1-1-2024

Artificial Intelligence, Big Data, and Regulation of Immunity: Challenges and Opportunities

Bhagirath Singh

Anthony M Jevnikar

Eric Desjardins

Archivum Immunologiae et Therapiae Experimentalis
72, 6 (2024), DOI: 10.2478/aite-2024-0006

REVIEW

Artificial Intelligence, Big Data, and Regulation of Immunity: Challenges and Opportunities

Bhagirath Singh[1,2,3] · Anthony M. Jevnikar[1,4] · Eric Desjardins[3,5]

Abstract
The immune system is regulated by a complex set of genetic, molecular, and cellular interactions. Rapid advances in the study of immunity and its network of interactions have been boosted by a spectrum of "omics" technologies that have generated huge amounts of data that have reached the status of big data (BD). With recent developments in artificial intelligence (AI), theoretical and clinical breakthroughs could emerge. Analyses of large data sets with AI tools will allow the formulation of new testable hypotheses open new research avenues and provide innovative strategies for regulating immunity and treating immunological diseases. This includes diagnosis and identification of rare diseases, prevention and treatment of autoimmune diseases, allergic disorders, infectious diseases, metabolomic disorders, cancer, and organ transplantation. However, ethical and regulatory challenges remain as to how these studies will be used to advance our understanding of basic immunology and how immunity might be regulated in health and disease. This will be particularly important for entities in which the complexity of interactions occurring at the same time and multiple cellular pathways have eluded conventional approaches to understanding and treatment. The analyses of BD by AI are likely to be complicated as both positive and negative outcomes of regulating immunity may have important ethical ramifications that need to be considered. We suggest there is an immediate need to develop guidelines as to how the analyses of immunological BD by AI tools should guide immune-based interventions to treat various diseases, prevent infections, and maintain health within an ethical framework.

Keywords
Artificial intelligence • Big data • Ethics of immunotherapy • Omics • Precision medicine • Regulation of immunity

Received: 14 December 2023 / Accepted: 30 January 2024/

1. Introduction: Regulation of Immune Responses in Health and Disease and the Relevance of Artificial Intelligence

Studies over the past 80 years have revealed the complexity and the inherently self-regulating capacity of the immune system. While an important aspect of this work has been a better understanding of the regulation of the immune system in health and disease states (Maecker et al. 2012; Kaczorowski et al. 2017; Singh et al. 2019), the interactive and multiple pathways of immunity involving innate and adaptive immune responses have hampered insight and translational research. Immune responses are driven by a network of molecular interactions involving biological response modifiers such as hormones, cytokines, and chemokines, as well as signaling pathways between diverse cells of the immune system. Mechanistically, both antibody-mediated and cellular immunity involve multiple cell types, with or without specific antigen receptors. Immune cells both resident and circulating are thus found in almost all sites and tissues in the body, and can alter homeostatic host responses as well as those that respond to "danger."

In the context of health, the initiation of immune responses as well as the magnitude and tempo of response is impacted by host genetics, nutritional and environmental factors. Furthermore, immune responses are also deeply impacted by factors such as commensal microbiota and infectious agents (Nikoopour and Singh 2014; Singh et al. 2015). Commensal microorganisms in the gastrointestinal tract, and on mucosal surfaces and skin maintain tissue hemostasis. Pathogens can trigger both innate and adaptive immune responses and cause dysbiosis (Nikoopour and Singh 2014; Singh et al. 2015). The complex relationship between the immune system and microbes needs to be considered in any analysis of regulating immune responses in health and disease. With an evolving focus on "Precision Medicine," there is also a need to better understand immune regulation for immunotherapy at an individual level (Bluestone and Tang 2015).

The vast and accumulating amount of information about the immune system constitutes big data (BD). The value of

[1] *Department of Microbiology and Immunology, University of Western Ontario, London, ON, Canada*
[2] *Robarts Research Institute, University of Western Ontario, London, ON, Canada*
[3] *Rotman Institute of Philosophy, University of Western Ontario, London, ON, Canada*
[4] *Department of Medicine, University of Western Ontario, London, ON, Canada*
[5] *Department of Philosophy, University of Western Ontario, London, ON, Canada*

✉ *bsingh@uwo.ca*

BD was stated as "the process of collecting, organizing and analyzing large sets of data to discover patterns and other useful information" (Heymann and Rodier 2004). Existing BD repositories, such as ImmPort, spanning very diverse data sets from clinical studies and trials to microarrays and peptide epitopes, have become important sites for the storage of publicly funded immunology datasets (Bhattacharya et al. 2018; Deng et al. 2022). These repositories contain vast and diverse amounts of information on cellular, molecular, environmental, and microbial interactions that drive innate and adaptive responses, along with the response modifiers, cytokines, and signaling molecules and pathways that regulate immunity and cellular interactions between lymphoid and non-lymphoid cells in all organs and tissues of the body. In the near future, the proposed Human Immunome Project (2023) will generate even more data using "multi-omics" technology over the next few years. The measurements, analyses, and cataloging of immunity BD promises to be fascinating, and challenging, and could be important in generating new information and strategies for regulating immune responses in health and disease (Yu et al. 2019; Chu et al. 2021). While such information and infrastructures are valuable, priorities and processes remain undefined in many cases.

