



BIOINFORMATICS TOOLS FOR PHARMACEUTICAL DRUG PRODUCT DEVELOPMENT

Edited By
Vivek Chavda
Krishnan Anand
Vasso Apostolopoulos

 Scrivener
Publishing

WILEY

Bioinformatics Tools
for Pharmaceutical Drug
Product Development

Scrivener Publishing

100 Cummings Center, Suite 541J
Beverly, MA 01915-6106

Publishers at Scrivener

Martin Scrivener (martin@scrivenerpublishing.com)
Phillip Carmical (pcarmical@scrivenerpublishing.com)

Bioinformatics Tools for Pharmaceutical Drug Product Development

Edited by

Vivek Chavda

*Department of Pharmaceutics and Pharmaceutical Technology,
L. M. College of Pharmacy, Ahmedabad, India*

Krishnan Anand

*Department of Chemical Pathology, School of Pathology, University of the Free
State, Bloemfontein, South Africa*

and

Vasso Apostolopoulos

*Institute for Health and Sport, Immunology and Translational Research Group,
Victoria University, Melbourne, Australia*



WILEY

This edition first published 2023 by John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA and Scrivener Publishing LLC, 100 Cummings Center, Suite 541J, Beverly, MA 01915, USA

© 2023 Scrivener Publishing LLC

For more information about Scrivener publications please visit www.scrivenerpublishing.com.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

Wiley Global Headquarters

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials, or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read.

Library of Congress Cataloging-in-Publication Data

ISBN 978-1-119-86511-7

Cover image: Pixabay.Com

Cover design by Russell Richardson

Set in size of 11pt and Minion Pro by Manila Typesetting Company, Makati, Philippines

Printed in the USA

10 9 8 7 6 5 4 3 2 1

Contents

Preface	xv
Part I: Bioinformatics Tools	1
1 Introduction to Bioinformatics, AI, and ML for Pharmaceuticals	3
<i>Vivek P. Chavda, Disha Vihol, Aayushi Patel, Elrashdy M. Redwan and Vladimir N. Uversky</i>	
1.1 Introduction	4
1.2 Bioinformatics	4
1.2.1 Limitations of Bioinformatics	8
1.2.2 Artificial Intelligence (AI)	8
1.3 Machine Learning (ML)	11
1.3.1 Applications of ML	12
1.3.2 Limitations of ML	14
1.4 Conclusion and Future Prospects	14
References	15
2 Artificial Intelligence and Machine Learning-Based New Drug Discovery Process with Molecular Modelling	19
<i>Isha Rani, Kavita Munjal, Rajeev K. Singla and Rupesh K. Gautam</i>	
2.1 Introduction	20
2.2 Artificial Intelligence in Drug Discovery	21
2.2.1 Training Dataset Used in Medicinal Chemistry	22
2.2.2 Availability and Quality of Initial Data	23
2.3 AI in Virtual Screening	24
2.4 AI for <i>De Novo</i> Design	25
2.5 AI for Synthesis Planning	26
2.6 AI in Quality Control and Quality Assurance	27
2.7 AI-Based Advanced Applications	28
2.7.1 Micro/Nanorobot Targeted Drug Delivery System	28

2.7.2	AI in Nanomedicine	29
2.7.3	Role of AI in Market Prediction	29
2.8	Discussion and Future Perspectives	30
2.9	Conclusion	31
	References	31
3	Role of Bioinformatics in Peptide-Based Drug Design and Its Serum Stability	37
	<i>Vivek Chavda, Prashant Kshirsagar and Nildip Chauhan</i>	
3.1	Introduction	37
3.2	Points to be Considered for Peptide-Based Delivery	38
3.3	Overview of Peptide-Based Drug Delivery System	40
3.4	Tools for Screening of Peptide Drug Candidate	41
3.5	Various Strategies to Increase Serum Stability of Peptide	42
3.5.1	Cyclization of Peptide	42
3.5.2	Incorporation of D Form of Amino Acid	44
3.5.3	Terminal Modification	44
3.5.4	Substitution of Amino Acid Which is Not Natural	46
3.5.5	Stapled Peptides	46
3.5.6	Synthesis of Stapled Peptides	47
3.6	Method/Tools for Serum Stability Evaluation	47
3.7	Conclusion	48
3.8	Future Prospects	49
	References	49
4	Data Analytics and Data Visualization for the Pharmaceutical Industry	55
	<i>Shalin Parikh, Ravi Patel, Dignesh Khunt, Vivek P. Chavda and Lalitkumar Vora</i>	
4.1	Introduction	56
4.2	Data Analytics	57
4.3	Data Visualization	58
4.4	Data Analytics and Data Visualization for Formulation Development	60
4.5	Data Analytics and Data Visualization for Drug Product Development	65
4.6	Data Analytics and Data Visualization for Drug Product Life Cycle Management	69
4.7	Conclusion and Future Prospects	71
	References	72

