

Review

Not peer-reviewed version

---

# Integrating Bioinformatics in Drug Discovery

---

[Mujahed Mustafa](#)<sup>\*</sup>, Ahmed Mohammed, [Awad Alzebair](#), Abdelrafie Makhawi

Posted Date: 24 October 2023

doi: 10.20944/preprints202310.1465.v1

Keywords: drug discovery; target identification; target validation; lead identification; bioinformatics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mini-review

# Integrating Bioinformatics in Drug Discovery

Mujahed I. Mustafa <sup>1,\*</sup>, Ahmed Mohammed <sup>2</sup>, Awad A. Alzebair <sup>2</sup> and Abdelrafie M. Makhawi <sup>1</sup>

<sup>1</sup> Department of Biotechnology, College of Applied and Industrial Sciences, University of Bahri, Khartoum, Sudan

<sup>2</sup> Department of Biotechnology, school of life sciences and technology, Omdurman Islamic University, Omdurman, Sudan

\* Correspondence: mujahedimustafa@gmail.com

**Abstract:** The integration of bioinformatics in drug discovery has revolutionized the field of pharmaceutical research. The use of computational tools and techniques has enabled researchers to analyze vast amounts of data, identify potential drug targets, and design new drugs with greater precision and efficiency. The integration of bioinformatics has also facilitated the development of personalized medicine, where drugs can be tailored to individual patients based on their genetic makeup. However, there are still challenges that need to be addressed, such as the need for more accurate predictive models and the ethical considerations surrounding the use of patient data. Overall, the integration of bioinformatics in drug discovery holds great promise for improving human health and advancing our understanding of disease mechanisms. In this mini-review, we discuss how Bioinformatics plays a crucial role in each step of drug discovery by providing tools and techniques to analyze large amounts of data generated from various sources, as well as the challenges and the opportunities offered by bioinformatics.

**Keywords:** drug discovery; target identification; target validation; lead identification; bioinformatics

---

## 1. Introduction

The integration of bioinformatics in drug discovery has revolutionized the pharmaceutical industry by accelerating the process of identifying potential drug candidates and optimizing their efficacy [1]. Bioinformatics, a multidisciplinary field that combines biology, computer science, and statistics, provides powerful tools and techniques for analyzing vast amounts of biological data. By leveraging these tools, researchers can gain valuable insights into the complex interactions between drugs and biological systems [2–4].

One of the key challenges in drug discovery is identifying suitable targets for therapeutic intervention. Bioinformatics plays a crucial role in this process by mining large-scale genomic and proteomic datasets to identify potential drug targets. By analyzing gene expression patterns, protein-protein interactions, and metabolic pathways, bioinformaticians can identify key molecules that are involved in disease processes. This information is then used to design drugs that specifically target these molecules, increasing the chances of therapeutic success [5].

Another important aspect of drug discovery is predicting the efficacy and safety of potential drug candidates [6]. Traditional methods for testing drugs *in vitro* or *in vivo* are time-consuming and expensive. Bioinformatics offers an alternative approach by using computational models to predict how a drug will interact with its target molecule. These models take into account factors such as molecular structure, binding affinity, and pharmacokinetics to estimate the effectiveness of a drug candidate. This allows researchers to prioritize promising candidates for further testing, saving time and resources [7].

Furthermore, bioinformatics plays a crucial role in optimizing the design of clinical trials. By analyzing patient data from previous trials or publicly available databases, researchers can identify biomarkers that are associated with treatment response or adverse reactions. This information can be

used to stratify patients into subgroups based on their genetic profiles or other relevant factors. By tailoring treatments to specific patient populations, researchers can increase the chances of success in clinical trials while minimizing side effects [8].

