Essays

Omics Sciences and Artificial Intelligence: Future Directions for Tailored Social Medicine

Marianna Talia, Eugenio Cesario, Rosamaria Lappano, Marcello Maggiolini *

ABSTRACT: Biomedical research is rapidly advancing through the convergence of omics sciences with artificial intelligence (AI) applications. Genomics, transcriptomics, proteomics, and metabolomics, among others, generate multidimensional data that embrace molecular complexity of diseases, whereas AI enables the integration, interpretation, and prediction from these datasets. Together, they contribute to enhance patient-tailored medicine by supporting biomarker discovery, disease classification, patient stratification, and personalized therapies. However, challenges such as data quality, cost, reproducibility, and model interpretability remain. Emerging strategies including federated learning and large language models provide promising solutions, bridging precision and social medicine to promote health equity, improve clinical decision-making, and maximize the societal impact of digital health innovations.

KEYWORDS: Omics data, artificial intelligence, machine learning, social medicine, personalized medicine

1. Introduction

Biomedical research is experiencing unprecedented progress through the integration of digital technologies. A wide array of digital health innovations, including interventions, applications and devices are being continuously developed and optimized in order to refine personalized medicine from the perspectives of patient, healthcare professional, healthcare infrastructures and industry stakeholders.[1] In this context, recent advancements in omics technologies, including genomics, transcriptomics and proteomics, is empowering our knowledge on various diseases, offering insights in-

* *Marianna Talia: Assistant Professor - RTDA, Department of Pharmacy, Health and Nutritional Sciences, University of Calabria. Mail: marianna.talia@unical.it; Eugenio Cesario: Associate Professor, Department of Cultures, Education and Society, University of Calabria. Mail: eugenio.cesario@unical.it; Rosamaria Lappano: Associate Professor, Department of Pharmacy, Health and Nutritional Sciences, University of Calabria. Mail: rosamaria.lappano@unical.it; Marcello Maggiolini, Full Professor, Department of Pharmacy, Health and Nutritional Sciences, University of Calabria. Mail: marcello.maggiolini@unical.it.* *This article was subject to a blind peer review process.*

[1] N. BIT-AVRAGIM, J. BOUSQUET, S. CANTÙ, S. OMBONI, E. RAVOT, P. TUNNAH, *The evolving reality of digital health*, in *Digit Health*, 10, 2024.

to the underlying molecular mechanisms.[2] Additionally, artificial intelligence (AI) has become an essential resource to interrogate the huge amount of omics data generated by high-throughput technologies, enabling to capture the full degree of disease-related complexity. In this scenario, the convergence of omics and AI holds great promise in precision and social medicine. Indeed, integrating patient-derived data can disclose hidden patterns that predict disease, inform diagnosis, guide care and support both patients and clinicians in decision making, potentially outperforming existing conventional protocols. While AI provides the analytical and statistical powers to interrogate datasets, omics sciences constitute the backbone of this information. By capturing multiple layers of biological organization in both health and disease, integrating DNA sequence data, RNA expression as well as protein and metabolite abundance, omics approaches create a bridge that links molecular profiling to phenotype manifestation.[3] A pivotal milestone in the rise of omics studies is represented by the human genome sequencing project, which was achieved in the early 2000s through two independent initiatives, the publicly funded Human Genome Project and a private effort led by J. Craig Venter. Genomics investigations catalogue the entire DNA of an organism, aiming to identify gene structures, functions and interactions, as well as their correlation to biological processes.[4] Beyond the initial view of describing the human genome through a single reference sequence, recent studies revealed a growing collection of human genomes that reflect population-specific variations, encompassing single nucleotide polymorphisms (SNPs), insertions and deletions, copy number variations (CNVs) and structural variants.[5] Beginning with genomics, subsequent omics disciplines, depicting additional layers of gene expression, have emerged and led to the generation of multi-dimensional datasets.[6] The genome dynamic expression, represented by the entire RNA set, reveals how genes are differentially regulated within specific tissues or conditions. Advances in RNA-sequencing (RNA-seq) approaches have greatly expanded our ability to study the transcriptome respect to previous microarray-based approaches. RNA-seq has provided critical insights into the complexity of gene expression regulation, including alternative splicing and the roles of non-coding and enhancer RNAs in modulating transcriptional activity.[7] Following these advancements, single-cell RNA-seq (scRNA-seq), was developed to explore the transcriptional dynamics at single cell resolution.[8] More recently, spatial transcriptomics has further implemented RNA-seq technology by preserving the spatial context of gene expression within tissues, allowing to map the cellular organization and interactions at unprecedented resolution. In oncology, these approaches represent precious resources able to reveal how tumor cells interact and arrange toward immune evasion, drug resistance and metastatic features, provid-