5	Mass Spectrometry, Protein Interaction and Amalgamation of Bioinformatics	77
	<i>Vivek Chavda, Kaustubh Dange and Madhav Joglekar</i>	
5.1	Introduction	77
5.2	Mass Spectrometry - Protein Interaction	79
5.2.1	The Prerequisites	80
5.2.2	Finding Affinity Partner (The Bait)	80
5.2.3	Antibody-Based Affinity Tags	80
5.2.4	Small Molecule Ligands	80
5.2.5	Fusion Protein-Based Affinity Tags	81
5.3	MS Analysis	81
5.4	Validating Specific Interactions	82
5.5	Mass Spectrometry – Qualitative and Quantitative Analysis	83
5.6	Challenges Associated with Mass Analysis	83
5.7	Relative vs. Absolute Quantification	85
5.8	Mass Spectrometry – Lipidomics and Metabolomics	86
5.9	Mass Spectrometry – Drug Discovery	87
5.10	Conclusion and Future Scope	88
5.11	Resources and Software	89
	Acknowledgement	89
	References	89
6	Applications of Bioinformatics Tools in Medicinal Biology and Biotechnology	95
	<i>Harshil Shah, Vivek Chavda and Moinuddin M. Soniwala</i>	
6.1	Introduction	96
6.2	Bioinformatics Tools	97
6.3	The Genetic Basis of Diseases	97
6.4	Proteomics	98
6.5	Transcriptomic	100
6.6	Cancer	101
6.7	Diagnosis	102
6.8	Drug Discovery and Testing	103
6.9	Molecular Medicines	105
6.10	Personalized (Precision) Medicines	106
6.11	Vaccine Development and Drug Discovery in Infectious Diseases and COVID-19 Pandemic	108
6.12	Prognosis of Ailments	109
6.13	Concluding Remarks and Future Prospects	110

Acknowledgement	111
References	111
7 Clinical Applications of “Omics” Technology as a Bioinformatic Tool	117
<i>Vivek Chavda, Rajashri Bezbaruah, Disha Valu, Sanjay Desai, Nildip Chauhan, Swati Marwadi, Gitima Deka and Zhiyong Ding</i>	
Abbreviations	118
7.1 Introduction	118
7.2 Execution Method	119
7.3 Overview of Omics Technology	121
7.4 Genomics	124
7.5 Nutrigenomics	127
7.6 Transcriptomics	128
7.7 Proteomics	129
7.8 Metabolomics	130
7.9 Lipomics or Lipidomics	133
7.10 Ayurgenomics	134
7.11 Pharmacogenomics	134
7.12 Toxicogenomic	135
7.13 Conclusion and Future Prospects	139
Acknowledgement	139
References	139

Part II: Bioinformatics Tools for Pharmaceutical Sector **147**

8 Bioinformatics and Cheminformatics Tools in Early Drug Discovery	149
<i>Palak K. Parikh, Jignasa K. Savjani, Anuradha K. Gajjar and Mahesh T. Chhabria</i>	
Abbreviations	150
8.1 Introduction	151
8.2 Informatics and Drug Discovery	152
8.3 Computational Methods in Drug Discovery	153
8.3.1 Homology Modeling	153
8.3.2 Docking Studies	155
8.3.3 Molecular Dynamics Simulations	158
8.3.4 <i>De Novo</i> Drug Design	159
8.3.5 Quantitative Structure Activity Relationships	160
8.3.6 Pharmacophore Modeling	161

8.3.7	Absorption, Distribution, Metabolism, Excretion and Toxicity Profiling	165
8.4	Conclusion	168
	References	169
9	Artificial Intelligence and Machine Learning-Based Formulation and Process Development for Drug Products	183
	<i>Vivek P. Chavda</i>	
9.1	Introduction	184
9.2	Current Scenario in Pharma Industry and Quality by Design (QbD)	185
9.3	AI- and ML-Based Formulation Development	187
9.4	AI- and ML-Based Process Development and Process Characterization	189
9.5	Concluding Remarks and Future Prospects	192
	References	193
10	Artificial Intelligence and Machine Learning-Based Manufacturing and Drug Product Marketing	197
	<i>Kajal Baviskar, Anjali Bedse, Shilpa Raut and Narayana Darapaneni</i>	
	Abbreviations	198
10.1	Introduction to Artificial Intelligence and Machine Learning	199
10.1.1	AI and ML in Pharmaceutical Manufacturing	200
10.1.2	AI and ML in Drug Product Marketing	201
10.2	Different Applications of AI and ML in the Pharma Field	202
10.2.1	Drug Discovery	202
10.2.2	Pharmaceutical Product Development	202
10.2.3	Clinical Trial Design	203
10.2.4	Manufacturing of Drugs	203
10.2.5	Quality Control and Quality Assurance	203
10.2.6	Product Management	203
10.2.7	Drug Prescription	204
10.2.8	Medical Diagnosis	204
10.2.9	Monitoring of Patients	204
10.2.10	Drug Synergism and Antagonism Prediction	204
10.2.11	Precision Medicine	205
10.3	AI and ML-Based Manufacturing	205
10.3.1	Continuous Manufacturing	205
10.3.2	Process Improvement and Fault Detection	209