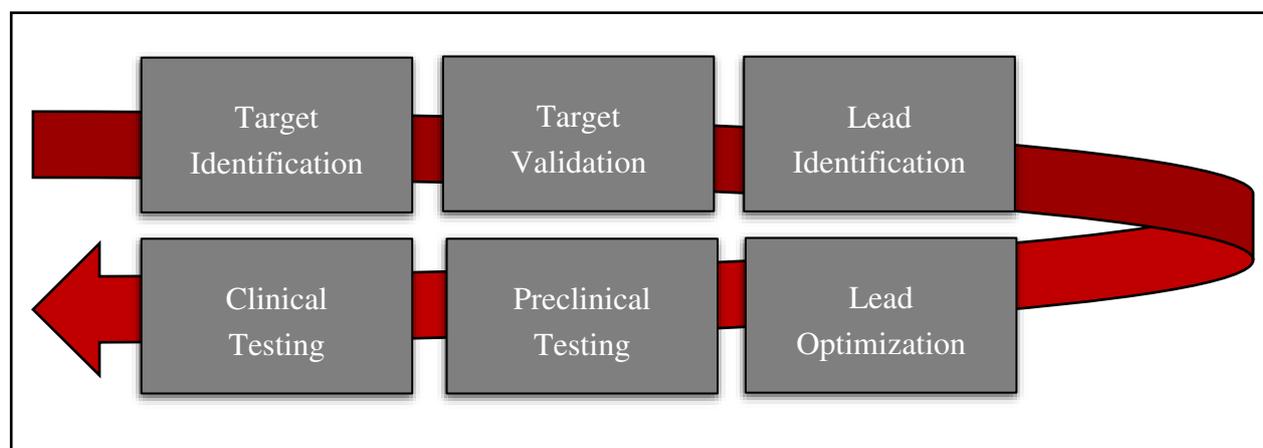
In addition to these applications, bioinformatics also contributes to the field of personalized medicine. By integrating genomic data with clinical information, bioinformaticians can identify genetic variations that are associated with drug response or disease susceptibility. This information can be used to develop personalized treatment plans that take into account an individual's unique genetic makeup [9]. This approach has the potential to revolutionize healthcare by improving treatment outcomes and reducing adverse reactions [10].

Despite its numerous advantages, the integration of bioinformatics in drug discovery also presents challenges. The analysis of large-scale biological datasets requires sophisticated computational infrastructure and expertise [11]. Additionally, the interpretation of complex data requires a deep understanding of both biology and statistics. Therefore, collaboration between biologists, computer scientists, and statisticians is crucial for successful implementation of bioinformatics in drug discovery.

In this mini-review, we discuss how Bioinformatics plays a crucial role in each step of drug discovery by providing tools and techniques to analyze large amounts of data generated from various sources, as well as the challenges and the opportunities offered by bioinformatics.

## 2. The crucial steps of drug discovery by using bioinformatics

Drug discovery is a complex and time-consuming process that involves the identification of new compounds that can be used to treat diseases [12]. Bioinformatics has become an essential tool in this process, allowing researchers to analyze large amounts of data and identify potential drug targets. In this article, we will discuss the crucial steps of drug discovery (Figure 1) by using bioinformatics:



**Figure 1.** The crucial steps of drug discovery by using bioinformatics approach.

### 2.1. Target identification

Identification of potential drug targets is a crucial step in the drug discovery process [13]. Bioinformatics has emerged as a powerful tool for identifying potential drug targets by analyzing large amounts of biological data. Bioinformatics involves the use of computational methods to analyze and interpret biological data, including genomic, proteomic, and metabolomic data.

One of the major advantages of using bioinformatics for drug target identification is that it allows researchers to analyze large amounts of data quickly and efficiently [14]. For example, genomic data can be analyzed to identify genes that are overexpressed or mutated in disease states. Proteomic data can be used to identify proteins that are differentially expressed or post-translationally modified in disease states. Metabolomic data can be used to identify metabolites that are altered in disease states [15].

Another advantage of using bioinformatics for drug target identification is that it allows researchers to identify potential drug targets that may not have been previously considered. For example, a protein that was previously thought to have no role in a particular disease may be identified as a potential drug target based on its differential expression or post-translational modification in the disease state [16].

Bioinformatics also allows researchers to predict the efficacy and safety of potential drugs before they are tested in clinical trials. This is done by using computational models to predict how a drug will interact with its target protein and other proteins in the body. This information can then be used to optimize the design of the drug and minimize any potential side effects.

Despite its many advantages, there are also some challenges associated with using bioinformatics for drug target identification. One challenge is the need for high-quality biological data. The accuracy and reliability of bioinformatics predictions depend on the quality of the input data, so it is important to ensure that the data used for analysis is accurate and reliable [17].

Another challenge is the need for expertise in both biology and computational methods. Bioinformatics requires expertise in both fields, so it can be challenging for researchers who are not familiar with both areas to effectively use bioinformatics for drug target identification [18].

## 2.2. Target validation

Target validation is the second step of drug discovery using bioinformatics. It is a crucial step in the drug discovery process as it helps to identify and validate potential drug targets that can be used to develop new drugs. Target validation involves the use of various bioinformatics tools and techniques to identify and validate potential drug targets.

The first step in target validation is to identify potential drug targets. This can be done by analyzing various biological data such as gene expression data, protein-protein interaction data, and pathway analysis data. Once potential drug targets have been identified, the next step is to validate them [19].

