IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Immunotherapy in the Age of Artificial Intelligence: Unlocking New Therapeutic Potentials

Ke Yixi[1], Liu Lantian1[1]

[1]Medical College, Southern University of Science and Technology, Shenzhen, Guangdong Province 518055 China

ABSTRACT Artificial Intelligence (AI) refers to the use of machines to mimic intelligent behavior, solving complex tasks with minimal human intervention. With the rapid development of technology, AI has become an important tool in the fields of biomedical research and clinical treatment. The application of AI in medicine has formed a new medical model that improves or enhances the level of medical diagnosis, treatment, and management, possessing tremendous potential to revolutionize and reshape medicine. Particularly in the field of immunotherapy, the application of AI is gradually changing our understanding and practice of disease treatment methods. This review aims to explore the application of AI in immunotherapy, especially in terms of target prediction, drug development, and clinical care prospects and challenges. In addition, we also summarized opinions on the challenges and opportunities related to the further application of AI in immunotherapy and its clinical practicality.

INDEX TERMS Artificial intelligence; Immunotherapy

1. INTRODUCTION

Artificial Intelligence involves creating algorithms and systems that can perform tasks that typically require human intelligence, which has evolved from basic computational models to complex algorithms capable of learning and adapting. Nowadays, the convergence of technological advancements, the proliferation of data, and the increasing acceptance and investment in AI have collectively contributed to its widespread application across various sectors and industries.

AI, particularly in the form of machine learning and deep learning, has become a pivotal tool in advancing biomedical research and therapy[1]. Machine learning (ML), a subset of AI, involves algorithms that learn from and make predictions or decisions based on data. Common machine learning algorithms include support vector machines (SVM), decision trees (DT), and Bayesian networks (BN). Deep learning (DL) is a more complex subset of machine learning in which a complex multi-layered neural network architecture automatically learns data by transforming input information into multiple levels of abstraction[2]. These technologies are highly relevant in the biomedical field, where vast and complex datasets are common. In the realm of biomedicine, AI technologies are used to analyze extensive datasets such as genomic sequences, proteomic profiles, and clinical data. These analyses often uncover patterns and relationships that are not immediately apparent to human researchers. This capability is particularly important in immunotherapy, where understanding the intricate relationship between cancer cells and the immune system is crucial.

Immunotherapy is a cutting-edge medical treatment that harnesses the innate power of the immune system by inducing, enhancing, or suppressing it to fight diseases, most notably cancer[3]. Unlike traditional treatments like chemotherapy, which directly target and kill cancer cells, immunotherapy works by either stimulating the immune system to attack cancer cells more effectively or by providing it with additional components, such as man-made immune system proteins. Immunotherapy has gained prominence for its potential to offer targeted, personalized treatment with potentially fewer side effects compared to conventional cancer therapies. Its applications extend beyond oncology, including treatments for autoimmune diseases, allergies, and even as preventive measures against certain infectious diseases. As research progresses, immunotherapy continues to evolve, showing promise as a pivotal approach in the future of medical treatments. However, challenges like variable patient responses

and drug resistance persist. Here, we delve into how AI is becoming an indispensable tool in overcoming these challenges and advancing the field of immunotherapy. AI, especially ML algorithms, can analyze vast amounts of biological data to identify patterns and relationships that might not be evident otherwise. This includes recognizing potential new targets which sets the stage for effective immunotherapy. Building upon the targets identified, the discovery of effective drugs will be accelerated. Through simulation and modeling, AI helps in understanding the efficacy and safety profiles as well as the synergistic combinations of drugs. Additionally, AI systems can continuously monitor patient responses to immunotherapy, enabling real-time adjustments to treatments, potential side effects prediction, and decision-making aids.

Overall, in this review, we summarized the multifaceted role of artificial intelligence and mainly focused on target prediction, drug development, and clinical care. The focus is on the synergy between AI and immunotherapy, highlighting the potential of AI to enhance the effectiveness and precision of cancer treatment. Immunotherapy in the age of artificial intelligence holds the promise of more personalized, efficient, and effective medical care, while it also presents challenges that need to be addressed to fully realize its potential.