Such an overwhelming amount and diversity of information has prompted an interest in AI, an umbrella term referring to the broad area of computed decision-making. Machine learning (ML), is a subset of AI in which programming enables computer task-learning for making predictions with increasing precision over time. Additionally, "augmented intelligence" relates to the complementary enhancement of human effort via computer systems (Rider et al. 2020). The hope is that these (and future) AI tools will assist us in analyzing BD, thereby further expanding and speeding up the processes of research, diagnostics, and therapeutics in all fields of science and medicine (Bhattacharya et al. 2014, 2018; Deng et al. 2022).

However, immunity-related databases continue to grow, and to be of value this mass of heterogeneous information must be carefully curated. Indeed, the use and accessibility of data are often limited due to the complex structure and lack of uniformity currently in these repositories and platforms. This issue of access must be addressed to make this BD available to the biomedical and clinical research community and to render it analyzable by AI-based tools (Deng et al. 2022). This lack of uniformity will exacerbate as increasingly diverse studies will come to be, such as microbiome research that looks at how molecules produced by microbiota in various body sites interact with lymphoid cells and impact health and disease outcomes. The complex interaction between immune cells and microbiota (Hooper et al. 2012; Nikoopour and Singh 2014; Singh et al. 2015) makes AI/BD analysis more challenging as well as necessary and this has yet to be fully appreciated in regulating immunity (Park et al. 2023).

2. From Basic Research to Diagnosis, Treatment, and Monitoring of Health Outcomes

By merging AI-based analysis with "multi-omics" technology, the next generation(s) of investigators and trainees are presented with an opportunity to advance fundamental research in immunology (Schultze 2015, Wang et al. 2023) and to translate discoveries into clinically relevant outcomes (Bhattacharya et al. 2014, 2018; Chu et al. 2021; Deng et al. 2022). Some recent developments seem promising in this regard, such as repository platforms like ImmPort (Bhattacharya et al. 2018; Deng et al. 2022) that have great potential to advance this process (Fu et al. 2020). A relevant example of improved clinical outcomes from data analysis was in COVID-19 patients with diagnosis and prediction of severity (Liu et al. 2023). We can also look at the rapidly growing field of cancer immunotherapy for promising developments. Significant research efforts are devoted toward finding biomarkers that can assist in disease monitoring and predicting the efficacy of the interventions (Naqa et al. 2018; Jiang et al. 2022; Sharma et al. 2023). The results have also been used to modify treatments although the current success rate of these interventions has been limited (Naqa et al. 2018). The use of chimeric antigen receptor T cells (Sadelain et al. 2017; Hong et al. 2020) and epitope-based precision vaccines (Parvizpour et al. 2020; Yurina and Adianingsih 2022) to target cancer cells are new potentially powerful strategies in cancer immunotherapy. However, the multitude of mutations in tumors represents a huge challenge. AI approaches have yet to address this issue and there remains a significant knowledge gap as to how these therapies target cancer. Hopefully, success rates may improve if Precision Medicine can become a reality. However, AI analysis of BD cannot currently be applied to individual subjects to guide personalized therapy (Garattini et al. 2019).

These few examples represent the tip of the proverbial iceberg. The list of promising avenues is already long, and as AI-based tools develop, this list will likely expand. The area of transplantation could be of particular interest. We might see emerging knowledge that could prevent the need for transplant by modifying the underlying disease state, improving the donor and recipient matching process, identifying organs that can be used that would otherwise be discarded, preventing organ rejection through insights into immunity and drug selection, and even improving postoperative care. Yet, what seems abundantly clear, especially from the few examples in cancer immunotherapy, is that designing effective strategies for immunoregulation will require going beyond data collection and AI pattern recognition. The value of BD and AI-assisted analysis will only materialize if we adhere to one of the pillars of modern science, namely that research strategies must go beyond revealing interesting correlations and

move us toward a better understanding of the mechanisms that drive the relevant biological processes (Editorial 2015; Naqa et al. 2018; Sharma et al. 2023; Xie et al. 2023). The ongoing search for a mechanistic understanding of immunoregulation will be furthered by new bioinformatics tools that can be integrated into the design of our research strategies and immunotherapy design (Zhang et al. 2014; Jabbari and Rezaei 2019), but in all these developments, AI should only assist (not replace) experimental research.