Validation of potential drug targets involves several steps. The first step is to confirm that the target is biologically relevant and plays a role in the disease or condition being targeted. This can be done by analyzing gene expression data, protein-protein interaction data, and pathway analysis data [20,21].

The next step in target validation is to determine whether the target can be modulated by a drug. This can be done by using various bioinformatics tools such as molecular docking, virtual screening, and molecular dynamics simulations [22]. These tools help to predict how a potential drug will interact with the target and whether it will be able to modulate its activity.

Once a potential drug has been identified, it is important to test its efficacy and safety in preclinical studies before moving on to clinical trials. Preclinical studies involve testing the drug on animal models to determine its efficacy and safety [23].

## 2.3. Lead identification

Drug discovery is a complex and time-consuming process that involves several steps, including target identification, lead generation, lead optimization, and preclinical testing. Bioinformatics plays a crucial role in each of these steps by providing tools and techniques to analyze large amounts of data generated from various sources.

Lead identification is the third step in the drug discovery process, where potential drug candidates are identified based on their ability to interact with the target molecule. This step involves screening large libraries of compounds to identify those that have the desired biological activity [24].

Bioinformatics tools are used extensively in lead identification to analyze and interpret data from various sources such as high-throughput screening (HTS), virtual screening, and structure-based drug design. HTS involves testing thousands of compounds simultaneously against a specific target molecule [25,26]. The data generated from HTS is analyzed using bioinformatics tools such as clustering algorithms, principal component analysis (PCA), and machine learning algorithms to identify compounds with the desired biological activity [26].

Virtual screening is another approach used in lead identification, where computer simulations are used to predict the binding affinity of compounds with the target molecule. This approach involves generating 3D models of both the target molecule and potential drug candidates and using molecular docking algorithms to predict their interaction [27].

Structure-based drug design is another approach used in lead identification that involves designing compounds based on the 3D structure of the target molecule. Bioinformatics tools such as molecular modeling software are used to generate 3D models of the target molecule, which can be used to design compounds that fit into specific binding sites.

#### 2.4. Lead optimization

Lead optimization is the fourth step in drug discovery, which involves the use of bioinformatics to identify and optimize potential drug candidates [28]. This process is critical in drug development as it helps to improve the efficacy, safety, and pharmacokinetic properties of a lead compound.

Bioinformatics is an interdisciplinary field that combines biology, computer science, and statistics to analyze and interpret biological data. In drug discovery, bioinformatics plays a crucial role in identifying potential drug targets, designing new compounds, and optimizing lead compounds.

The lead optimization process begins with the identification of a lead compound that has shown promising activity against a specific target. The lead compound is then subjected to various modifications to improve its potency, selectivity, and pharmacokinetic properties.

Bioinformatics tools are used to analyze the structure-activity relationship (SAR) of the lead compound. SAR analysis involves studying how changes in the chemical structure of a compound affect its biological activity. This information is used to design new compounds with improved activity against the target [29]. In addition to SAR analysis, bioinformatics tools are also used for molecular modeling and simulation studies. Molecular modeling involves using computer software to predict the three-dimensional structure of a molecule. This information can be used to design new compounds that fit better into the target site.

Simulation studies involve using computer models to simulate how a drug interacts with its target *in vivo* [26]. This information can be used to optimize the pharmacokinetic properties of a lead compound by predicting how it will be absorbed, distributed, metabolized, and excreted by the body.

Another important aspect of lead optimization is toxicity prediction. Bioinformatics tools can be used to predict potential toxicities associated with a lead compound before it enters clinical trials. This helps researchers identify potential safety issues early on in drug development [30].

#### 2.5. Preclinical testing

Preclinical testing is the fifth step in drug discovery, and it is a crucial stage in the development of new drugs. This stage involves testing the drug candidate in animals to determine its safety and efficacy before it can be tested in humans. Bioinformatics plays a vital role in preclinical testing by providing tools and techniques for analyzing large amounts of data generated during this stage.

The goal of preclinical testing is to identify potential safety issues and determine the optimal dose for the drug candidate [31]. This stage involves conducting various tests on animals, including toxicity studies, pharmacokinetic studies, and efficacy studies. The results of these tests are used to determine whether the drug candidate is safe and effective enough to move on to clinical trials.

Bioinformatics provides a powerful set of tools for analyzing data generated during preclinical testing. These tools include data mining, machine learning, and statistical analysis techniques that can help researchers identify patterns and relationships within large datasets. By using bioinformatics tools, researchers can gain insights into how the drug candidate interacts with biological systems, identify potential side effects, and optimize dosing regimens.