10.3.3	Predictive Maintenance (PdM)	210
10.3.4	Quality Control and Yield	211
10.3.5	Troubleshooting	211
10.3.6	Supply Chain Management	212
10.3.7	Warehouse Management	213
10.3.8	Predicting Remaining Useful Life	214
10.3.9	Challenges	215
10.4	AI and ML-Based Drug Product Marketing	217
10.4.1	Product Launch	217
10.4.2	Real-Time Personalization and Consumer Behavior	218
10.4.3	Better Customer Relationships	219
10.4.4	Enhanced Marketing Measurement	220
10.4.5	Predictive Marketing Analytics	220
10.4.6	Price Dynamics	221
10.4.7	Market Segmentation	222
10.4.8	Challenges	223
10.5	Future Prospects and Way Forward	223
10.6	Conclusion	224
	References	225
11	Artificial Intelligence and Machine Learning Applications in Vaccine Development	233
	<i>Ali Sarmadi, Majid Hassanzadeganroudsari and M. Soltani</i>	
11.1	Introduction	234
11.2	Prioritizing Proteins as Vaccine Candidates	237
11.3	Predicting Binding Scores of Candidate Proteins	238
11.4	Predicting Potential Epitopes	243
11.5	Design of Multi-Epitope Vaccine	244
11.6	Tracking the RNA Mutations of a Virus	245
	Conclusion	248
	References	249
12	AI, ML and Other Bioinformatics Tools for Preclinical and Clinical Development of Drug Products	255
	<i>Avinash Khadela, Sagar Popat, Jinal Ajabiya, Disha Valu, Shrinivas Savale and Vivek P. Chavda</i>	
	Abbreviations	256
12.1	Introduction	257
12.2	AI and ML for Pandemic	258
12.3	Advanced Analytical Tools Used in Preclinical and Clinical Development	259

12.3.1	Spectroscopic Techniques	260
12.3.2	Chromatographic Techniques	263
12.3.3	Electrochemical Techniques	263
12.3.4	Electrophoretic Techniques	264
12.3.5	Hyphenated Techniques	264
12.4	AI, ML, and Other Bioinformatics Tools for Preclinical Development of Drug Products	265
12.4.1	Various Computational Tools Used in Pre-Clinical Drug Development	266
12.5	AI, ML, and Other Bioinformatics Tools for Clinical Development of Drug Products	268
12.5.1	Role of AI, ML, and Bioinformatics in Clinical Research	270
12.5.2	Role of AI and ML in Clinical Study Protocol Optimization	272
12.5.3	Role of AI and ML in the Management of Clinical Trial Participants	272
12.5.4	Role of AI and ML in Clinical Trial Data Collection and Management	272
12.6	Way Forward	275
12.7	Conclusion	276
	References	277

Part III: Bioinformatics Tools for Healthcare Sector **285**

13 Artificial Intelligence and Machine Learning in Healthcare Sector **287**

Vivek P. Chavda, Kaushika Patel, Sachin Patel and Vasso Apostolopoulos

	Abbreviations	288
13.1	Introduction	288
13.2	The Exponential Rise of AI/ML Solutions in Healthcare	289
13.3	AI/ML Healthcare Solutions for Doctors	291
13.4	AI/ML Solution for Patients	293
13.5	AI Solutions for Administrators	295
13.6	Factors Affecting the AI/ML Implementation in the Healthcare Sector	297
13.6.1	High Cost	297
13.6.2	Lack of Creativity	298
13.6.3	Errors Potentially Harming Patients	298
13.6.4	Privacy Issues	298

13.6.5	Increase in Unemployment	299
13.6.6	Lack of Ethics	299
13.6.7	Promotes a Less-Effort Culture Among Human Workers	299
13.7	AI/ML Based Healthcare Start-Ups	299
13.8	Opportunities and Risks for Future	304
13.8.1	Patient Mobility Monitoring	305
13.8.2	Clinical Trials for Drug Development	305
13.8.3	Quality of Electronic Health Records (EHR)	305
13.8.4	Robot-Assisted Surgery	305
13.9	Conclusion and Perspectives	306
	References	307
14	Role of Artificial Intelligence in Machine Learning for Diagnosis and Radiotherapy	315
	<i>Sanket Chintawar, Vaishnavi Gattani, Shivaneer Vyas and Shilpa Dawre</i>	
	Abbreviations	316
14.1	Introduction	317
14.2	Machine Learning Algorithm Models	318
14.2.1	Supervised Learning	318
14.2.2	Unsupervised Learning	319
14.2.3	Semi-Supervised Learning	319
14.2.4	Reinforcement Learning (RL)	320
14.3	Artificial Learning in Radiology	321
14.3.1	Types of Radiation Therapy	321
14.3.1.1	External Radiation Therapy	322
14.3.1.2	Internal Radiation Therapy	323
14.3.1.3	Systemic Radiation Therapy	323
14.3.2	Mechanism of Action	323
14.4	Application of Artificial Intelligence and Machine Learning in Radiotherapy	324
14.4.1	Delineation of the Target	324
14.4.2	Radiotherapy Delivery	325
14.4.3	Image Guided Radiotherapy	327
14.5	Implementation of Machine Learning Algorithms in Radiotherapy	328
14.5.1	Image Segmentation	328
14.5.2	Medical Image Registration	329
14.5.3	Computer-Aided Detection (CAD) and Diagnosis System	329
14.6	Deep Learning Models	331