As noted above, immunoregulation is inherently extremely complex and involves a host and environmental factors. The analysis of how we naturally respond to pathogens and treat or prevent infection with vaccines or immunotherapies must include a detailed understanding of the role of gut flora in immunity (Jabbari and Rezaei 2019). This is but one of the obvious ramifications that recent research has highlighted. We should not only encourage avenues that will lead to a mechanistic understanding of immunoregulation, but one that foster the discovery of a diversity of pathways and therapies. This means creating databases with sufficiently rich sources of information that include factors occurring at various scales.

3. Ethical Considerations

Amassing mountains of information is of course one of the outcomes of the newly developing approaches to data analysis. But BD also stands for data, which are of substantial variety, great volume, and accumulates with great velocity. However, while a "faster" rate of accumulation can hopefully translate into faster discoveries and clinical outputs, speed should not be mistaken for progress. Even if autonomous vehicles could travel at 300 km/h on our highways, it does not mean that we should push travel to these limits. Improving knowledge, health outcomes, and saving lives matter of course, but technological development should not be a race to publishing outcomes and producing innovations. Ultimately, if the goal of pursuing AI-based analyses of immunological BD is sustainable flourishing and improved well-being of humanity as a whole, then we must ensure that every step taken remains in line with this overarching good. This might mean intentionally introducing inertia to slow the process, like keeping research and development out of the hands of general intelligence (at least temporarily). Also needed is an enhancement of safety protocols already utilized where we have a more comprehensive assessment and equitable access to innovations to come.

We present some ethical considerations that we think are essential (albeit not exhaustive) as we move forward. We composed these with the assumption that regardless of the means, regulating immunity will always be a formidable problem with unintended consequences, some good and bad ones. We must thus proceed with caution. This implies keeping options open for solutions to address problems as they emerge, maintaining humans as final decision-makers, and moving beyond risk assessment. Running clinical trials is expected to reveal the side effects of interventions and when they are likely to occur. Having this information is crucial as it helps to prevent or mitigate some detrimental situations. Risk assessment allows to manage the situations by taking into consideration the probability of undesirable events. It utilizes classic statistical methods and substantial human inference, whereas AI and ML can automatically classify inputs without much human intervention. This has great power and requires caution, just as entering data into statistical packages without understanding limitations can generate incorrect outputs. Regardless of whether it is AI-assisted or not, the province of risk assessments remains in the realm of what we can know, predict, and hopefully do to reduce the chance and/or impact of these undesirable events. While essential, these measures are insufficient. Science and technology scholars often emphasize that innovations are not neutral, i.e., mere tools that we simply add, remove, and control. They change the world, including us, in substantial and often unpredictable and irreversible ways (Jasanoff 2016). The knowledge, techniques, and medical innovations that will come from using AI and immunological BD should be considered in the same light. Consequently, our ethical evaluations must embrace unpredictability and include broader social and environmental impacts. Table 1 outlines some of the major ethical challenges and possible approaches to address the regulation

Table 1. *Ethical considerations for the use of AI-based analysis of BD for regulating immunity and advancing immunotherapy*

Challenges	Approach
Autonomy	Informed consent of the participating subjects should be the cornerstone of data collection, storage, and usage. Enhanced risk of confidentiality breaches should be further emphasized and mitigated.
Public engagement at all stages	A diversity of experts and all potential end-users should be actively consulted in the development and implementation of these new analytical tools. Given the open-ended nature of BD, impact assessment should be an ongoing and adaptive process.
Equity and managing biases	The AI-based analysis should give adequate weight to all the relevant populations and socio-environmental determinants of immunoregulation to avoid biases and potential injustice.
Protect vulnerable populations	The new analytics should strive to create equitable opportunities by researching illnesses that especially affect vulnerable populations and develop treatments that are well-tailored to their means and values. At the same time, these efforts must not add a greater burden on vulnerable populations.
Reliability and trust	The AI-generated models and treatments are based on partially opaque processes and offer a limited understanding of the mechanisms involved. Unless high standards in research and care are maintained, this has the potential to hinder reliability and trust toward experts and institutions using AI-assisted analyses and decision-making.

AI, artificial intelligence; BD, big data.

of immune responses and advancing immunotherapy using AI-based analyses of immunological BD.