One example of how bioinformatics is used in preclinical testing is through pharmacokinetic modeling [32]. Pharmacokinetics refers to how drugs are absorbed, distributed, metabolized, and excreted by the body. By using bioinformatics tools to model pharmacokinetics, researchers can

predict how a drug will behave in different animal models or human populations. This information can be used to optimize dosing regimens or identify potential safety issues.

Another example of how bioinformatics is used in preclinical testing is through toxicology screening [33]. Toxicology screening involves testing a drug candidate for potential toxic effects on different organs or systems within an animal's body. By using bioinformatics tools to analyze toxicology data, researchers can identify patterns that may indicate potential toxicity issues or help them understand how a drug interacts with different biological systems.

## 2.6. Clinical trials

Clinical trials are an essential part of the drug discovery process, and they represent the sixth step in this complex and challenging journey. In these trials, researchers test the safety and efficacy of new drugs on human subjects, with the goal of bringing them to market and improving patient outcomes.

Bioinformatics plays a critical role in clinical trials, helping researchers to analyze large amounts of data generated during these studies. By using computational tools and algorithms, bioinformaticians can identify patterns in patient data that may be relevant to drug development.

One key area where bioinformatics is used in clinical trials is in the analysis of genomic data. By sequencing patients' DNA, researchers can identify genetic variations that may impact how a drug is metabolized or how it interacts with specific proteins in the body. This information can help guide dosing decisions and improve patient outcomes [34,35].

Another important application of bioinformatics in clinical trials is in the analysis of proteomic data. Proteomics involves studying the structure and function of proteins, which play a critical role in many biological processes. By analyzing protein expression levels and interactions within patients' bodies, researchers can gain insights into how drugs are working at a molecular level.

In addition to genomics and proteomics, bioinformatics is also used to analyze other types of data generated during clinical trials, such as imaging data or electronic health records. By integrating these different sources of information, researchers can gain a more comprehensive understanding of how drugs are affecting patients' health [36].

Overall, clinical trials represent a crucial step in the drug discovery process, and bioinformatics plays an essential role in making these studies more efficient and effective. By leveraging computational tools to analyze large amounts of data generated during these trials, researchers can identify new insights into drug development that may ultimately lead to better treatments for patients.

## 3. Bioinformatics tools and techniques

Bioinformatics tools and techniques are essential for analyzing large-scale genomic, proteomic, and metabolomic data sets. These tools include databases for storing and sharing biological data, algorithms for analyzing complex data sets, and visualization tools for interpreting results.

### 3.1. Genomics, Proteomics and Metabolomics

Bioinformatics is a rapidly growing field that combines biology, computer science, and statistics to analyze and interpret biological data. In drug discovery, bioinformatics tools and techniques are used to identify potential drug targets, design new drugs, and optimize drug efficacy.

Genomics is the study of an organism's complete set of DNA, including all of its genes [37]. In drug discovery, genomics is used to identify genetic variations that may be associated with disease susceptibility or drug response. This information can be used to develop personalized medicine approaches that target specific genetic mutations.

Proteomics is the study of an organism's complete set of proteins [38]. In drug discovery, proteomics is used to identify proteins that are involved in disease pathways or drug targets. Proteomics can also be used to study protein-protein interactions and post-translational modifications that may affect protein function.

Metabolomics is the study of an organism's complete set of metabolites, which are small molecules produced by cellular metabolism [39]. In drug discovery, metabolomics is used to identify metabolic pathways that are dysregulated in disease states or affected by drugs. Metabolomics can also be used to identify biomarkers for disease diagnosis or monitoring.

One example of a bioinformatics tool used in drug discovery is the Kyoto Encyclopedia of Genes and Genomes (KEGG), which provides a comprehensive database of biological pathways and molecular interactions [40]. KEGG can be used to identify potential drug targets based on their involvement in disease pathways [41].

Another example is the Protein Data Bank (PDB), which provides a database of three-dimensional structures for proteins and other biomolecules. PDB can be used to design new drugs that target specific protein structures or interactions [42].

### 3.2. *Molecular Docking*

In recent years, bioinformatics has become an essential approach in drug discovery and development. One of the most important tools of bioinformatics is molecular docking.

Molecular docking is a computational technique that plays a vital role in drug discovery using bioinformatics. It involves the prediction of the binding affinity and orientation of small molecules, such as drugs, to their target proteins. This technique has revolutionized the drug discovery process by reducing the time and cost required for drug development [43].

One of the significant benefits of molecular docking is its ability to predict the binding affinity between a drug and its target protein accurately. This information is crucial in determining the efficacy of a drug candidate before it undergoes further testing. Molecular docking can also predict potential off-target effects, which can help researchers identify potential side effects early in the drug development process.

