciplinary and has the highest coverage number of reputable journals in the world.<sup>10</sup> This research analyzes all publications without any date and time limitation. The keyword included in the search term was "immunoinformatics". The results yielded a total of 780 papers which were used to develop our knowledgebase. For each document, a list of metadata including citation information (Authors, title, year, source title, volume, issue, pages, citation count, source, document type, and DOI), bibliographical information (affiliation, identifier, PubMed ID, publisher, editors, language, correspondence address and abbreviated source title), and abstract and keywords (abstract, author keywords, and index keywords) were fetched from database. So,

our analysis focused on these metadata statistical and mathematical reports. We used SciMAT to analysis the generated knowledgebase.<sup>11</sup> It is an open source and free software with powerful algorithms to preprocess and process data. To preprocess, deduplication of documents, grouping of keywords and sorting the documents were applied. Finally, a rapid and basic bibliometrics analysis was performed. The earliest paper was published in 2000, so we categorized documents in four time periods every five to six years (2000-2005, 2006-2010, 2011-2015, 2015-2021). The details of each analysis will be clarified in the next section of the paper. More detailed information could be found in SciMAT manual.<sup>12</sup>

## Results

### *Journals and immunoinformatics*

Table 1 shows the most interesting journals to publish papers with the theme of immunoinformatics. According to the analysis, we reported the top ten to top twenty in each section based on the importance and count of materials. The results showed that diversity existed in journal titles and rank. Scientific reports journal has the highest citescore with 14 papers and the International Journal of Peptide Research and Therapeutics has published 37 papers with a 1.60 citescore .

### *Authors and Immunoinformatics*

Totally, 2813 distinct authors were contributed in 780 papers. Figure 1 displays the most influencers with more than 10 papers in the field of immunoinformatics according to our knowledgebase. Annie S. De Groot with Scopus H-index of 40 is the most influencing scientist in the immunoinformatics. She works in EpiVax, a privately-held biotechnology company based in Providence, RI that is dedicated to the development and implementation of computational immunology tools to benefit the biotechnology industry. The second researcher is Marie-Paule Lefranc, who is the founder and director of the International ImMunoGeneTics information system. Her

**Table 1.** Top 12 journals that published papers with immunoinformatics keyword in title, abstracts and keywords in the Scopus

Title of Journal	No of papers	Citescore	Q
International Journal of Peptide Research and Therapeutics	37	1.60	Q3
Methods in Molecular Biology	36	1.80	Q4
Journal of Biomolecular Structure and Dynamics	31	4.50	Q3
Infection, Genetics and Evolution	29	5	Q1
Frontiers in Immunology	24	5.40	Q2
Human Vaccines and Immunotherapeutics	20	4.70	Q2
Vaccine	18	5.50	Q1
Microbial Pathogenesis	15	4.40	Q2
Scientific Reports	14	7.20	Q1
Molecular Immunology	12	4.90	Q2
PLoS ONE	11	5.20	Q1
BMC Bioinformatics	11	4.59	Q1

major research areas are on the human immunoglobulin and T cell receptor genes.

### Affiliations and Immunoinformatics

Table 2 shows the 15 most motivated organizations around the immunoinformatics research. EpiVax as a knowledge-based company has 33 research projects based on keyword analysis. They reported their achievements and products in form of a published paper to improve the level of evidences in the immunology projects. The second is the school of pharmacy at Shiraz university of medical sciences in Iran. They generated some evidences about immunology and applied informatics. Next one is an institute in Rhode Island university. A specific institute was established in 2008 under the leadership of Annie De Groot for immunology and informatics to handle research and development projects. It seems the EpiVax and they have a strong relationship in this research field.

### Articles and Immunoinformatics

The top cited papers (more than 50 citations from Scopus) without limitation or any filters are listed in Table 3. Also most interesting papers within the last years (2019-2021) were extracted and displayed in Table 4. As expected most of highly cited papers are related to the recent pandemic caused by SARS-CoV-2.