2. Immunotherapy Targets Prediction

In the modern medical field, particularly in cancer treatment, the integration of artificial intelligence is opening up new possibilities for therapies. Immunotherapy, which utilizes a patient's immune system to treat cancer, relies heavily on accurately identifying effective targets to improve treatment outcomes. In this area, AI helps in scanning large molecular datasets to pinpoint potential therapeutic targets.

Through machine learning models, AI identifies genetic mutations and protein expression patterns related to cancer progression. These analyses help reveal biomarkers and potential therapeutic targets for cancer. For example, using ML models, particularly classification algorithms based on single-cell RNA sequencing (scRNA-Seq) data, a cancer stemness signature named Stem. Sig was developed to predict different cancer patients' responses to ICI treatment. Additionally, key immune-resistant genes in Stem. Sig were identified through statistical analysis of CRISPR datasets, highlighting them as potential targets for cancer immunotherapy[4]. Another study demonstrates the effectiveness of Extreme Gradient Boosting (XGBoost) algorithms in predicting immunotherapy targets by identifying a set of 88 predictive genes from the data of 853 non small cell lung cancer (NSCLC) patients, showing improved precision over traditional tumor mutational burden (TMB) methods[5]. A framework based on Deep Convolutional Neural Networks (CNN), named DeepLRHE, utilizes whole slide images (WSI) of lung cancer from The Cancer Genome Atlas (TCGA) for model training and validation to predict the efficacy of immunotherapy and identify biomarkers related to treatment response[6]. Meanwhile, mass spectrometry is a crucial method for analyzing the proteome, and the development of serum proteomic tests using machine learning algorithms through training datasets from cancer patients can be widely applied to identify biomarkers that predict the outcomes of immunotherapy in NSCLC patients[7].

AI also demonstrates immense potential in predicting tumor antigens. Tumor-specific antigens (TSAs) and tumor-associated antigens (TAAs) are key targets for immunotherapy. Deep learning algorithms enable the effective prediction and selection of tumor antigens with therapeutic potential. Neoantigens are a type of antigen that arise from tumor-specific mutations. They are essentially new epitopes formed when cancer cells undergo genetic mutations[8]. Recent advances in genomic and bio information technology have made it possible to profile immune responses to personalized neoantigens encoded by tumor-specific mutations. However, timely and effective identification of neoantigens remains a major obstacle to personalized neoantigen-based cancer immunotherapy[9]. Generally, the identification of potential neoantigens involves several steps, including the identification of somatic mutations, human leukocyte antigen (HLA) typing, peptide processing, and peptide-MHC binding prediction[10]. On the one hand, AI-based models can predict peptide-MHC binding specific to certain HLA class I alleles, an essential part of neoantigen identification. For example, the integration of ML algorithms such as MixMHCpred 2.0.1, NetMHCpan 4.0, and NetMHCcons 1.1 enhances the prediction of immunogenic peptides by considering additional factors like proteasome cleavage and peptide transport[11] because an important issue contributing to the relatively high false-positive rate is that current tools for predicting antigen presentation are mostly trained on in vitro binding affinity data, thus ignoring other factors[12]. On the other hand, complex computational algorithms based on ML have been developed to explore the nonlinear relationship between peptide sequences and the binding affinity with homologous MHC molecules, representing the high predictive accuracy of artificial neural networks[13].

Moreover, understanding how tumors evade immune system attacks is a challenge in immunotherapy. The data analysis capability of AI makes it a powerful tool for exploring and deciphering immune evasion mechanisms, thereby aiding in the discovery of new therapeutic targets[14]. Research shows that in high-grade serous ovarian cancer, the spatial variation of the intra-patient immune microenvironment shapes intraperitoneal clonal dispersion of tumors, where epithelial infiltration of CD8+ T cells scored using AI and DP negatively correlates with malignant diversity, reflecting the immunological pruning of tumor clones[15].