Another benefit of molecular docking is its ability to identify novel drug candidates. By screening large databases of compounds, researchers can identify molecules that have similar structures to known drugs but have not yet been tested for their therapeutic potential. This approach has led to the discovery of several new drugs, including anti-cancer agents and antibiotics.

Molecular docking also allows researchers to optimize existing drugs by predicting how modifications to their chemical structure will affect their binding affinity with target proteins. This approach has led to the development of more potent and selective drugs with fewer side effects [44].

Moreover, molecular docking can be used to study protein-protein interactions, which play a crucial role in many biological processes. By predicting how proteins interact with each other, researchers can gain insights into disease mechanisms and develop new therapeutic strategies.

## 4. **Case studies on successful integration of bioinformatics in drug discovery**

The use of computational tools and techniques has enabled researchers to analyze vast amounts of data, identify potential drug targets, and design new drugs with greater precision and efficiency. As an examples, two case studies on successful integration of bioinformatics in drug discovery: discovery of anti-cancer drugs using genomic data analysis and development of HIV protease inhibitors through molecular docking simulations. Genomic data analysis has emerged as a powerful tool in identifying genetic mutations and alterations that contribute to cancer development and progression [45]. The Cancer Genome Atlas (TCGA) project is one of the largest cancer genomics initiatives that has generated comprehensive genomic data for various types of cancers. The integration of TCGA data with bioinformatics tools such as pathway analysis, machine learning algorithms, and network analysis has led to the discovery of several potential targets for anti-cancer drug development [46,47].

One such example is the discovery of a new class of anti-cancer drugs, called PARP inhibitors, for the treatment of breast and ovarian cancers [48]. PARP inhibitors work by targeting cancer cells with defects in the DNA repair pathway. The identification of these targets was made possible by analyzing TCGA data for genetic mutations and alterations in the DNA repair pathway. Furthermore, bioinformatics tools were used to identify potential drug candidates that could selectively target

cancer cells with DNA repair defects. Several PARP inhibitors have now been approved by the FDA for the treatment of breast and ovarian cancers [49].

Inhibition of HIV protease has been a successful strategy in the development of anti-HIV drugs. Molecular docking simulations are a type of bioinformatics tool that can predict the binding affinity and interactions between small molecules and target proteins. This approach has been widely used in the development of HIV protease inhibitors [50–52].

One such example is the development of a highly potent HIV protease inhibitor, called darunavir. Darunavir was developed using a combination of molecular docking simulations and X-ray crystallography. The structure of the HIV protease was used as a template for virtual screening of millions of small molecules. Bioinformatics tools were used to analyze the binding affinity and interactions between the small molecules and the HIV protease. Darunavir was found to have a high binding affinity and specificity for HIV protease and was subsequently approved by the FDA for the treatment of HIV [52].

## 5. Challenges and future directions

Bioinformatics has revolutionized the field of drug discovery by providing a wealth of data and tools for analyzing biological systems. However, there are still many challenges to be addressed in order to fully integrate bioinformatics into the drug discovery process.

One of the biggest challenges in integrating bioinformatics into drug discovery is the need to integrate data from multiple sources. This includes data from different types of experiments, such as genomics, proteomics, metabolomics, and imaging studies. In addition, there are often differences in data formats and standards used by different research groups [53].

To address these issues, there is a need for better data integration and standardization. This can be achieved through the development of common data formats and ontologies that allow different types of data to be easily integrated and compared. In addition, there is a need for better tools for data visualization and analysis that can help researchers make sense of complex datasets.

Another challenge in integrating bioinformatics into drug discovery is the need for more accurate predictive models. While there have been significant advances in machine learning algorithms and other computational methods for predicting drug targets and efficacy, these models are still far from perfect.

To improve predictive models, there is a need for more comprehensive datasets that include information on both positive and negative outcomes. In addition, there is a need for better methods for validating predictive models using independent datasets.

Finally, one of the most important factors in integrating bioinformatics into drug discovery is interdisciplinary collaboration. Drug discovery involves many different disciplines, including biology, chemistry, pharmacology, and computer science.

To fully leverage the power of bioinformatics in drug discovery, it is essential that researchers from these different disciplines work together closely. This includes sharing data and expertise across disciplines, as well as developing new tools and methods that are tailored to the specific needs of drug discovery.