14.6.1	Deep Neural Networks	331
14.6.2	Convolutional Neural Networks	332
14.7	Clinical Implementation of AI in Radiotherapy	332
14.8	Current Challenges and Future Directions	339
	References	339
15	Role of AI and ML in Epidemics and Pandemics	345
	<i>Rajashri Bezbaruah, Mainak Ghosh, Shuby Kumari, Lawandashisha Nongrang, Sheikh Rezzak Ali, Monali Lahiri, Hasmi Waris and Bibhuti Bhushan Kakoti</i>	
15.1	Introduction	346
15.2	History of Artificial Intelligence (AI) in Medicine	347
15.3	AI and ML Usage in Pandemic and Epidemic (COVID-19)	348
15.3.1	SARS-CoV-2 Detection and Therapy Using Machine Learning and Artificial Intelligence	349
15.3.2	SARS-Cov-2 Contact Tracing Using Machine Learning and Artificial Intelligence	350
15.3.3	SARS-CoV-2 Prediction and Forecasting Using Machine Learning and Artificial Intelligence	350
15.3.4	SARS-CoV-2 Medicines and Vaccine Using Machine Learning and Artificial Intelligence	350
15.4	Cost Optimization for Research and Development Using AI and ML	351
15.5	AI and ML in COVID 19 Vaccine Development	352
15.6	Efficacy of AI and ML in Vaccine Development	357
15.7	Artificial Intelligence and Machine Learning in Vaccine Development: Clinical Trials During an Epidemic and Pandemic	358
15.8	Clinical Trials During an Epidemic	360
15.8.1	Ebola Virus	360
15.8.2	SARS-CoV-2	361
15.9	Conclusion	361
	References	362
16	AI and ML for Development of Cell and Gene Therapy for Personalized Treatment	371
	<i>Susmit Mhatre, Somanshi Shukla, Vivek P. Chavda, Lakshmikanth Gandikota and Vandana Patravale</i>	
16.1	Fundamentals of Cell Therapy	372
16.1.1	Stem Cell Therapies	374
16.1.1.1	Mesenchymal Stem Cells (MSCs)	375
16.1.1.2	Hematopoietic Stem Cells (HSCs)	375

16.1.1.3	Mononuclear Cells (MNCs)	375
16.1.1.4	Endothelial Progenitor Cells (EPCs)	375
16.1.1.5	Neural Stem Cells (NSCs) or Neural Progenitor Cells (NPCs)	376
16.1.2	Adoptive Cell Therapy	376
16.1.2.1	Tumor-Infiltrating Lymphocyte (TIL) Therapy	376
16.1.2.2	Engineered T-Cell Receptor (TCR) Therapy	377
16.1.2.3	Chimeric Antigen Receptor (CAR) T Cell Therapy	377
16.1.2.4	Natural Killer (NK) Cell Therapy	377
16.2	Fundamentals of Gene Therapy	378
16.2.1	Identification	378
16.2.2	Treatment	379
16.3	Personalized Cell Therapy	381
16.4	Manufacturing of Cell and Gene-Based Therapies	382
16.5	Development of an Omics Profile	385
16.6	ML in Stem Cell Identification, Differentiation, and Characterization	387
16.7	Machine Learning in Gene Expression Imaging	389
16.8	AI in Gene Therapy Target and Potency Prediction	390
16.9	Conclusion and Future Prospective	391
	References	392
17	Future Prospects and Challenges in the Implementation of AI and ML in Pharma Sector	401
	<i>Prashant Pokhriyal, Vivek P. Chavda and Mili Pathak</i>	
17.1	Current Scenario	402
17.2	Way Forward	406
	References	407
Index		417

Preface

For a new drug to be developed and brought to market, approximately US\$1.8 billion and a minimum of 15 years in development are required. In most instances, only a few drugs make it to market because the process of creating a new drug can fail at different steps along the way, with most of them failing in the final stages of development. Some reasons for this can be attributed to the lack of extensive clinical data, unexpected toxicities and long-term side effects; as well as the highly competitive market, which puts a strain on the development of new drugs. A way to reduce some of the costs and increase the likelihood of success is to maximize the information gained via basic science and the design of better translational approaches and clinical trials. As such, bioinformatics approaches are becoming more essential in drug discovery and vaccine design, not only in academia, but also in the pharmaceutical industry. Bioinformatics involves the use of software tools and computer programming to understand biological data, particularly when the data is large and complex. The development of large data warehouses and algorithms to analyze large data, the identification of biomarkers and novel drug targets, computational biochemistry, genomics, drug discovery and design have all been at the forefront of translational drug discovery in recent years. Bioinformatics has revolutionized disease-based and drug-based approaches as well as improved knowledge of biological targets. It has ushered in a new era of research with the aim to accelerate the design and development of drug and vaccine targets, improve validation approaches as well as facilitate in identifying side effects and predict drug resistance. As such, this will aid in more successful drug candidates from discovery to clinical trials to the market and, most importantly, make it a more cost-effective process overall.