### Time trends and immunoinformatics

Figures 2 and 3 display the time trend of publications with a focus on immunoinformatics. At first, we visualize it by one year interval in Figure 2. Then the length of time was categorized in four periods (every five years consisted of a period). The trend dramatically increased in 2020 due to the Covid-19 and vaccine development studies around this pandemic. It was clear that the immunoinformatics

transformed to a popular and critical scope in the immunology. It will be expected that the trend continues its growth in the future.

### Most frequent keywords and immunoinformatics

Last analysis extracts the most frequent keywords related to immunoinformatics in the generated knowledgebase. This analysis can help to visualize a theme of important

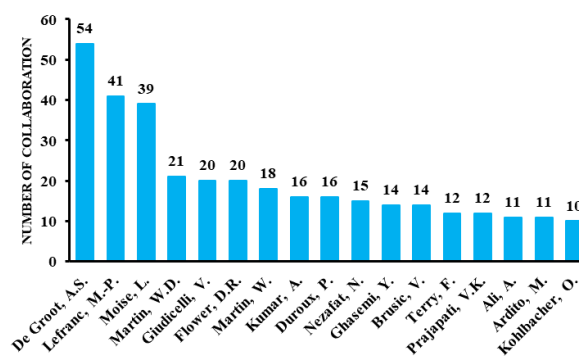


Figure 1. No of collaborations by distinct authors in the immunoinformatics

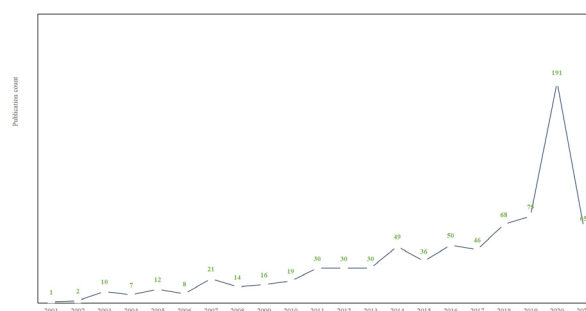


Figure 2. The count of publications in time interval = 1 year

Table 2. The 15 most motivated organizations around the immunoinformatics research

Affiliation	Doc Count
EpiVax, Inc., Providence, RI, United States	33
School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran	30
Institute for Immunology and Informatics, University of Rhode Island, Providence, RI, United States	20
Department of Genetic Engineering and Biotechnology, School of Life Sciences, Shahjalal University of Science and Technology, Sylhet, Bangladesh	5
Cellular and Molecular Biology Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran	5
Department of Bioinformatics, Rajendra Memorial Research Institute of Medical Sciences, Patna, 800007, India	5
Nutritional Immunology and Molecular Medicine Laboratory, Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA, United States	5
Department of Bioinformatics and Biotechnology, Government College University, Faisalabad, Pakistan	4
Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Islamic University, Kushtia, 7003, Bangladesh	4
School of Medicine, Western Sydney University, Campbelltown, NSW 2560, Australia	4
Department of Hepatitis and AIDS, Pasteur Institute of Iran, Tehran, Iran	4
Student Research Committee, Babol University of Medical Sciences, Babol, Iran	4
Iranian Comprehensive Hemophilia Care Center, Tehran, Iran	4
Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, 226028, India	4
Jenner Institute, University of Oxford, Compton, Berkshire, United Kingdom	4