In brief, AI has revolutionized the field of immunotherapy target prediction with its precision, speed, and data-handling capabilities. This precision aids in the development of personalized treatment strategies, targeting cancer cells more effectively while minimizing damage to healthy cells.

3. Drug Development

3.1 Druggability Assessment

Druggability is a criterion for assessing whether a drug can bind to a corresponding protein target and induce biological activity[16], a crucial aspect of immunotherapeutic drug development. Statistically, the human genome encompasses a "druggable genome," containing approximately 30,000 genes, but only a small subset of these are suitable for drug design[17], making the evaluation of target druggability exceedingly important.

3.1.1 AI-Based Web Server for Druggability Assessment

Traditional analyses of druggability are resource-intensive and time-consuming[18]; web-based analyses can significantly expedite this process. PockDrug-Server, a novel web server, predicts the druggability of holo and apo proteins[19]. Borrel et al. optimized a linear discriminant analysis model by utilizing fpocket for the estimation of the Non-Redundant Ligandable Domain (NRDLD) dataset and by selecting pocket descriptors from a set of 52 geometric and physicochemical descriptors. Seven stable and effective models were chosen through cross-validation and with four different pocket estimation methods using an independent NRDLD test set. These models contributed to the development of PockDrug[20]. The PockDrug-Server, based on PockDrug, can predict pocket druggability. Users can input protein pocket coordinates and receive druggability evaluations upon submitting the protein's PDB file[19]. Researchers have used PockDrug-Server to predict the druggability of BCL-2 Interacting Killer (BIK) protein, identifying five protein pockets with three showing a druggability probability greater than 0.5, offering data support for monoclonal antibody drug development targeting BIK[21]. In summary, as protein pocket research advances, more web-based biological analysis applications are developed, enabling accurate assessment of anticancer targets' druggability and providing reliable druggable targets for immunotherapy in cancer treatment.

3.1.2 AI-Based Machine Learning for Druggability Assessment

AI based on machine learning aids in predicting protein structures to accurately identify protein pockets and target sites, as well as analyze drug molecule docking and target affinity. PockDrug relies heavily on precise 3D protein structures, the construction of which is time-intensive, thus greatly limiting its application. The widespread use of machine learning has made predictions of protein structures based on gene sequences increasingly viable, exemplified by the continually improving accuracy of AlphaFold. AlphaFold's innovative neural network architecture and training programs, informed by the evolutionary, physical, and geometric constraints of protein structures, have significantly enhanced the accuracy of structural predictions. In AlphaFold 2, released in 2021, the model achieved a median backbone accuracy of 0.96Å rmsd95 (Cα root-mean-square deviation at 95% residue coverage)[22]. Researchers like Yang Jie have used deep learning with co-evolutionary data to predict contacts and distances between protein residues, developing the trRosetta method. This method extends the machine learning-based prediction of protein residue distances to predict residue orientations and is optimized using Rosetta. In CASP13 FM target tests, trRosetta outperformed existing prediction methods, thus proving its enhanced accuracy in protein structure prediction[23]. The progress in these protein structure prediction methods undoubtedly aids druggability prediction models like PockDrug.

Molecular docking is a computational method for drug discovery, predicting ligand-target binding at the molecular level[24]. It first requires predicting the molecular orientation of ligands within the receptor, then uses scoring functions to estimate their complementarity[25]. For instance, aldose reductase, an enzyme linked to diabetes complications, was targeted by researchers using the ADAM&EVE program to search a 3D database, defining a 16 × 16 × 20 Å3 rectangular region encompassing the catalytic and G6P binding sites for ligand binding site conception. The CALGRID computer

program was employed to calculate van der Waals interactions, electrostatic interactions, and hydrogen bonding energy at each grid point within the region (with a spacing of 0.4Å). Of the molecules ultimately selected, 30% exhibited inhibitory effects on aldose reductase, demonstrating the powerful capability of the molecular docking method[26].