3.1. Autonomy

There is a growing culture of presumed consent when it comes to personal data collection by insurance companies, financial institutions, and "big tech" companies. Consumers continue to use companies and products that seek personal information, as they determine that the utility of the product is greater than the risk of sharing personal information. In healthcare studies, informed consent is one of the cornerstones of research ethics and this should remain so as much as possible for the collection, storage, and usage of data for AI-generated and assisted immunotherapies. However, this ideal of autonomy is clearly challenged by the open-ended nature of BD (Parker et al. 2019). One of the expected premises of BD is the possibility of discovering the unexpected. Moreover, it is common to reanalyze large datasets multiple times, and by multiple parties (Beaulieu and Leonelli 2021). It is not possible to predict what will come out of serial analyses. Will conclusions change with AI methodology changes and BD dataset merging? The consent regarding usage has important limitations. Once an AI has been trained with the data, it may not be possible to easily reset or undo the learning that has been done up to that point. Going back to our "safety over speed" principle, we contend that cutting corners and agreeing to lower autonomy standards to facilitate the use of AI for BD, despite the importance of the question cannot be justified. A minimal requirement should be that research participants and patients be made fully aware of the data being used, by whom, what database will be used to store the information, analyzed by which AI platform, for what questions, who will be allowed access, and whether participants can have their data removed in future. Measures should be put in place to minimize as much as possible a general consent to the unknown. For example, adding into the process a step where participants are informed of preliminary results before publication, with an option to remove their data. A related ethical concern is the protection of participants' and patients' confidentiality. When genetic information and multiple personal details are collected, (re)identification becomes a genuine concern, even for anonymized data (Parker et al. 2019). For example, what legislative protection exists to firewall insurance companies or financial institutions from data that relates to the risk of future health issues? This risk may be unavoidable, especially with increasingly powerful AI to come. Anyone giving permission to data usage should therefore be clearly made aware of this and governments need to openly debate what protections need to be built for their citizens. Finally, the burden and responsibility of maintaining minimal autonomy requirements should remain on the side of data users, and reuse should be stringently regulated to involve data providers, even if it means slowing down the process of discovery.

3.2. Public engagement

Not only experts but potential end-users should be actively consulted in the development and implementation of these new analytical tools. AI-generated models will expand the realm of possibilities by suggesting new interactions between the immune system and a host of factors (genes, hormones, drugs, microbiota, environmental agents, behaviors, etc.). These models will have to be tested rigorously and communicated to the public and caregivers as needed. This is important so that the new immunotherapies, be they based on vaccine design, gene therapy, or drugs, are economically, socially, and environmentally acceptable, especially for treatments that are likely to be scaled up. The open-ended nature of BD adds further complications to this engagement requirement because the network of "stakeholders" will be revealed and changed over time. Consultation should become an ongoing and dynamic process. Flexible and adaptive approaches should be favored instead of overly standardized and uniform ones.

3.3. Equity and managing biases

Ignoring the effect of diversity can result in blind spots, biases, and potential injustices (Beaulieu and Leonelli 2021). All relevant populations and determining factors should be given adequate weight in AI-based analyses intended for the development of models and immunotherapies. As noted current immunological BD is heavily weighted on "omics" research with the expectation the Human Immunome Project will further generate big datasets. These initiatives require advanced and expensive technology that can only be supported in resource-rich environments. Without vigilance, omics-centered approaches risk producing molecular models and treatments such as for precision medicine and drug efficacy predictions (Johnson et al. 2021; Li et al. 2021a, 2021b). Furthermore, even with the essential use of omics-rich data, humans are complex adaptive entities with immunity biology that cannot be accurately interpreted without considering social and environmental interactions. These will not appear in data sets that are curated to be primarily based on "omics" information. If we are to keep options open and in line with the various needs and demands of diverse populations, new analytical resources should be developed such that we gain greater knowledge of all the determinants of immunoregulation. This also has implications for the validation of models and treatments. Equivalence principles are commonly used in judging the efficacy and safety of treatments. AI should not replace experimental approaches here as well. Models

should be validated for human populations in their relevant socio-ecological contexts.

3.4. Protection of vulnerable populations

The new analytics should strive to create equitable opportunities by focusing research on illnesses that especially affect vulnerable populations and searching for treatments that are tailored to their means and values. At the same time, the follow-up of AI analysis outcomes must not add a greater burden on vulnerable populations. Arguably anyone with an immunological illness is vulnerable, but some populations have more limited means than others to deal with or compensate for their illnesses, while others can be more easily lured or coerced into research/treatment without proper informed consent. These include historically recognized vulnerable groups, such as children, prisoners, pregnant women, fetuses, elderly, mentally disabled persons, and economically or educationally disadvantaged individuals (Park and Grayson 2008). In the context of BD, vulnerable populations also include any group that could be subject to the misuse of data. Bioinformatics data analytics tools can be used to identify vulnerable populations (Cullen and Garcia 2021). This should be developed in conjunction with the other ethical considerations discussed in this section.