**Table 3.** Top cited papers in the field of immunoinformatics based on keyword analysis.

Article Title	Pub Year	Citation Count
IMGT®, the international ImMunoGeneTics information system®	2009	481
IMGT® tools for the nucleotide analysis of immunoglobulin (IG) and t cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS	2012	253
IMGT R, the international ImMunoGeneTics information system R 25 years on	2015	212
T-cell epitope vaccine design by immunoinformatics	2013	192
IMGT/3dstructure-DB and IMGT/domaingalign: A database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF	2009	149
Immunoinformatics comes of age	2006	140
Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach	2020	130
Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein	2007	107
Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV	2020	104
Immunoinformatics and epitope prediction in the age of genomic medicine	2015	101
Recent advances in B-cell epitope prediction methods	2010	101
Emerging vaccine informatics	2010	95
Quantitative prediction of mouse class I MHC peptide binding affinity using support vector machine regression (SVR) models	2006	91
Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach	2017	88
Immunoglobulin and T cell receptor genes: IMGT® and the birth and rise of immunoinformatics	2014	88
Immunoinformatics: An integrated scenario	2010	87
Prediction of promiscuous peptides that bind HLA class I molecules	2002	86
Prediction of immunogenicity for therapeutic proteins: State of the art	2007	85
Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection	2017	83
Immunoinformatic comparison of T-cell epitopes contained in novel swine-origin influenza A (H1N1) virus with epitopes in 2008-2009 conventional influenza vaccine	2009	81
IMGT-ontology for immunogenetics and immunoinformatics	2004	81
Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study	2020	79
POPI: Predicting immunogenicity of MHC class I binding peptides by mining informative physicochemical properties	2007	79
HIV vaccine development by computer assisted design: The GAIA vaccine	2005	76
Toward more accurate pan-specific MHC-peptide binding prediction: A review of current methods and tools	2012	75
From genome to vaccine - New immunoinformatics tools for vaccine design	2004	75
Computer aided selection of candidate vaccine antigens	2010	71
In silico quantitative prediction of peptides binding affinity to human MHC molecule: An intuitive quantitative structure-activity relationship approach	2009	70
IMGT-Choreography for immunogenetics and immunoinformatics	2005	70
Novel Immunoinformatics Approaches to Design Multi-epitope Subunit Vaccine for Malaria by Investigating Anopheles Salivary Protein	2018	69
Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus	2017	69
NetMHCstab - Predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery	2014	68
Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines	2006	68
MHCPred: bringing a quantitative dimension to the online prediction of MHC binding.	2003	68
Mapping IgE binding epitopes of major shrimp (Penaeus monodon) allergen with immunoinformatics tools	2011	67
Low immunogenicity predicted for emerging avian-origin H7N9: Implication for influenza vaccine design	2013	66
T cell epitope: Friend or Foe? Immunogenicity of biologics in context	2009	64
Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins	2016	62
Designing an efficient multi-epitope peptide vaccine against Vibrio cholerae via combined immunoinformatics and protein interaction based approaches	2016	62
Genome-derived vaccines	2004	62
Integrated modeling of the major events in the MHC class I antigen processing pathway	2005	59

Table 3. Continued

Article Title	Pub Year	Citation Count
Diversity of <i>Francisella tularensis</i> Schu4 antigens recognized by T lymphocytes after natural infections in humans: Identification of candidate epitopes for inclusion in a rationally designed tularemia vaccine	2007	58
Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches	2017	57
IMGT-ONTOLOGY and IMGT databases, tools and Web resources for immunogenetics and immunoinformatics	2004	57
The two-faced T cell epitope: Examining the host-microbe interface with JanusMatrix	2013	56
MACppred: A support vector machine-based meta-predictor for identification of anticancer peptides	2019	54
A review of reverse vaccinology approaches for the development of vaccines against ticks and tick borne diseases	2016	54
Strength in numbers: Achieving greater accuracy in MHC-I binding prediction by combining the results from multiple prediction tools	2007	53

Table 4. Most interesting papers (Top five) in the last two years (2019-2021- Scopus)

Article Title	Pub Year	Citation Count	Journal	Ref
Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach	2020	130	Journal of Medical Virology	13
Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV	2020	104	Journal of medical virology	14
Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study	2020	79	Journal of Biomolecular Structure and Dynamics	15
MACppred: A support vector machine-based meta-predictor for identification of anticancer peptides	2019	54	International Journal of Molecular Sciences	16
Structure-based drug designing and immunoinformatics approach for SARS-CoV-2	2020	33	Science Advances	17

topics in immunoinformatics. Figure 4 shows the pie diagram of the keywords that repeated more than 125 times in the papers.

## Discussion

A basic bibliometrics analysis was conducted to identify and explain the most influential persons, journals and, papers which led to development of immunoinformatics term in scientific settings. Some statistical methods and tools were used to achieve these goals and some keynotes were seen in the extracted information and patterns.