## 5. Conclusion

The integration of bioinformatics in drug discovery has revolutionized the field of pharmaceutical research. The use of computational tools and techniques has enabled researchers to analyze vast amounts of data, identify potential drug targets, and design new drugs with greater precision and efficiency. The integration of bioinformatics has also facilitated the development of personalized medicine, where drugs can be tailored to individual patients based on their genetic makeup. However, there are still challenges that need to be addressed, such as the need for more accurate predictive models and the ethical considerations surrounding the use of patient data. Overall, the integration of bioinformatics in drug discovery holds great promise for improving human health and advancing our understanding of disease mechanisms.

**Author Contributions:** MM, & MA performed overall writing of the paper including research, literature review, and analysis of data; AA literature review, and analysis of data; AM Supervision and editing of the review paper. All authors read and approved the final manuscript.

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Acknowledgements:** The authors acknowledge the Deanship of Scientific Research at University of Bahri for the supportive cooperation.

**Conflicts of interest:** None declared.

## References

1. David, E.; Tramontin, T.; Zimmel, R. Pharmaceutical R&D: the road to positive returns. *Nat. Rev. Drug Discov.*, 2009, 8, 609-610. <https://doi.org/10.1038/nrd2948>
2. Drews, J.; Ryser, S. The role of innovation in drug development. *Nat. Biotechnol.*, 1997, 15, 1318-1319. <https://doi.org/10.1038/nbt1297-1318>
3. Gilbert, D. Bioinformatics software resources. *Brief. Bioinform.* 2004, 5, 300-304. <https://doi.org/10.1093/bib/5.3.300>
4. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009, 37, 1-13. <https://doi.org/10.1093/nar/gkn923>
5. Paananen J, Fortino V. An omics perspective on drug target discovery platforms. *Briefings in bioinformatics.* 2020 Nov;21(6):1937-53. <https://doi.org/10.1093/bib/bbz122>
6. Reichel A, Lienau P. Pharmacokinetics in drug discovery: an exposure-centred approach to optimising and predicting drug efficacy and safety. *New approaches to drug discovery.* 2016:235-60.
7. Malathi K, Ramaiah S. Bioinformatics approaches for new drug discovery: a review. *Biotechnology and Genetic Engineering Reviews.* 2018 Jul 3;34(2):243-60. <https://doi.org/10.1080/02648725.2018.1502984>
8. Gill SK, Christopher AF, Gupta V, Bansal P. Emerging role of bioinformatics tools and software in evolution of clinical research. *Perspectives in clinical research.* 2016 Jul;7(3):115. doi: 10.4103/2229-3485.184782
9. Xue Y, Lameijer EW, Ye K, Zhang K, Chang S, Wang X, Wu J, Gao G, Zhao F, Li J, Han C. Precision medicine: what challenges are we facing?. *Genomics, Proteomics & Bioinformatics.* 2016 Oct;14(5):253. doi: 10.1016/j.gpb.2016.10.001
10. Fernald GH, Capriotti E, Daneshjou R, Karczewski KJ, Altman RB. Bioinformatics challenges for personalized medicine. *Bioinformatics.* 2011 Jul 1;27(13):1741-8. <https://doi.org/10.1093/bioinformatics/btr295>
11. Schadt EE, Linderman MD, Sorenson J, Lee L, Nolan GP. Computational solutions to large-scale data management and analysis. *Nature reviews genetics.* 2010 Sep;11(9):647-57. <https://doi.org/10.1038/nrg2857>
12. Sinha S, Vohora D. Drug discovery and development: An overview. *Pharmaceutical medicine and translational clinical research.* 2018 Jan 1:19-32. <https://doi.org/10.1016/B978-0-12-802103-3.00002-X>
13. Kopec KK, Bozyczko-Coyne D, Williams M. Target identification and validation in drug discovery: the role of proteomics. *Biochemical Pharmacology.* 2005 Apr 15;69(8):1133-9. <https://doi.org/10.1016/j.bcp.2005.01.004>
14. Xia X. Bioinformatics and drug discovery. *Current topics in medicinal chemistry.* 2017 Jun 1;17(15):1709-26. DOI: 10.2174/1568026617666161116143440
15. Mamas M, Dunn WB, Neyses L, Goodacre R. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Archives of toxicology.* 2011 Jan;85:5-17. <https://doi.org/10.1007/s00204-010-0609-6>
16. Ramazi S, Zahiri J. Post-translational modifications in proteins: resources, tools and prediction methods. *Database.* 2021 Jan 1;2021:baab012. <https://doi.org/10.1093/database/baab012>
17. Golestan Hashemi FS, Razi Ismail M, Rafii Yusop M, Golestan Hashemi MS, Nadimi Shahraki MH, Rastegari H, Miah G, Aslani F. Intelligent mining of large-scale bio-data: Bioinformatics applications. *Biotechnology & Biotechnological Equipment.* 2018 Jan 2;32(1):10-29. <https://doi.org/10.1080/13102818.2017.1364977>
18. Malathi K, Ramaiah S. Bioinformatics approaches for new drug discovery: a review. *Biotechnology and Genetic Engineering Reviews.* 2018 Jul 3;34(2):243-60. <https://doi.org/10.1080/02648725.2018.1502984>