Since there has been much emphasis placed on developing bioinformatics tools for pharmaceutical drug development, this book is a timely and important addition to the evolving field. The 17 chapters are categorized into three sections. The first section presents the latest information on bioinformatics tools, artificial intelligence, machine learning, computational

methods, protein interactions, peptide-based drug design and omics technologies. The following two sections include bioinformatics tools for the pharmaceutical and healthcare sectors.

In this book, experts from around the world provide comprehensive overviews of the many important steps involved in—and the critical insights needed for—the successful development of therapeutic drug products with the help of bioinformatics, artificial intelligence and machine learning. The amount of work put into these 17 chapters required significant collaboration and input, and there are many people who are worthy of our thanks. We thank the support of over 20 thought leaders in AI/ML from across the globe who contributed to the chapters. Without their contributions, the book would not have been possible. We offer our sincere gratitude to the Department of Pharmaceutics and Pharmaceutical Technology, L.M. College of Pharmacy, Ahmedabad-Gujarat, India; the Department of Chemical Pathology, School of Pathology, University of the Free State, Bloemfontein campus, Free State, South Africa; and Victoria University, Institute for Health and Sport for their ongoing support in publishing this volume. We also thank Scrivener Publishing and their staff for their editorial support throughout the publication process.

The Editors
Vivek Chavda
K. Anand
Vasso Apostolopoulos
December 2022

Part I
BIOINFORMATICS TOOLS

Introduction to Bioinformatics, AI, and ML for Pharmaceuticals

Vivek P. Chavda^{1*}, Disha Vihol², Aayushi Patel³, Elrashdy M. Redwan^{4,5}
and Vladimir N. Uversky^{6†}

¹*Department of Pharmaceutics and Pharmaceutical Technology, L. M. College of Pharmacy, Ahmedabad, Gujarat, India*

²*Department of Phytopharmacy and Phytomedicine, School of Pharmacy, Gujarat Technological University, Ahmedabad, Gujarat, India*

³*Pharmacy Section, L. M. College of Pharmacy, Ahmedabad, Gujarat, India*

⁴*Department of Biological Sciences, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia*

⁵*Therapeutic and Protective Proteins Laboratory, Protein Research Department, Genetic Engineering and Biotechnology Research Institute, City of Scientific Research and Technological Applications, New Borg EL-Arab, Alexandria, Egypt*

⁶*Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA*

Abstract

Bioinformatics is a growing field that has emerged in recent years. The use of computational applications for protein sequence analysis in the early 1960s created the groundwork for bioinformatics. Alongside, developments in molecular biology techs evolved DNA analysis, leading to simpler manipulation of DNA, its sequencing, and computer science, suggesting the development of compatible and more powerful computers with innovative software for performing bioinformatics tasks. Biological Big Data collection when analyzed with bioinformatics tools leads to powerful predictive results with repeatability. Due to advancements in the merging of computer science and biology, even subdisciplines like synthetic biology, systems biology, and whole-cell modeling are emerging rapidly.

*Corresponding author: vivek7chavda@gmail.com

†Corresponding author: vuversky@usf.edu; ORCID: <https://orcid.org/0000-0002-4037-5857>

Keywords: Bioinformatics, artificial intelligence, machine learning, AI/ML, pharmaceuticals, drug product development

1.1 Introduction

In the context of Artificial Intelligence (AI), Machine Learning (ML), and Big Data, the healthcare industry explores the medication research process, evaluating how emerging technologies can enhance efficacy [1]. Artificial intelligence and machine learning are seen as the way of the future in a variety of fields and sectors, including pharmaceuticals. In a world, where a single authorized medicine costs millions of dollars and requires years of rigorous testing before being licensed, saving money and time is a priority.

Producing new pharmacological compounds to combat any disease is an expensive and time-consuming procedure, yet it goes unchecked. The most important aspect of drug design is to take advantage of the collected data and seek fresh and unique leads. Once the medication target has been determined, numerous multidisciplinary domains collaborate to develop enhanced pharmaceuticals using AI and ML technologies [2]. These technologies are utilized at every phase of the computer-aided drug discovery process, and combining them results in a proven track record of success in finding hit molecules. Furthermore, the fusion of AI and ML with high-dimension data enhanced the capabilities of computer-aided drug discovery and design [3]. Clinical trial output prediction using AI/ML integrated models might decrease the costs of the clinical trial, while simultaneously increasing their success rates. In this study, we will be covering the potential of AI and ML technologies in enabling computer-aided drug creation, along with challenges and opportunities for the pharmaceutical sector.