[2] A. Sharma, A. Lysenko, S. Jia, K.A. Boroevich, T. Tsunoda, *Advances in AI and machine learning for predictive medicine*, in *J Hum Genet*, 69, 2024, 487–497.
[3] K.Y.X. Wang, G.M. Pupo, V. Tembe, E. Patrick, D. Strbenac, S-J. Schramm, *et al.*, *Cross-Platform Omics Prediction procedure: a statistical machine learning framework for wider implementation of precision medicine*, in *NPJ Digital Med.*, 5, 2022, 85.
[4] C. Manzoni, D.A. Kia, J. Vandrovcova, J. Hardy, N. W. Wood, P. A. Lewis, *et al.*, *Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences*, in *Brief Bioinform*, 19, 2018, 286–302.
[5] R.M. Sherman, S.L. Salzberg, *Pan-genomics in the human genome era*, in *Nat Rev Genet*, 21, 2020, 243–254.
[6] M. Mann, C. Kumar, W.F. Zeng, M.T. Strauss, *Artificial intelligence for proteomics and biomarker discovery*, in *Cell Syst*, 12, 2021, 759–770.
[7] X. Li, C.-Y. Wang., *From bulk, single-cell to spatial RNA sequencing*, in *Int J Oral Sci*, 13, 2021, 36.
[8] D. Jovic, X. Liang, H. Zeng, L. Lin, F. Xu, Y. Luo, *Single-cell RNA sequencing technologies and applications: A brief overview*, in *Clin Transl Med.*, 12, 2022, e694.

ing critical insights for precision medicine.[9] Mass spectrometry (MS)-based proteomics and metabolomics finally represents the functional output of the genome and transcriptome, assessing the dynamic molecular layers linked to cellular phenotypes in health and disease conditions.[10] Although the comprehensive understanding of biological systems derived from omics studies, the high dimensionality of data, as well as the intricate relationships existing between data layers represent analytical challenges. In this vein, machine learning (ML) and deep learning (DL) approaches enable the extraction of biological and clinically meaningful insights from complex datasets, thereby paving the way for precision and social medicine. AI facilitates biomarker discovery, disease classification, patient stratification, disease risk prediction, and target identification toward personalized therapies. ML, a branch of AI, learns from data to improve prediction tasks, whereas DL, a sub-field of ML, uses multi-layered models for automated feature extraction and pattern recognition in complex datasets. Nevertheless, most DL applications hold limited transparency and explainability that restrain their applicability in omics studies.[11] Beyond model interpretability, additional challenges are represented by data quality and standardization since large differences exist across different protocols and platforms regarding specificity, sensitivity, chemistry of library construction and bioinformatics.[12] Additionally, in many cases, datasets contain a limited number of patients, for example due to phenotype rarities, making the reproducibility a critical concern.[13] Another limitation is related to the high costs for generating omics data that often limits cohort size, reducing the statistical power and the reproducibility of ML models. Aiming to maximize the value of clinical and related omics data, politicians, funders, and publishers should support and implement data sharing policies that further restrict biomedical advancements. Simultaneously, researchers should be trained to effectively reuse existing datasets to strengthen their studies and the robustness of their conclusions.[14] In this context, federated learning (FL), a distributed ML approach, where data is decentralized and models are trained locally, is emerging as a promising strategy for omics data analysis further supporting personalized and social medicine. FD overcomes the problem of collecting and integrating data from medical institutions, enabling collaborative model training without the need to centralize sensitive patient data.[15] At the same time, large language models (LLMs) could enhance trust and facilitate clinical adoption thanks to their capability in language understanding and the ability to tackle new tasks through in-context learning.[16]

---

[9] X. Li, C.Y. Wang, *op. cit*.

[10] T. Kowalczyk, M. Ciborowski, J. Kisluk, A. Kretowski, C. Barbas, *Mass spectrometry based proteomics and metabolomics in personalized oncology*, in *Biochim Biophys Acta Mol Basis Dis*, 1866, 2020, 165690.

[11] S. Shafi, A.V. Parwani, *Artificial intelligence in diagnostic pathology*, in *Diagn Pathol*, 18, 2023, 109.