3.2 Vaccine Drug Development

3.2.1 Personalized Vaccine Development

Artificial Intelligence can be harnessed to predict proteomic peptides and identify which of them can be presented by MHC molecules, thereby facilitating the development of personalized vaccines. Current prediction methods for MHC II molecules are predominantly based on machine learning. These techniques are bifurcated into two main categories: allele-specific methods and pan-specific methods[27]. Allele-specific methods involve training models with a limited series of experimental data on peptide-MHC molecular binding affinities, which include approaches like simple motif sequences, hidden Markov models, and neural network models[28].

Due to the high polymorphism of MHC molecules, allele-specific methods exhibit lower accuracy. In contrast, pan-specific methods infer structural similarities or shared physicochemical binding determinants between HLA genes to predict the affinity of alleles not included in the training dataset[27]. Researchers have constructed a cross-allele predictive model that can utilize the amino acid sequences and structures of MHC II molecules to predict peptide-MHC II molecular interactions at three MHC II sites. They trained the model using a benchmark experimental dataset and measured its predictive performance with new data. The results demonstrated high accuracy, providing a theoretical framework for predicting peptide-MHC II interactions[27].

Artificial Intelligence can also distinguish between peptidic epitopes and non-epitopes using methods such as SVM, aiding in the assessment of a protein's immunogenicity. Gandharva and others collected experimentally identified immunomodulatory peptides to construct a positive set for antigen-presenting cell (APC) cell epitopes and gathered endogenous cyclic peptides from humans to form a negative set for APC cell epitopes. Combining the positive and negative sets into a complete training dataset, a computational model based on SVM was developed, which could classify new query peptides as either A-cell epitopes or non-epitopes[29], thereby aiding in the assessment of the antigenicity of target vaccine molecules.

3.2.2 Vaccine Molecule Screening

AI-based vaccine molecule screening can greatly enhance the speed of selection. In 2019, a newly developed influenza vaccine in the United States utilized AI for oligonucleotide prediction during its experimental phase, becoming the first vaccine identified using AI. This oligonucleotide sequence activates Toll-like receptors (TLR-9) in the human body, thereby promoting the production of cytokines and the expression of interferons, enhancing the vaccine's effectiveness. The construction of the oligonucleotides began with the generation of over 10^16 oligonucleotide candidates by a computer program, followed by TLR-9 affinity selection using software trained on various TLR-9s, ultimately yielding the most suitable oligonucleotide sequence[30]. Traditional manual experimental methods for screening these nucleotides would take years, but AI algorithms can complete the process within weeks, significantly increasing efficiency.

Furthermore, the involvement of AI also greatly improves the screening efficiency for protective immunogens. Protective immunogens are antigens that can produce memory cells in the human body. Dimitrov and colleagues utilized six different supervised machine learning algorithms (Partial Least Squares-Discriminant Analysis (PLS-DA), k-Nearest Neighbors (kNN), Random Forest (RF), Support Vector Machine (SVM), Random Subspace Method with kNN estimator (RSM), and XGBoost) to test 317 known bacterial immunogens and 317 bacterial non-immunogens. Validated through 10-fold internal cross-validation on the training set and an external test set, these models showed good performance, with the XGBoost model standing out for identifying immunogens and the RSM-1NN model being particularly effective for filtering non-immunogens[31]. These models aid in better and faster screening for protective immunogens, enhancing the efficiency of selection.

4. Clinical Applications

4.1 Healthcare

The widespread adoption of Electronic Health Records (EHR) has resulted in a voluminous amount of health information that exceeds the processing capability of clinical doctors[32]. While these large datasets pose challenges for human analysis, they also create opportunities for the large-scale application of artificial intelligence and machine learning in healthcare. Marina and colleagues compared three machine learning models—Random Forests (RFs), XGBoost, and