1.2 Bioinformatics

When biological data along with genetic information is analyzed using computer technology for calculating and obtaining mathematically and statistically approved results, is called Bioinformatics. The computational means are utilized for addressing data-intensive, large-scale biological challenges [4]. It includes the development and application of databases, algorithms, computational and statistical tools, and theory to tackle formal and practical issues emerging from biological data administration and analysis [5, 6].

Bioinformatics allows rapid molecular modeling of biological processes from the collected big data for obtaining meaningful conclusions through various stages such as compilation of the statistical information from biological data, creation of a computational model, the resolution of computational modeling issues, and the assessment of the resulting computational algorithm [7]. Genomics and proteomics are the branches of bioinformatics that aim at understanding the organizing principles encoded inside the nucleic acid and protein sequences, respectively. Image and signal processing enables usable conclusions to be extracted from vast volumes of raw data [5]. It helps in decoding genetics by facilitating text mining of biological literature, comparing, analyzing, and interpreting genetic, genomic, and proteomic data, assessing gene and protein expression, detecting mutations, sequencing and annotating genomes, and interpreting evolutionary aspects with disease pathways [8, 9]. Moreover, it helps to simulate and model DNA, RNA, proteins, and biomolecular interactions in structural biology. All of these can be achieved by correlating the biological data for understanding the effect of the diseased condition on the body's normal cellular functions [10]. Hence, bioinformatics has progressed now to the point, where analysis and interpretation of diverse forms of data is the most important challenge.

Omics technologies offer a chance to investigate changes in DNA, RNA, proteins, and other biological components across intraspecies and interspecies. The analysis of these components is interesting from a toxicity assessment view as they can alter in response to chemical or drug exposure in cells or tissues. Genomic research, which generates enormous amounts of data, is one area, where bioinformatics is very valuable. Bioinformatics aids in the interpretation of data, which may then be used to provide a diagnosis for a patient suffering from a rare ailment, track and monitor infectious organisms as they spread across a community, or determine the best therapy for a cancer patient [11]. Genomics sequences assemble and analyze the structure and function of genomes using recombinant DNA, DNA sequencing technologies, and bioinformatics. Various software used in bioinformatics includes accessibility of protein surface and secondary structure predictions using NetSurfP, prediction of beta-turn sites in protein sequences using NetTurnP, and AutoDock for Automated Docking Tool Suite [4].

Rapid advances in genomics and other molecular research tools, along with advances in information technology, have resulted in a massive volume of molecular biology knowledge during the last few decades [12]. Bioinformatics will continue to advance as a result of the integration of many technologies and techniques that bring together professionals from

Table 1.1 Various bioinformatics and AI-driven tools applied in the pharmacy department and industry.

Computational tools	Application	Reference
BLAST	The Basic Local Alignment Search Tool (BLAST) is used for searching local similarity regions between sequences and compares to the available database for calculating the statistical significance of matches. The matching infers functional and evolutionary relationships between sequences and identifies genetically related families.	[15]
ChEMBL	ChEMBL is designed manually to maintain a database of bioactive molecules. It correlates genomic data with chemical structure and bioactivity for the development of new drugs.	[16]
geWorkbench	The genomics Workbench (geWorkbench) comprises the tools for performing management, analysis, visualization, and annotation of biomedical data. It supports data for microarray gene expression, DNA and Protein Sequences, pathways, Molecular structure – prediction, Sequence Patterns, Gene Ontology, and Regulatory Networks	[11]
GROMACS	It is software for high-performance molecular dynamics and output analysis, especially for proteins and lipids.	[7]
IGV	Integrative Genomics Viewer (IGV) is a high-performance, user-friendly, interactive tool for the visualization and exploration of genomic data.	[17]
MODELLER	A protein three-dimensional structural homology or comparative modeling tool.	[18]
SwissDrugDesign	SwissDrugDesign provides a collection of tools required for Computer-Aided Drug Design (CADD).	[19]

(Continued)

Table 1.1 Various bioinformatics and AI-driven tools applied in the pharmacy department and industry. (*Continued*)

Computational tools	Application	Reference
UCSF Chimera	UCSF Chimera is an interactive tool for the visualization and analysis of molecular structures.	[20]
AlphaFold	It is an AI system developed to computationally predict protein structures with unprecedented accuracy and speed.	[21]
Cyclica	The ML tool address challenges faced across the drug discovery life cycle by correlating biophysics, medicinal chemistry, and systems biology.	[21]
DeepChem	It is a deep learning framework for drug discovery.	[18]
DeltaVina	Gives docking scores for protein-ligand binding affinity	[17]
Exscientia	The AI engineers precision medicines more rapidly and efficiently by accelerating pre-clinical drug discovery phases through monitoring and analysis of drug design and experiments.	[13]
Hit Dexter	It estimates how likely a small molecule is to trigger a positive response in biochemical and biological assays.	[4]
ORGANIC	Objective-Reinforced Generative Adversarial Network for Inverse-design Chemistry (ORGANIC) is a tool for creating molecules with desired properties.	[8]
Somatix	It is a real-time gesture detection technology that enables passive data collection of indoor and outdoor patients for enhancing medication adherence rates, data reliability, and cost-effectiveness.	[22]

other domains to build cutting-edge computational and informational tools tailored to the biomedical research business [4, 9]. Table 1.1. represents various bioinformatics and AI-driven tools, which can be applied in the pharmacy department and industry.