[12] M. Wysocka, O. Wysocki, M. Zufferey, D. Landers, *A.* Freitas, *A systematic review of biologically-informed deep learning models for cancer: fundamental trends for encoding and interpreting oncology data*, in *BMC Bioinformatics*, 24, 2023, 198.

[13] X. Chen, Z. Yang, W. Chen, Y. Zhao, A. Farmer, B. Tran, *et al.*, *A multi-center cross-platform single-cell RNA sequencing reference dataset*, in *Sci Data*, 8, 2021, 39.

[14] Y. Liang, A. Kelemen, A. Kelemen, *Reproducibility of biomarker identifications from mass spectrometry proteomic data in cancer studies*, in *Stat Appl Genet Mol Biol*, 18, 2019.

[15] C. Locher, G. Le Goff, A. Le Louarn, U. Mansmann, F. Naudet, *Making data sharing the norm in medical research*. In *BMJ*, 382, 2023, 1434.

[16] Q. Wang, M. He, L. Guo, H. Chai, *AFEI: adaptive optimized vertical federated learning for heterogeneous multi-omics data integration*, in *Brief Bioinform*, 24, 2023.

Social medicine, which focuses on social, economic and cultural conditions that impact health, disease and the practice of medicine[17], may take useful advantages from these technological innovations. It frames biomedical progress within the context of health equity, public health priorities and the reduction of disparities, ensuring that data-driven advances comprehensively allow individual and society benefit.

2. Methods

High-throughput technologies advent has represented a turning point in medical science, enabling the generation of multi-omics datasets.[18] Next generation sequencing (NGS) approaches produce vast amounts of DNA sequence data, offering detailed insights into genetic variation, mutations, as well as DNA methylation profiles.[19] RNA-seq measures average transcript abundance within a sample, encompassing both coding and non-coding RNAs.[20] MS-based methods allow the detection and quantification of proteins and metabolites, thereby completing the multi-omics framework toward a comprehensive phenotype characterization.[21] Advances in genomics, transcriptomics and proteomics have been achieved through the development of single-cell technologies, which capture the complexity of clinical conditions that are characterized by molecular heterogeneity at the single cell level, such as in tumors.[22]
Genomics. Genome sequencing has accelerated the discovery of disease-associated genes, novel gene variants and their related phenotypes.[23] After DNA extraction from biological samples, typical NGS pipelines involve DNA fragmentation and ligation to platform-specific adapters, followed by amplification and sequencing. Next steps comprise quality control (QC) of the reads and alignment to the reference genome.[24] A base-pair resolution across the entire genome is provided by whole genome sequencing (WGS), whereas whole exome sequencing (WES) only targets protein-coding regions.[25] The resulting data allow comprehensive characterization of genomic variation and integrative analyses with further omics layers.

[17] *Z.* Wang, H. Wang, B. Danek, Y. Li, C. Mack, L. Arbuckle, et al., *A perspective for adapting generalist AI to specialized medical AI applications and their challenges*, in *NPJ Digital Med*., 8, 2025, 429.
[18] N. Nour, D. Onchonga, S. Neville, P. O'donnell, M. E. Abdalla, *Integrating the social determinants of health into graduate medical education training: a scoping review*, in *BMC Med Educ*., 24, 2024, 565.
[19] B. Hwang, J.H. Lee, D. Bang, *Single-cell RNA sequencing technologies and bioinformatics pipelines*, in *Exp Mol Med*., 50, 2018, 1–14.
[20] X. Li, C.Y. Wang, *op. cit*.
[21] Y.J. Heo, C. Hwa, G.H. Lee, J.M. Park, J.Y. An, *et al.*, *Integrative multi-omics approaches in cancer research: From biological networks to clinical subtypes*, in *Mol Cells*, 44, 2021, 433–443.
[22] B. Aslam, M. Basit, M.A. Nisar, M. Khurshid, M.H. Rasool, *Proteomics: Technologies and their applications*, in *J Chromatogr Sci.*, 55, 2017, 182–196.
[23] D.T. Melnekoff, A. Laganà, *Single-cell sequencing technologies in precision oncology,* in *Adv Exp Med Biol.*, 1361, 2022, 269–282.
[24] M.L. Metzker, *Sequencing technologies - the next generation*, in *Nat Rev Genet.*, 11, 2010, 31–46.
[25] R. Ramakrishnan, A. Washington, S. Suveena, J.R. Rani, O.V. Oommen, *From DNA to big data: NGS technologies and their applications*, in *Methods Mol Biol.*, 2952, 2025, 459–482.