Classification and Regression Tree—with Logistic Regression (LR) for predicting the risk of Inborn Errors of Immunity (IEI). A cohort of 128 participants was divided into a training/validation set comprising 80% of the samples (102 individuals) and a test set containing the remaining 20% (26 individuals). The performance of the different models on the test dataset was evaluated using Receiver Operating Characteristic (ROC) curves, Area Under the ROC Curve (AUROC), Area Under the Precision-Recall Curve (AUPRC), and F1 scores. A 10-fold cross-validation framework was employed within the training/validation set to construct the models, which were subsequently assessed on the test set. All tested ML models exhibited high sensitivity and low specificity, indicating a significant advantage in screening for IEI[33]. The enhanced predictive capability provided by the ML models could serve as a resource for tracking IEI, offering reliable support for subsequent medical diagnosis, reducing referral delays, and thereby delivering better healthcare outcomes.

4.2 Disease Diagnosis

At the cellular level, deep convolutional neural networks have been utilized to identify different immune cells in frozen sections and their interactions, aiding in a better understanding of the immune system[34]. At the disease level, artificial intelligence can assist in the auxiliary diagnosis of immunological diseases by combining various clinical tests and examinations. Latent Tuberculosis Infection (LTBI) has become a major source of active tuberculosis (ATB). Due to the absence of clinical signs or symptoms, the application of machine learning in the discrimination and diagnosis of LTBI relies solely on biomarker data, not clinical symptoms or medical imaging. Patients infected with Mycobacterium tuberculosis initiate a specific immune response, resulting in variations in the types and quantities of mRNA in different cells. Machine learning methods based on microarrays offer a means to accurately differentiate between ATB and LTBI. In a study, differential gene expression analysis using microarrays and qPCR identified four biomarkers associated with ATB and LTBI differentiation: nuclear export mediator factor, asunder spermatogenesis regulator, DEAH (Asp-Glu-Ala-His) box polypeptide 29, and protein tyrosine phosphatase receptor type C. Additionally, researchers utilized four classifiers—Decision Tree, Random Forest, SVM, and Naïve Bayes—comparing the capacity of the four biomarkers to classify individuals through a 5-fold cross-validation method. It was found that models constructed with a combination of biomarkers exhibited higher accuracy, sensitivity, and specificity compared to models constructed using a single candidate gene[35]. Moreover, machine learning can be applied to analyze RNA-seq and RT-PCR data. In a study conducted in India, researchers employed RT-PCR and machine learning techniques to establish and validate a classification model. Upon cross-validation and on a separate validation dataset, the combination of just four genes was able to distinguish tuberculosis patients from healthy individuals with high accuracy, enhancing detection sensitivity while also reducing cost implications[36].

5. Discussion

Immunotherapy presents a promising approach compared to traditional therapies, particularly in targeting malignancies to overcome the limitations of conventional treatments. However, challenges such as large-scale screening of drug targets and immune evasion by tumor cells persist. The rapid advancement of artificial intelligence, especially machine learning and deep learning algorithms, enables the swift processing of vast immunological datasets, like RNA-seq and protein structure data. This assists in molecular screening and predicting unknown protein structures after supervised training with partial experimental data, thus improving the identification of drug targets, antigen screening, and the development of targeted immunotherapies or prophylactic drugs. AI's progress in image recognition also aids in identifying medical imaging data and analyzing changes in immune cells and active substances, enhancing clinical diagnostic efficiency and accuracy. Despite AI's advantages in sensitivity, accuracy, and high-throughput, issues such as patient privacy, dependence on pre-written algorithms, and lack of adaptability remain. Most current machine learning models in immunology focus on inherent errors and target analysis in conditions like asthma and atopic dermatitis. Although AI has advanced to be directly applicable in other fields like radiology with many software applications approved by the FDA as medical devices, applications in immunotherapy are still nascent. Future efforts should strengthen data privacy protection, enhance AI's learning capabilities, and leverage AI's computational power for immune drug simulation and expanded applications to accelerate clinical trials and minimize harm to subjects. The prospect of AI in immunotherapy is broad, and more advanced algorithms are expected to accelerate the evolution of immunotherapy.