1.2.1 Limitations of Bioinformatics

In the case of drug discovery by applying *in silico* approaches, prediction of ligand affinity and inhibitory activity, where adequate training data are available is possible at certain levels but search for the mechanism-based inhibitors remains a highly challenging task. In structure-based approaches to various binding affinities, free energy calculations without extensive data collection at times leads to adequate results but high chances of false-positive rates remain a limiting factor. While chemical structure mining, the bioinformatics tools are efficiently able to produce a large number of metabolites but it is by no means possible to find ways of ranking metabolites accurately. Prediction of metabolic rates is generally not possible with bioinformatics tools. The *in vivo* predictions by applying bioinformatics for assessing biological activity and toxicological effects detect most toxicophores but the prediction of time-dependent inhibitors remains a difficult task [13].

Furthermore, it requires a sophisticated laboratory for the in-depth study and collection of biomolecules, and the establishment of such laboratories requires significant funds. The system operation is a complex mechanism so specially trained individuals are required for handling computer-based biological data. Uninterrupted electricity supply for biological investigations using computational applications is the basic requirement, abrupt interruption can lose huge data from the system memory. The system must also be virus-protected, which, if not controlled, can lead to the deletion of data and corruption of the programs [14].

1.2.2 Artificial Intelligence (AI)

Artificial intelligence (AI) is a branch of computer science that unlike natural intelligence is demonstrated by machines or computers and is defined as a system that analyzes its environment and performs a series of actions to achieve goals similar to a human-being process. The objectives of AI include gathering information, establishing rules on usage of the information, reasoning, data representation, organizing, planning, learning,

- Application of artificial intelligence and machine learning in radiotherapy, 324
 - delineation of the target, 324–325
 - image guided radiotherapy, 327
 - radiotherapy delivery, 325–326
- Applications of bioinformatics tools in,
 - cancer, 101
 - diagnosis, 102
 - drug discovery and testing, 103–104
 - molecular medicines, 105
 - personalized (precision) medicines, 106
- Artificial intelligence (AI), 402–406
- Artificial intelligence in drug discovery,
 - availability and quality of initial data, 23
 - training dataset used in medicinal chemistry, 22
- Artificial learning in radiology, 321
 - external radiation therapy, 322
 - internal radiation therapy, 323
 - systemic radiation therapy, 323
 - types of radiation therapy, 321
- Artificial neural networks, 379, 383, 388
- Augmented reality, 403
- Ayurgenomics, 134
- Bioinformatics, 152, 402
 - artificial intelligence (AI), 8–10
 - limitations of bioinformatics, 8
- CFD, 403
- Challenged associated with mass analysis, 83–85
- Cheminformatics, 152
- Chemometrics, 383
- Chimeric antigen receptor (CAR) T cell therapy, 377
- Chromatography modeling, 403
- Clinical implementation of AI in radiotherapy, 332–339
- Clinical trials during an epidemic, 360
- Clinical trials for drug development, 305
- Cost optimization for research and development using AI and ML, 351
- CRISPR, 379, 380
- Current scenario in pharma industry and quality by design (QbD), 185–187
- Data analytics, 57
- Data analytics and data visualization for drug product development, 65
- Data analytics and data visualization for drug product life cycle management, 69
- Data analytics and data visualization for formulation development, 60
- Data mining, 402–403
- Data visualization, 58
- De novo* ligand design, 152, 159–160
- Deep learning, 379, 390, 402, 404–405
- Deep learning models, 331
 - convolutional neural networks, 332
 - deep neural networks, 331
- Deep neural networks (DNNs), 291–293
- Design of multi-epitope vaccine, 244
- Different applications of AI and ML in the pharma field,
 - clinical trial design, 203
 - drug discovery, 202
 - drug prescription, 204
 - drug synergism and antagonism prediction, 204–205
 - manufacturing of drugs, 203
 - medical diagnosis, 204
 - pharmaceutical product development, 202
 - precision medicine, 205
 - product management, 203–204
 - quality control and quality assurance, 203