3.5. Reliability and trust

In general, initial attempts to use AI-generated models and treatment recommendations will be based on a limited mechanistic understanding of the regulation of immunity. Furthermore, due to its inherent complexity and lack of tractability, AI-assisted decision-making has a certain degree of opacity (black-box problem) (Wadden 2021). These features raise a wide range of issues that fall under the umbrella of reliability and trust (Lockey et al. 2021). Can we justify starting clinical trials based on models and hypotheses that have little or no good causal basis? How reliant should health professionals be on AI while searching for diagnostics and treatments? For that matter, how will health professionals deal with non-medically trained patients armed with AI analyses pertaining to their illness? How will the trust relationships between healthcare professionals, caregivers, and patients be affected? Answering these questions will depend in part on how much success AI-assisted research and therapies will have. However, "success" should be operationalized carefully if knowledge and trust are to be part of the equation. High standards in research and care are not limited to finding testable models and effective cures. Can AI-assisted research yield explanatory models? Can AI decision-making achieve an operational equivalent of judgments considerate of patients' contexts, emotions, values, and general well-being? Until it does accomplish these tasks, AI should be used with augmentation and not automation in mind. Currently, there are clearly more questions than answers in this context.

4. Strategic Recommendations for the Application of AI in BD Analysis for Immune Regulation

In addition to the ethical considerations listed in the previous section, the use of new analytical tools in research and immunotherapy calls for the development of strategic guidelines. The following points will be relevant in this context and Table 2 summarizes the strategies and potential recommendations with the regulation of immunity and enhancing immunotherapy using AI/BD analyses:

4.1. Streamline BD repositories for various immunological platforms

Develop guidelines for current and future immunological data repositories that will streamline the collection, storage, and retrieval of data from omics and other platforms in a transparent manner. This will help with the standardization of analysis and use of immunological datasets by AI/BD-based tools. This will allow for better workflow, it can enhance transparency, and in return improve global acceptance of recommendations for clinical interventions in regulating immune responses and for data collection, storage, and accessibility to all users.

Table 2. *Strategic recommendations for the use of AI-based analysis of BD for regulating immunity and advancing immunotherapy*

Strategy	Recommendation to respond
Streamline BD repositories	Develop guidelines that will streamline current and future immunological data repositories to help workflow and transparency.
Establish decision-making strategies for the use of AI-based big omics data analysis	Develop guidelines for decision-making as to how the analysis of "omics" data by AI/BD will use biomarkers for immunotherapy and regulation of immunity.
Advanced patient-centric "Precision Medicine" approaches	Develop approaches that are primarily patient--centric and do not depend uniquely on aggregated data from BD sets for AI analysis.
Develop strategies to address unforeseen adverse effects	Develop guidelines as to what steps and alternatives should be considered if AI-guided analysis predicts undocumented side effects or fails to predict side effects.
Incorporate the role of microbiota-immune cell interaction in immunoregulation	Develop complementary tracks to analyze BD involving interaction between innate and adaptive immune cells with microbiota.
Develop strategies to use AI/BD analysis to address knowledge gaps in the regulation of immunity and advancing immunotherapy	Develop strategies to use the analysis of omics and other immunological datasets by AI/BD tools to understand molecular and cellular mechanisms to address knowledge gaps for regulating immunity and immunotherapy in health and disease states.

AI, artificial intelligence; BD, big data.

4.2. Establish decision-making strategies for using AI-based BD analysis to regulate immunity

Develop guidelines for decision-making as to how AI analysis of BD in health *vs*. disease states can be used for regulating immunity. This is important because to define what is "immunity" and what immune parameters are relevant to maintaining good health is still an open question. Clinical biomarkers such as the level of antibodies, cytokines, chemokines, and other immune cell-derived biological response modifier compounds are highly variable in different subjects. The diversity of immune parameters and interrelated cell signaling pathways create significant challenges to apply AI approaches to analyze BD data for clinical decision-making. Moreover, various cell types in innate and adaptive immunity produce similar compounds that regulate immunity. This presents challenges to the use of aggregated BD data for AI-based decision-making to regulate immunity as it should be guided by biomarkers and other acceptable health parameters. To minimize the risk of false positives and negatives, it is important that there are reliable parameters to define healthy and diseased immune states/responses. However, this is very difficult because the biomarkers used to characterize immunological states are so variable from person to person that the two categories (healthy *vs*. diseased) overlap significantly. The implication of this overlap means that there is a high risk of error because it is typically dealing with aggregated data, the use of AI might be more at risk of diagnosing a condition as a disease while it may correspond to a healthy immune response to a threat (false positive) and the converse is also true. The AI might fail to recognize a situation as a malfunctioning immunological response (false negative) because of a lack of statistical difference with the aggregated healthy state in the population. Therefore, efforts should be made to reduce the risk of these errors. If AI-based models and diagnostics come from the analysis of aggregated data, then the risk of error for personal medicine is even greater.