Immunoinformatics involves the research and development of algorithms for mapping possible B- and T-cell epitopes, which reduce the time and expense of pathogen gene product analysis in the lab. An immunologist may use this knowledge to investigate potential binding sites, which contributes to the creation of new vaccines.<sup>1</sup> As can be seen, the epitope is the most frequent keyword due to its effective role in the immunology reactions which knowing the epitope type can be helpful in immunoinformatic calculations. If we discuss more specifically, antigen-antibody reaction is the basic interaction in immunoassays that arise from epitope – paratope interactions. Actually, the epitope is part of antigen that is capable of stimulating an immune response and like other proteins, it is composed of amino acid sequence which also has a high number of the most frequently repeated words in published studies.

According to the World Health Organization, the COVID-19 pandemic has resulted in many deaths worldwide, especially among the elderly, and has resulted in an increasingly growing number of publications.<sup>18</sup> In the case of the recent SARS-CoV-2 outbreak,

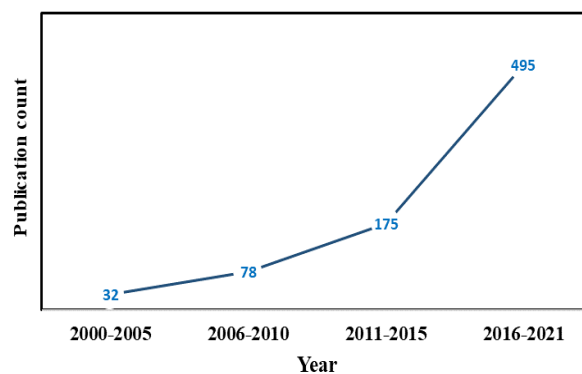


Figure 3. The count of publications in time interval = 5 years

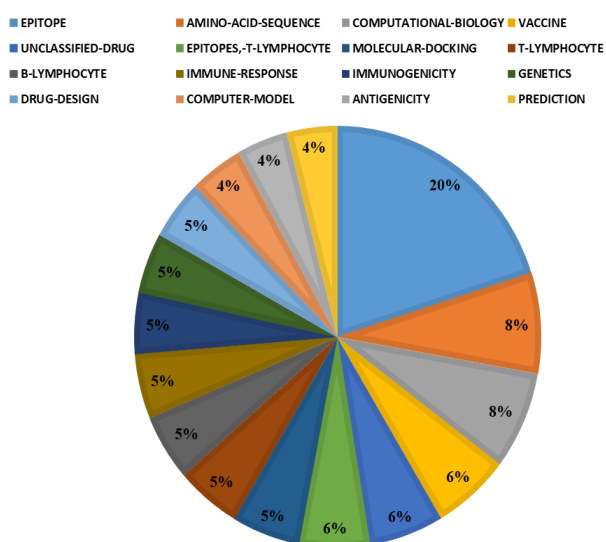


Figure 4. Most frequent keywords in the published works

bibliometric analysis and visualized mapping will aid in the monitoring of research methodologies based on omics or computational sciences.<sup>19</sup> Due to the nature of immunoinformatics in vaccine development and the emergence of Covid-19 in recent years, the popularity of the immunoinformatics to Covid-19 vaccine design and development was increased. Therefore, it is not surprising that the four of five articles with the highest citations described some level of this process.

### Conclusion

On basis of a basic bibliometric indicator analyses, for the period of 2000–2021, the role of knowledge-based companies and operational research institutes in the popularity of the immunoinformatics term was identified. Also, it was revealed that the immunoinformatics is a relatively novel field and the time trends showed that the strategic research plans can be directed to expand this field. We can conclude that the future is the age of immunoinformatics for immunologists and related specialties. Our study suggests a deeper bibliometrics analysis to extract hidden knowledge behind the network of keywords, collaborators and institutions.

### Conflicts of Interests

There are no conflicts of interest to declare.

### Acknowledgment

Thank our colleagues from Tabriz University of Medical Sciences who provided insight and expertise that greatly assisted the research.