- Docking, 152, 155–158, 166
 - scoring function, 158
 - tools used for docking studies, 156–157
- Drug discovery, 151
- Drug discovery process, 151
- Efficacy of AI and ML in vaccine development, 357
- Electronic laboratory notebooks, 405
- Embryonic stem cell, 374
- Endothelial progenitor cells, 375
- Engineered T-cell receptor (TCR) therapy, 377
- Epigenomics, 386
- Eroom law, 402
- Errors in AI/ML potentially harming patients, 298
- Execution method, 119–121
- Flexible docking, 158
- Gene expression, 384, 389–391
- Gene set analysis, 384
- Genetic basis of diseases, 97
- Genome scale modeling, 403
- Genomics, 124–127
- Hematopoietic stem cell (HSC), 372, 375
- History of artificial intelligence (AI) in medicine, 347
- Homeobox (HOX) genes, 109
- Homology modeling, 153–155
 - steps involved in homology modeling, 153
 - tools used for homology modeling, 154–155
- Hybrid model, 402–403
- Implementation of machine learning algorithms in radiotherapy, 328
- computer-aided detection (CAD) and diagnosis system, 329–330
- image segmentation, 328–329
- medical image registration, 329
- In silico*, 402–403
- Internet of Things (IoT), 405
- Introduction to artificial intelligence and machine learning,
 - AI and ML in pharmaceutical manufacturing in drug product marketing, 201–202
 - AI and ML in pharmaceutical manufacturing in pharmaceutical manufacturing, 200–201
- Kinetic modeling, 403
- L. M. College of Pharmacy, 149
- Ligand-based drug design, 152
- Lipomics or lipidomics, 133
- Machine learning (ML), 404–406
 - applications of ML, 12–14
 - limitations of ML, 14
- Machine learning algorithm models, 318
 - reinforcement learning (RL), 320
 - semi-supervised learning, 319
 - supervised learning, 318–319
 - unsupervised learning, 319
- Mass spectrometry - drug discovery, 87–88
- Mass spectrometry - lipidomics and metabolomics, 86–87
- Mass spectrometry - protein interaction,
 - antibody-based affinity tags, 80
 - finding affinity partner (the bait), 80
 - fusion protein-based affinity tags, 81
 - small molecule ligands, 80–81
 - the prerequisites, 80
- Mass spectrometry - qualitative and quantitative analysis, 83
- Mechanism of action, 323–324
- Mechanistic models, 403

- Mesenchymal stem cells (MSC), 375, 388
- Metabolic flux analysis, 384
- Metabolomics, 130–133
- Method/tools for serum stability evaluation, 47–48
- miRNA, 109–110
- Molecular dynamics (MD)
 - simulations, 152, 158–159, 166
 - common workflow of MD simulation, 159
- Mononuclear cells, 375
- MS analysis, 81
- Multivariate data analysis, 383, 384

- Natural killer (NK) cell therapy, 377
- Natural language processing (NLP), 295, 405
- Neural stem cell, 376
- Nutrigenomics, 127–128

- Omics, 385, 386
- Opportunities and risks of AI/ML for the future, 304–306
- Overview of omics technology, 121–124
- Overview of peptide-based drug delivery system, 40–41

- Patient mobility monitoring, 305
- Personalized cell therapy, 381
- Perspectives of AI/ML, 306–307
- Pharmacogenomics, 134
- Pharmacophore modeling, 152, 161, 163–165
 - ligand-based pharmacophore modeling, 163
 - structure-based pharmacophore modeling, 163–164
 - tools used for pharmacophore modeling, 164–165
- Points to be considered for peptide-based delivery, 38–39

- Predicting binding scores of candidate proteins, 238
- Predicting potential epitopes, 243
- Predictive maintenance, 403
- Principal component analysis (PCA), 383, 389
- Prioritizing properties as vaccine candidates, 237
- Process analytical technologies (PAT), 384, 403
- Prognosis of ailments, 109–110
- Protein Data Bank (PDB), 152, 156, 158
- Proteomics, 98, 129

- Quality assurance, 403
- Quality control, 403
- Quality management system (QMS), 403
- Quality of electronic health records (EHR), 305
- Quantitative structure-activity relationship (QSAR), 152, 160–163, 166
 - general process of QSAR modeling, 160–161
 - statistical methods or machine learning techniques, 161
 - tools for QSAR modeling, 162–163

- Ramachandran plot, 155
- Random forest algorithms, 386
- Relative vs. absolute quantification, 85–86
- Reverse vaccinology, 108–109
- Rigid docking, 156, 158
- Robot-assisted surgery, 305
- Role of AI and ML in clinical study protocol optimization, 272
- Role of AI and ML in clinical trial data collection and management, 272
- Role of AI and ML in the management of clinical trial participants, 272
- RPA (robotic process automation), 296

- Semi-flexible docking, 156, 158
- Structure-based drug design, 152
- Support vector machine (SVM), 379, 386

- The exponential rise of AI/ML solutions in healthcare, 289–291
- Tools for screening for peptide drug candidate, 41–42
- Toxicogenomic, 135–136
- Tracking the RNA mutations of a virus, 245–247
- Transcriptomics, 100, 128
- Tumor-infiltrating lymphocyte (TIL) therapy, 376

- Vaccine development in COVID-19 pandemic, 108
- Validating specific interactions, 82–83
- Various strategies to increase serum stability of peptide,
 - cyclization of peptide, 42–44
 - incorporation of D form of amino acid, 44
 - stapled peptide, 46–47
 - substitution of amino acid which is not natural, 46
 - synthesis of stapled peptide, 47
 - terminal modification, 44–46
- Virtual screening, 152, 155, 156, 161, 165