4.3. Advanced patient-centric "Precision Medicine" approaches for using BD

Create a knowledge-based framework for immunotherapy to allow precision medicine approaches that are primarily patient-centric and do not depend on aggregated data resulting from AI/BD analysis of large data sets. This should include a possibility to override AI-based recommendations and more fine-grained adjustments of the timing of immunotherapies for various patients according to their health status and context. As discussed, the disclosure of patient information will be necessary to adopt the application of Precision Medicine approaches to design appropriate therapy or treatment in each case. This will impact decision-making as to what therapy options are available, accessible, or affordable to an individual or as part of an identifiable vulnerable group.

4.4. Develop strategies to address (unforeseen) adverse effects in regulating immunity using AI-based BD analysis

Ideally, AI-assisted analysis and diagnostics will inform users of the potential adverse effects of various interventions. If AI predicts potential side effects, then this information should already be in the database, which implies that there has been research already, and there should be guidelines in place regarding their severity and safety. If the side effects predicted are not documented and no guidelines are available, then evidence-based protocols must be developed. This should include developing safe alternatives that will consider the unexpected side effects and initiate research aimed at elucidating the validity of the inference and (if applicable) the mechanisms involved. If adverse effects manifest themselves but have not been predicted by AI, then it means that the models have not been tested rigorously enough and that researchers must return to the drawing board before the treatments continue to be administered in clinical contexts.

4.5. Incorporate the role of microbiota in AI-based BD analysis in regulating immunity

As discussed above, commensal microbiota and infectious agents such as bacteria and viruses are directly involved in regulating immunity in health and disease states (Nikoopour and Singh 2014; Singh et al. 2015; Berg et al. 2020; Zheng et al. 2020; Cao et al. 2022). It is essential that there are complementary tracks to analyze BD involving interaction between the innate and adaptive immune system cells with the host microbiota. This also includes the development of metabolic, gastrointestinal, autoimmune, infectious, and chronic diseases including cancer.

4.6. Develop strategies to address knowledge gaps at the mechanistic level for using AI-based BD data analysis

Develop strategies to consider how the AI-based BD analysis will help to advance a basic mechanistic understanding of immune regulation and immunotherapy at the molecular and cellular level in normal healthy and diseased states. This should address knowledge gaps in regulating immunity. It is not sufficient to just collect data and analyze it by using AI tools but its application to advance the conceptual framework of immune regulation and regulate immunity in health and disease. This will help in advancing the mechanistic understanding of immunity using AI-driven BD data analysis to address knowledge gaps (some theoretical, others

more practical) at the molecular and cellular level. As mentioned above the microbiome-host immune cell interactions driven immunoregulation seem like an important example of a knowledge gap that is important in AI-mediated decision using the BD analysis.

5. Conclusion

The field of immunology has always been driven by conceptual advances that help the discovery research to address substantial and long-standing knowledge gaps. This is evident from the lack of a clear understanding of why cancer immunotherapy has relatively low efficacy for example. It can be extended to other immune system-related knowledge gaps such as why vaccines work or do not work, why the immune system fails to control tumor development, why some transplant patients achieve operational tolerance and do not require drugs while others do not, and why autoimmune diseases or allergies progress despite various immune-regulatory mechanisms that should be operating in vivo. In most cases, this can be attributed to the lack of mechanistic understanding and AI-based analyses may help to design new strategies to modulate immune responses and develop rationale for immunotherapy. However, the collection of more big datasets and chasing AI algorithms is not going to make the field of immune regulation and immunotherapy provide answers as to how our immune system works at a mechanistic or individual level, or what immunotherapies should be considered. AI will not deliver more than it can. While AI may allow us to develop better vaccines, it is unlikely to provide answers to what drives vaccine hesitancy. In summary, AI using BD has the potential to suggest new pathways in regulating immunity and advancing immunotherapy through a better understanding of the basic immunological mechanisms, and to do so in an ethical manner. There is an urgent need to develop and refine how to apply AI/BD-driven ML to advance the field of immunology and immunotherapy in diverse diseases. Despite the pressure to move "faster," this should be guided by an ethical and rationale framework that will guide the search for new knowledge and develop strategies to maintain health and prevent and cure diseases.