### References

- Tomar N, De RK. Immunoinformatics: an integrated scenario. *Immunology*. 2010;131(2):153–68. doi:10.1111/j.1365-2567.2010.03330.x
- Bragazzi NL, Gianfredi V, Villarini M, Rosselli R, Nasr A, Hussein A, et al. Vaccines Meet Big Data: State-of-the-Art and Future Prospects. From the Classical 3Is (“Isolate–Inactivate–Inject”) Vaccinology 1.0 to Vaccinology 3.0, Vaccinomics, and Beyond: A Historical Overview. *Front Public Heal*. 2018;6. doi:10.3389/fpubh.2018.00062
- Chakraborty C, Sharma AR, Bhattacharya M, Sharma G, Lee S-S. Immunoinformatics Approach for the Identification and Characterization of T Cell and B Cell Epitopes towards the Peptide-Based Vaccine against SARS-CoV-2. *Arch Med Res*. 2021. doi:10.1016/j.arcmed.2021.01.004
- Oli AN, Obialor WO, Ifeanyichukwu MO, Odimegwu DC, Okoyeh JN, Emechebe GO, et al. Immunoinformatics and Vaccine Development: An Overview. *ImmunoTargets Ther*. 2020;Volume 9:13–30. doi:10.2147/ITT.S241064
- Raoufi E, Hemmati M, Eftekhari S, Khaksaran K, Mahmodi Z, Farajollahi MM, et al. Epitope Prediction by Novel Immunoinformatics Approach: A State-of-the-art Review. *Int J Pept Res Ther*. 2020;26(2):1155–63. doi:10.1007/s10989-019-09918-z
- Hegde NR, Gauthami S, Sampath Kumar HM, Bayry J. The use of databases, data mining and immunoinformatics in vaccinology: where are we? *Expert Opin Drug Discov*. 2018;13(2):117–30. doi:10.1080/17460441.2018.1413088
- Kokol P, Blažun Vošner H, Završnik J. Application of bibliometrics in medicine: a historical bibliometrics analysis. *Heal Inf Libr J*. 2020;hir.12295. doi:10.1111/hir.12295
- Harande YI. Exploring the literature of diabetes in Nigeria: a bibliometrics study. *African J Diabetes Med*. 2011;19(2).
- Stelzer B, Meyer-Brötz F, Schiebel E, Brecht L. Combining the scenario technique with bibliometrics for technology foresight: The case of personalized medicine. *Technol Forecast Soc Change*. 2015;98:137–56. doi:10.1016/j.techfore.2015.06.008
- Burnham JF. Scopus database: a review. *Biomed Digit Libr*. 2006;3(1):1. doi:10.1186/1742-5581-3-1
- Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. SciMAT: A new science mapping analysis software tool. *J Am Soc Inf Sci Technol*. 2012;63(8):1609–30. doi:10.1002/asi.22688
- SciMAT User Guide 2020. Available from: <https://sci2s.ugr.es/scimat/software/v1.01/SciMAT-v1.0-userGuide.pdf>
- Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, et al. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. *J Med Virol*. 2020;92(6):618–31. doi:10.1002/jmv.25736
- Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J Med Virol*. 2020;92(5):495–500. doi:10.1002/jmv.25698
- Enayatkhani M, Hasaniazad M, Faezi S, Gouklani H, Davoodian P, Ahmadi N, et al. Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study. *J Biomol Struct Dyn*. 2021;39(8):2857–72. doi:10.1080/07391102.2020.1756411
- Boopathi V, Subramaniyam S, Malik A, Lee G, Manavalan B, Yang D-C. mACPpred: A Support Vector Machine-Based Meta-Predictor for Identification of Anticancer Peptides. *Int J Mol Sci*. 2019;20(8):1964. doi:10.3390/ijms20081964
- Panda PK, Arul MN, Patel P, Verma SK, Luo W, Rubahn H-G, et al. Structure-based drug designing and immunoinformatics approach for SARS-CoV-2. *Sci Adv*. 2020;6(28):eabb8097. doi:10.1126/sciadv.abb8097
- Safieddine M, Kassir R. COVID 19 and the race to publish: an ethical issue. *J Br Surg*. 2020;107(11):e504–e504.
- Murillo J, Villegas LM, Ulloa-Murillo LM, Rodríguez AR. Recent trends on omics and bioinformatics approaches to study SARS-CoV-2: A bibliometric analysis and mini-review. *Comput Biol Med*. 2021;128:104162. doi:10.1016/j.combiomed.2020.104162