6. Reference

1 Yang, Y., Zhao, Y., Liu, X. & Huang, J. Artificial intelligence for prediction of response to cancer immunotherapy. *Semin Cancer Biol* **87**, 137-147 (2022). https://doi.org/10.1016/j.semcancer.2022.11.008

2 Chan, H. P., Samala, R. K., Hadjiiski, L. M. & Zhou, C. Deep Learning in Medical Image Analysis. *Adv Exp Med Biol* **1213**, 3-21 (2020). https://doi.org/10.1007/978-3-030-33128-3_1

3 Yin, X. *et al.* Artificial intelligence-based prediction of clinical outcome in immunotherapy and targeted therapy of lung cancer. *Semin Cancer Biol* **86**, 146-159 (2022). https://doi.org/10.1016/j.semcancer.2022.08.002

4 Zhang, Z. *et al.* Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. *Genome Med* **14**, 45 (2022). https://doi.org/10.1186/s13073-022-01050-w

5 Liu, Z. *et al.* Predictive mutation signature of immunotherapy benefits in NSCLC based on machine learning algorithms. *Front Immunol* **13**, 989275 (2022). https://doi.org/10.3389/fimmu.2022.989275

6 Yang, Y. *et al.* Identification and Validation of Efficacy of Immunological Therapy for Lung Cancer From Histopathological Images Based on Deep Learning. *Front Genet* **12**, 642981 (2021). https://doi.org/10.3389/fgene.2021.642981

7 Park, Y. *et al.* Role of mass spectrometry-based serum proteomics signatures in predicting clinical outcomes and toxicity in patients with cancer treated with immunotherapy. *J Immunother Cancer* **10** (2022). https://doi.org/10.1136/jitc-2021-003566

8 Peng, M. *et al.* Neoantigen vaccine: an emerging tumor immunotherapy. *Mol Cancer* **18**, 128 (2019). https://doi.org/10.1186/s12943-019-1055-6

9 Chen, F. *et al.* Neoantigen identification strategies enable personalized immunotherapy in refractory solid tumors. *J Clin Invest* **129**, 2056-2070 (2019). https://doi.org/10.1172/jci99538

10 Richters, M. M. *et al.* Best practices for bioinformatic characterization of neoantigens for clinical utility. *Genome Med* **11**, 56 (2019). https://doi.org/10.1186/s13073-019-0666-2

11 Mei, S. *et al.* A comprehensive review and performance evaluation of bioinformatics tools for HLA class I peptide-binding prediction. *Brief Bioinform* **21**, 1119-1135 (2020). https://doi.org/10.1093/bib/bbz051

12 Xu, Y. *et al.* Technological advances in cancer immunity: from immunogenomics to single-cell analysis and artificial intelligence. *Signal Transduct Target Ther* **6**, 312 (2021). https://doi.org/10.1038/s41392-021-00729-7

13 Zhao, W. & Sher, X. Systematically benchmarking peptide-MHC binding predictors: From synthetic to naturally processed epitopes. *PLoS Comput Biol* **14**, e1006457 (2018). https://doi.org/10.1371/journal.pcbi.1006457

14 Sobhani, F. *et al.* Artificial intelligence and digital pathology: Opportunities and implications for immuno-oncology. *Biochim Biophys Acta Rev Cancer* **1875**, 188520 (2021). https://doi.org/10.1016/j.bbcan.2021.188520

15 Zhang, A. W. *et al.* Interfaces of Malignant and Immunologic Clonal Dynamics in Ovarian Cancer. *Cell* **173**, 1755-1769.e1722 (2018). https://doi.org/10.1016/j.cell.2018.03.073

16 Zhang, Y., Cao, H. & Liu, Z. Binding cavities and druggability of intrinsically disordered proteins. *Protein Sci* **24**, 688-705 (2015). https://doi.org/10.1002/pro.2641

17 Hopkins, A. L. & Groom, C. R. The druggable genome. *Nature Reviews Drug Discovery* **1**, 727-730 (2002). https://doi.org/10.1038/nrd892