Acknowledgements

Our published studies cited in this article were supported by grants from the Canadian Institutes of Health Research (CIHR) and the International Human Microbiome Standards grant (HEALTH-F4-2010-261376) from the European Community's Seventh Framework Program.

Author Contributions

All authors have made direct intellectual and conceptual contributions to the article and have approved it for publication.

Conflict of Interest

The authors have no financial conflict of interest to report.

References

Beaulieu A, Leonelli S (2021). Data and Society: A Critical Introduction. SAGE Publications Sage CA: Los Angeles, CA.

Berg G, Rybakova D, Fischer D et al (2020) Microbiome definition re-visited: Old concepts and new challenges. Microbiome 8:103. https://doi.org/10.1186/s40168-020-00875-0

Bhattacharya S, Andorf S, Gomes L et al (2014) ImmPort: Disseminating data to the public for the future of immunology. Immunol Res 58:234–239. https://doi.org/10.1007/s12026-014-8516-1

Bhattacharya S, Dunn P, Thomas CG et al (2018) ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Sci Data 5:180015. https://doi.org/10.1038/sdata.2018.15

Bluestone JA, Tang Q (2015) Immunotherapy: Making the case for precision medicine. Sci Transl Med 7:280ed3. https://doi.org/10.1126/scitranslmed.aaa9846

Cao Z, Sugimura N, Burgermeister E et al (2022) The gut virome: A new microbiome component in health and disease. EBioMedicine 81:104113. https://doi.org/10.1016/j.ebiom.2022.104113

Chu X, Zhang B, Koeken VACM et al (2021) Multi-omics approaches in immunological research. Front Immunol 12:668045. https://doi.org/10.3389/fimmu.2021.668045

Cullen T, Garcia JE (2021) Data mining, data analytics, and bioinformatics. In: Okpaku SO (ed) Innovations in Global Mental Health. Springer, Cham, pp 455–488. https://doi.org/10.1007/978-3-030-57296-9_141

Deng N, Wu C, Yaseen A et al (2022) ImmuneData: An integrated data discovery system for immunology data repositories. Database 2022:baac003. https://doi.org/10.1093/database/baac003

Editorial (2015) Big data meets mechanism. Nat Med 21:673. https://doi.org/10.1038/nm.3903

Fu J, Li K, Zhang W et al (2020) Large-scale public data reuse to model immunotherapy response and resistance. Genome Med 12:21. https://doi.org/10.1186/s13073-020-0721-z

Garattini C, Raffle J, Aisyah DN et al (2019) Big data analytics, infectious diseases and associated ethical impacts. Philos Technol 32:69–85. https://doi.org/10.1007/s13347-017-0278-y

Heymann DL, Rodier GR (2004) SARS: A global response to an international threat. Brown J World Aff 10:185–197.

Hong M, Clubb JD, Chen YY (2020) Engineering CAR-T cells for next-generation cancer therapy. Cancer Cell 38:473–488. https://doi.org/10.1016/j.ccell.2020.07.005

Hooper LV, Littman DR, Macpherson AJ (2012) Interactions between the microbiota and the immune system. Science 336: 1268–1273. https://doi.org/10.1126/science.1223490

Human Immunome Project (2023) A new model for global health: Harnessing the power of the immune system to improve health for all. Available at: https://www.humanimmunomeproject.org/ and was accessed on 27 October 2023.

Jabbari P, Rezaei N (2019) Artificial intelligence and immunotherapy. Expert Rev Clin Immunol 15:689–691. https://doi.org/10.1080/1744666X.2019.1623670

Jasanoff S (2016) The Ethics of Invention: Technology and the Human Future. WW Norton & Company Inc, New York, NY.

Jiang P, Sinha S, Aldape K et al (2022) Big data in basic and translational cancer research. Nat Rev Cancer 22:625–639. https://doi.org/10.1038/s41568-022-00502-0

Johnson KB, Wei WQ, Weeraratne D et al (2021) Precision medicine, AI, and the future of personalized health care. Clin Transl Sci 14:86–93. https://doi.org/10.1111/cts.12884

Kaczorowski KJ, Shekhar K, Nkulikiyimfura D et al (2017) Continuous immunotypes describe human immune variation and predict diverse responses. Proc Natl Acad Sci USA 114:E6097–E6106. https://doi.org/10.1073/pnas.1705065114

Li J, Wu J, Zhao Z et al (2021a) Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: A narrative review. J Thorac Dis 13:7021–7033. https://doi.org/10.21037/jtd-21-864