2.1. Transcriptomics

RNA-seq workflow starts from raw sequencing data and culminates with the identification of differentially expressed genes (DEGs) across experimental groups. It involves mRNA-only library and whole transcriptome library that includes all RNA species except for rRNA.[26] Total RNA is extracted from biological samples, RNA quality is evaluated, and RNA is then fragmented and reverse-transcribed to produce double stranded complementary DNA (cDNA). Sequencing libraries are constructed through adaptor ligation and amplification, followed by sequencing to produce millions of reads.[27] The generated raw sequences undergo QC and filtering to remove low-quality reads and adaptors. Processed reads are then aligned to the reference genome and quantified. The resulting data is then normalized and DEGs are identified by comparing normalized expression profiles across conditions.

2.2. Proteomics and Metabolomics

To comprehensively assess the molecular alterations in biological samples, including protein abundance, peptide composition, and post-translational modifications, high-resolution MS offers the most robust approach.[28] Proteins are extracted using appropriate lysis buffers and enzymatically digested to generate peptides. These peptides are separated by liquid chromatography, ionized through electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI), and analyzed by MS, which measures their mass-to-charge ratios and provide sequence information. The most commonly used mass analyzers are quadrupole, time of flight (ToF), and Orbitrap. Identification and quantification are carried out by comparison with protein databases and spectral matching. Similar approaches are employed to identify metabolic intermediates. Metabolites are extracted using solvent-based protocols and profiled by liquid or gas chromatography coupled to high-resolution MS.[29] Specialized software that detects peaks in the spectra enables alignment, annotation and quantification. In both proteomic and metabolomic analyses, data are subjected to QC, normalization, and statistical evaluation.[30]

2.3. Single-Cell Technologies

Omics studies are increasingly moving to single cell resolution with scRNA-seq being the most employed technique, particularly in oncology.[31] Compared to bulk approaches, individual cells are first isolated through microfluidic or droplet-based systems, lysed to release RNA that is captured through primers or barcoded beads.[32] RNA is then reverse-transcribed into cDNA, amplified, and subjected to high-throughput sequencing. Downstream processing, including QC, read alignment, normalization, and di-

[26] X. Li, C.Y. Wang, *op. cit*.
[27] C. Manzoni, D.A. Kia, J. Vandrovcova, J. Hardy, N.W. Wood, P. A. Lewis, *et al.*, *op. cit.*
[28] S. Rozanova, K. Barkovits, M. Nikolov, C. Schmidt, H. Urlab, K. Marcus, *Quantitative mass spectrometry-based proteomics: An overview*, in *Methods Mol Biol*, 2228, 2021, 85-116.
[29] S. Alseekh, A. Aharoni, Y. Brotman, K. Contrepois, J. D'auria, J. Ewald, *et al.*, *Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices*, in *Nat Methods*, 18, 2021, 747–756.
[30] J. Sun, Y. Xia, *Pretreating and normalizing metabolomics data for statistical analysis*, in *Genes Dis*, 11, 2024, 100979.
[31] Y. Zhang, D. Wang, M. Peng, L. Tang, J. Ouyang, F. Xiong, *et al.*, Single-cell RNA sequencing in cancer research, in J Exp Clin Cancer Res, 40, 2021, 81.
[32] B. Hwang, J. H. Lee, D. Bang, *op. cit.*

mensionality reduction, enable the identification of cell clusters with peculiar transcriptional profiles, allowing the reconstruction of developmental trajectories and functional annotation. A further RNA-seq implementation is spatial transcriptomics, that dissects spatially patterned gene expression across tissue sections.[33] In this context, tissue sections are fixed, stained and imaged; RNA is released, captured by spatially barcoded probes and converted to cDNA for library preparation. Downstream analyses follow approaches developed in single-cell studies. Further single cell techniques include single-cell DNA-sequencing, single-cell proteomics and single-cell metabolomics, although these are less routinely applied due to technical limitations such as low input material, and limited sensitivity.[34]

The advent of high-throughput omics technologies has generated richly layered, high-dimensional datasets that reveal the intricate dynamics of biological systems. AI, and in particular ML and DL techniques, provides advanced approaches and methodologies to extract meaningful patterns and knowledge from high-dimensional, heterogeneous, and complex data.