18 Agache, I. *et al.* Advances and highlights in asthma in 2021. *Allergy* **76**, 3390-3407 (2021). https://doi.org/10.1111/all.15054

19 Hussein, H. A. *et al.* PockDrug-Server: a new web server for predicting pocket druggability on holo and apo proteins. *Nucleic Acids Res* **43**, W436-442 (2015). https://doi.org/10.1093/nar/gkv462

20 Borrel, A., Regad, L., Xhaard, H., Petitjean, M. & Camproux, A. C. PockDrug: A Model for Predicting Pocket Druggability That Overcomes Pocket Estimation Uncertainties. *J Chem Inf Model* **55**, 882-895 (2015). https://doi.org/10.1021/ci5006004

21 Liu, P. *et al.* Identification of BIK as an unfavorable prognostic marker and novel therapeutic target in microsatellite stable colorectal cancer harboring KRAS mutations. *Am J Cancer Res* **12**, 5300-5314 (2022).

22 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org/10.1038/s41586-021-03819-2

23 Yang, J. *et al.* Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci U S A* **117**, 1496-1503 (2020). https://doi.org/10.1073/pnas.1914677117

24 Pinzi, L. & Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int J Mol Sci* **20** (2019). https://doi.org/10.3390/ijms20184331

25 Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery* **3**, 935-949 (2004). https://doi.org/10.1038/nrd1549

26 Iwata, Y. *et al.* Discovery of Novel Aldose Reductase Inhibitors Using a Protein Structure-Based Approach: 3D-Database Search Followed by Design and Synthesis. *Journal of Medicinal Chemistry* **44**, 1718-1728 (2001). https://doi.org/10.1021/jm000483h

27 Degoot, A. M., Chirove, F. & Ndifon, W. Trans-Allelic Model for Prediction of Peptide:MHC-II Interactions. *Front Immunol* **9**, 1410 (2018). https://doi.org/10.3389/fimmu.2018.01410

28 Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol* **199**, 3360-3368 (2017). https://doi.org/10.4049/jimmunol.1700893

29 Nagpal, G., Chaudhary, K., Agrawal, P. & Raghava, G. P. S. Computer-aided prediction of antigen presenting cell modulators for designing peptide-based vaccine adjuvants. *Journal of Translational Medicine* **16**, 181 (2018). https://doi.org/10.1186/s12967-018-1560-1

30 Ahuja, A. S., Reddy, V. P. & Marques, O. Artificial intelligence and COVID-19: A multidisciplinary approach. *Integrative Medicine Research* **9**, 100434 (2020). https://doi.org/https://doi.org/10.1016/j.imr.2020.100434

31 Dimitrov, I., Zaharieva, N. & Doytchinova, I. Bacterial Immunogenicity Prediction by Machine Learning Methods. *Vaccines* **8** (2020).

32 Obermeyer, Z. & Lee, T. H. Lost in Thought - The Limits of the Human Mind and the Future of Medicine. *N Engl J Med* **377**, 1209-1211 (2017). https://doi.org/10.1056/NEJMp1705348

33 Takao, M. M. V. *et al.* Artificial Intelligence in Allergy and Immunology: Comparing Risk Prediction Models to Help Screen Inborn Errors of Immunity. *International Archives of Allergy and Immunology* **183**, 1226-1230 (2022). https://doi.org/10.1159/000526204

34 Liarski, V. M. *et al.* Quantifying in situ adaptive immune cell cognate interactions in humans. *Nat Immunol* **20**, 503-513 (2019). https://doi.org/10.1038/s41590-019-0315-3

35 Lee, S. W. *et al.* Gene expression profiling identifies candidate biomarkers for active and latent tuberculosis. *BMC Bioinformatics* **17 Suppl 1**, 3 (2016). https://doi.org/10.1186/s12859-015-0848-x

36 Maertzdorf, J. *et al.* Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol Med* **8**, 86-95 (2016). https://doi.org/10.15252/emmm.201505790