Li Y, Ma L, Wu D et al (2021b) Advances in bulk and single-cell multi-omics approaches for systems biology and precision medicine. Brief Bioinform 22:bbab024. https://doi.org/10.1093/bib/bbab024

Liu X, Hasan MR, Ahmed KA et al (2023) Machine learning to analyse omic-data for COVID-19 diagnosis and prognosis. BMC Bioinformatics 24:7. https://doi.org/10.1186/s12859-022-05127-6

Lockey S, Gillespie N, Holm D et al (2021) A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions. In: Proceedings of the 54th Hawaii International Conference on System Sciences, 5463–5472. https://hdl.handle.net/10125/71284

Maecker HT, McCoy JP, Nussenblatt R (2012) Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol 12:191–200. https://doi.org/10.1038/nri3158

Naqa IE, Kosorok MR, Jin J et al (2018) Prospects and challenges for clinical decision support in the era of big data. JCO Clin Cancer Inform 2:CCI.18.00002. https://doi.org/10.1200/CCI.18.00002

Nikoopour E, Singh B (2014) Reciprocity in microbiome and immune system interactions and its implications in disease and health. Inflamm Allergy Drug Targets 13:94–104. https://doi.org/10.2174/1871528113666140330201056

Park JS, Gazzaniga FS, Kasper DL et al (2023) Microbiota-dependent regulation of costimulatory and coinhibitory pathways via innate immune sensors and implications for immunotherapy. Exp Mol Med 55:1913–1921. https://doi.org/10.1038/s12276-023-01075-0

Park SS, Grayson MH (2008) Clinical research: Protection of the “vulnerable”? J Allergy Clin Immunol 121:1103–1107.

Parker DM, Pine SG, Ernst ZW (2019) Privacy and informed consent for research in the age of big data. Penn State Law Rev 123:703–733.

Parvizpour S, Pourseif MM, Razmara J et al (2020) Epitope-based vaccine design: A comprehensive overview of bioinformatics approaches. Drug Discov Today 25:1034–1042. https://doi.org/10.1016/j.drudis.2020.03.006

Rider NL, Srinivasan R, Khoury P (2020) Artificial intelligence and the hunt for immunological disorders. Curr Opin Allergy Clin Immunol 20:565–573. https://doi.org/10.1097/ACI.000000000000069

Sadelain M, Rivière I, Riddell S (2017) Therapeutic T cell engineering. Nature 545:423–431. https://doi.org/10.1038/nature22395

Schultze JL (2015) Teaching ‘big data’ analysis to young immunologists. Nat Immunol 16:902–905. https://doi.org/10.1038/ni.3250

Sharma P, Goswami S, Raychaudhuri D et al (2023) Immune checkpoint therapy-current perspectives and future directions. Cell 186:1652–1669. https://doi.org/10.1016/j.cell.2023.03.006

Singh B, Qin N, Reid G (2015) Microbiome regulation of autoimmune, gut and liver associated diseases. Inflamm Allergy Drug Targets 14:84–93. https://doi.org/10.2174/1871528114666160128150747

Singh B, Summers K, Barker G et al (2019) Emergence of human immunoprofiling in health and disease. Curr Trends Immunol 20:11–19.

Wadden JJ (2021) Defining the undefinable: The black box problem in healthcare artificial intelligence. J Med Ethics 48:764–768. https://doi.org/10.1136/medethics-2021-107529

Wang X, Fan D, Yang Y et al (2023) Integrative multi-omics approaches to explore immune cell functions: Challenges and opportunities. iScience 26:106359. https://doi.org/10.1016/j.isci.2023.106359

Xie J, Luo X, Deng X et al (2023) Advances in artificial intelligence to predict cancer immunotherapy efficacy. Front Immunol 13:1076883. https://doi.org/10.3389/fimmu.2022.1076883

Yu J, Peng J, Chi H (2019) Systems immunology: Integrating multi-omics data to infer regulatory networks and hidden drivers of immunity. Curr Opin Syst Biol 15:19–29. https://doi.org/10.1016/j.coisb.2019.03.003

Yurina V, Adianingsih OR (2022) Predicting epitopes for vaccine development using bioinformatics tools. Ther Adv Vaccines Immunother 10:25151355221100218. https://doi.org/10.1177/25151355221100218

Zhang GL, Sun J, Chitkushev L et al (2014) Big data analytics in immunology: A knowledge-based approach. Biomed Res Int 2014:437987. https://doi.org/10.1155/2014/437987

Zheng D, Liwinski T, Elinav E (2020) Interaction between microbiota and immunity in health and disease. Cell Res 30:492–506. https://doi.org/10.1038/s41422-020-0332-7