19. Wang S, Sim TB, Kim YS, Chang YT. Tools for target identification and validation. *Current opinion in chemical biology*. 2004 Aug 1;8(4):371-7. <https://doi.org/10.1016/j.cbpa.2004.06.001>
20. Rao VS, Srinivas K, Sujini GN, Kumar GN. Protein-protein interaction detection: methods and analysis. *International journal of proteomics*. 2014;2014. <http://dx.doi.org/10.1155/2014/147648>
21. Yang Y, Adelstein SJ, Kassis AI. Target discovery from data mining approaches. *Drug discovery today*. 2012 Feb 1;17:S16-23. <https://doi.org/10.1016/j.drudis.2011.12.006>
22. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, Karlsson A, Al-Lazikani B, Hersey A, Oprea TI, Overington JP. A comprehensive map of molecular drug targets. *Nature reviews Drug discovery*. 2017 Jan;16(1):19-34. DOI <https://doi.org/10.1038/nrd.2016.230>
23. Singh SS. Preclinical pharmacokinetics: an approach towards safer and efficacious drugs. *Current drug metabolism*. 2006 Feb 1;7(2):165-82. DOI: <https://doi.org/10.2174/138920006775541552>
24. Schenone M, Dančik V, Wagner BK, Clemons PA. Target identification and mechanism of action in chemical biology and drug discovery. *Nature chemical biology*. 2013 Apr;9(4):232-40. DOI <https://doi.org/10.1038/nchembio.1199>
25. Dhasmana A, Raza S, Jahan R, Lohani M, Arif JM. High-throughput virtual screening (HTVS) of natural compounds and exploration of their biomolecular mechanisms: An in silico approach. In *New look to phytomedicine 2019* Jan 1 (pp. 523-548). Academic Press. <https://doi.org/10.1016/B978-0-12-814619-4.00020-3>
26. Badrinarayan P, Narahari Sastry G. Virtual high throughput screening in new lead identification. *Combinatorial chemistry & high throughput screening*. 2011 Dec 1;14(10):840-60.
27. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*. 2004 Nov;3(11):935-49. DOI <https://doi.org/10.1038/nrd1549>
28. Rossi T, Braggio S. Quality by Design in lead optimization: a new strategy to address productivity in drug discovery. *Current Opinion in Pharmacology*. 2011 Oct 1;11(5):515-20. <https://doi.org/10.1016/j.coph.2011.05.004>
29. Ghosh A, Chakraborty M, Chandra A, Alam MP. Structure-activity relationship (SAR) and molecular dynamics study of withaferin-A fragment derivatives as potential therapeutic lead against main protease (M pro) of SARS-CoV-2. *Journal of molecular modeling*. 2021 Mar;27:1-7. DOI <https://doi.org/10.1007/s00894-021-04703-6>
30. Johnson DE, Wolfgang GH. Predicting human safety: screening and computational approaches. *Drug discovery today*. 2000 Oct 1;5(10):445-54. [https://doi.org/10.1016/S1359-6446\(00\)01559-2](https://doi.org/10.1016/S1359-6446(00)01559-2)
31. Kraljevic S, Stambrook PJ, Pavelic K. Accelerating drug discovery: Although the evolution of '-omics' methodologies is still in its infancy, both the pharmaceutical industry and patients could benefit from their implementation in the drug development process. *EMBO reports*. 2004 Sep;5(9):837-42. <https://doi.org/10.1038/sj.embor.7400236>
32. Hsiao Y, Su BH, Tseng YJ. Current development of integrated web servers for preclinical safety and pharmacokinetics assessments in drug development. *Briefings in Bioinformatics*. 2021 May;22(3):bbaa160. <https://doi.org/10.1093/bib/bbaa160>
33. Afshari CA, Hamadeh HK, Bushel PR. The evolution of bioinformatics in toxicology: advancing toxicogenomics. *Toxicological Sciences*. 2011 Mar 1;120(suppl\_1):S225-37. <https://doi.org/10.1093/toxsci/kfq373>
34. Varshney S, Bharti M, Sundram S, Malviya R, Fuloria NK. The Role of Bioinformatics Tools and Technologies in Clinical Trials. In *Bioinformatics Tools and Big Data Analytics for Patient Care 2022* Aug 31 (pp. 1-16). Chapman and Hall/CRC.
35. Loding WT, editor. *Bioinformatics and computational biology in drug discovery and development*. Cambridge University Press; 2016 Mar 17.
36. Peraman R, Sure SK, Dusthacker VA, Chilamakuru NB, Yiragamreddy PR, Pokuri C, Kutagulla VK, Chinni S. Insights on recent approaches in drug discovery strategies and untapped drug targets against drug resistance. *Future Journal of Pharmaceutical Sciences*. 2021 Dec;7:1-25. <https://doi.org/10.1186/s43094-021-00196-5>
37. Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. *Nature*. 2000 Jun 15;405(6788):827-36. DOI <https://doi.org/10.1038/35015701>
38. Pandey A, Mann M. Proteomics to study genes and genomes. *Nature*. 2000 Jun 15;405(6788):837-46. DOI <https://doi.org/10.1038/35015709>

39. Idle JR, Gonzalez FJ. *Metabolomics. Cell metabolism.* 2007 Nov 7;6(5):348-51. DOI 10.1016/j.cmet.2007.10.005
40. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research.* 1999 Jan 1;27(1):29-34. <https://doi.org/10.1093/nar/27.1.29>
41. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research.* 2017 Jan 4;45(D1):D353-61. <https://doi.org/10.1093/nar/gkw1092>
42. Goodsell DS, Zardecki C, Di Costanzo L, Duarte JM, Hudson BP, Persikova I, Segura J, Shao C, Voigt M, Westbrook JD, Young JY. RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Science.* 2020 Jan;29(1):52-65. <https://doi.org/10.1002/pro.3730>
43. Stanzione F, Giangreco I, Cole JC. Use of molecular docking computational tools in drug discovery. *Progress in Medicinal Chemistry.* 2021 Jan 1;60:273-343. <https://doi.org/10.1016/bs.pmch.2021.01.004>
44. Saikia S, Bordoloi M. Molecular docking: challenges, advances and its use in drug discovery perspective. *Current drug targets.* 2019 Apr 1;20(5):501-21. <https://doi.org/10.2174/1389450119666181022153016>
45. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nature genetics.* 2003 Mar;33(3):238-44. DOI <https://doi.org/10.1038/ng1107>
46. Cooper LA, Demicco EG, Saltz JH, Powell RT, Rao A, Lazar AJ. PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. *The Journal of pathology.* 2018 Apr;244(5):512-24. <https://doi.org/10.1002/path.5028>
47. Cline MS, Craft B, Swatloski T, Goldman M, Ma S, Haussler D, Zhu J. Exploring TCGA pan-cancer data at the UCSC cancer genomics browser. *Scientific reports.* 2013 Oct 2;3(1):2652. <https://doi.org/10.1038/srep02652>
48. Mittica G, Ghisoni E, Giannone G, Genta S, Aglietta M, Sapino A, Valabrega G. PARP inhibitors in ovarian cancer. Recent patents on anti-cancer drug discovery. 2018 Nov 1;13(4):392-410. <https://doi.org/10.2174/1574892813666180305165256>
49. Zheng F, Zhang Y, Chen S, Weng X, Rao Y, Fang H. Mechanism and current progress of Poly ADP-ribose polymerase (PARP) inhibitors in the treatment of ovarian cancer. *Biomedicine & Pharmacotherapy.* 2020 Mar 1;123:109661. <https://doi.org/10.1016/j.biopha.2019.109661>
50. Imamichi T. Action of anti-HIV drugs and resistance: reverse transcriptase inhibitors and protease inhibitors. *Current pharmaceutical design.* 2004 Dec 1;10(32):4039-53. DOI: <https://doi.org/10.2174/1381612043382440>
51. Ghosh AK, Chapsal BD, Weber IT, Mitsuya H. Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance. *Accounts of chemical research.* 2008 Jan 15;41(1):78-86. <https://doi.org/10.1021/ar7001232>
52. Zhan P, Pannecouque C, De Clercq E, Liu X. Anti-HIV drug discovery and development: current innovations and future trends: miniperspective. *Journal of medicinal chemistry.* 2016 Apr 14;59(7):2849-78. <https://doi.org/10.1021/acs.jmedchem.5b00497>
53. Merelli I, Pérez-Sánchez H, Gesing S, D'Agostino D. Managing, analysing, and integrating big data in medical bioinformatics: open problems and future perspectives. *BioMed research international.* 2014 Sep 1;2014. <https://doi.org/10.1155/2014/134023>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.