One of the most widely used AI methodologies in omics is supervised ML, where algorithms learn to predict outcomes based on labeled datasets. For example, classifiers such as support vector machines (SVMs), random forests, and gradient boosting are frequently adopted for biomarker identification, disease classification, and patient stratification, leveraging annotated datasets to guide model inference.[35] These models are particularly valuable for precision medicine, where the objective is to predict patient responses to treatments or to stratify patients according to molecular signatures.

In addition to supervised learning, unsupervised learning plays a crucial role in omics data exploration. In particular, dimensionality reduction methods, clustering algorithms, and self-organizing maps support exploratory analysis by identifying latent structure and molecular subtypes without reliance on labeled data.[36] More precisely, dimensionality reduction methods, such as principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) are frequently applied to visualize complex datasets and to reduce redundancy before further modeling. Moreover, both hierarchical and partitional clustering approaches allow researchers to group samples or genes with similar expression patterns, revealing hidden structures and biological subtypes.

A particularly promising area of AI is the application of DL. Neural networks, including convolutional and recurrent architectures, can model non-linear and hierarchical relationships in omics data. DL has shown success in tasks such as predicting gene–disease associations, inferring regulatory networks, and integrating multi-omics layers. Moreover, autoencoders are often used to perform robust nonlinear transformations for dimensionality reduction and feature extraction. For example, techniques such as denoising autoencoders and variational autoencoders (VAEs) have been deployed to denoise data, stratify

[33] A. Rao, D. Barkley, G.S. França, I. Yanai, *Exploring tissue architecture using spatial transcriptomics*, in *Nature*, 596, 2021, 211–220.

[34] H.M. Bennett, W. Stephenson, C.M. Rose, S. Darmanis, *Single-cell proteomics enabled by next-generation sequencing or mass spectrometry*, in *Nat Methods*, 20, 2023, 363–374.

[35] A. Yetgin, *Revolutionizing multi-omics analysis with artificial intelligence and data processing*, in *Quant Biol.*, 13, 2025.

[36] N. Vahabi, G. Michailidis, *Unsupervised multi-Omics data integration methods: A comprehensive review*, in *Front Genet*, 13, 2022, 854752.

patients, and discover latent omic embeddings.[37] Complementarily, the Multi-view Factorization Auto-Encoder (MAE) integrates biological network constraints into autoencoders to improve the integration of multi-omic data and to boost prediction accuracy.[38]

Another important methodological aspect is multi-omics integration, where AI approaches are used to combine information from different omics layers in order to provide a more comprehensive understanding of biological systems. Methods such as multi-view learning, Bayesian networks, and graph-based models enable the integration of genomics, epigenomics, transcriptomics, and proteomics data, supporting the identification of cross-level interactions and pathways. For example, Graph Neural Networks (GNNs) map multiple omics layers onto graph structures and capture inter-entity relationships via message passing, enabling prediction and inference within network contexts.[39] Also, concatenation-based, transformation-based, and network-based strategies have been recently proposed to harness complementary layers of biological information.[40] Finally, Bayesian relational learning frameworks further facilitate integrative analysis by discovering latent interactions across omics layers via graph-encoded relationships.[41]

3. Results

AI, encompassing ML and DL techniques, has found multiple applications in omics sciences and is profoundly transforming biomedical research. These approaches are enabling the generation of novel results and the extraction of actionable biological insights with direct relevance to medicine. Below, we provide a concise overview of the most significant applications and advancements in this domain.

Biomarker Discovery and Disease Diagnosis. Biomarker discovery is the process of identifying measurable indicators, such as molecules or physiological changes, that signal the presence of a disease, predict its risk, or monitor treatment response. This multi-stage process involves high-throughput screening of biological samples like blood or tissue using techniques such as genomics and proteomics, followed by data analysis and rigorous validation to confirm the candidate markers. The identification of biomarkers across multiple omics layers has become a challenging and innovative task in omics science, with the aim of improving diagnostic and prognostic precision in diseases like cancer and liver disorders. This task is faced by the application of supervised ML algorithms such as SVMs, random forests, and neural networks, which are extensively applied to detect molecular signatures associated with disease.[42]

---

[37] Y. Wen, L. Zheng, D. Leng, C. Dai, J. Lu, Z. Zhang, *et al., Deep learning-based multiomics data integration methods for biomedical application,* in *Adv Intell Syst*, 5, 2023.

[38] T. Ma, A. Zhang, *Integrate multi-omics data with biological interaction networks using Multi-view Factorization AutoEncoder (MAE)*, in *BMC Genomics*, 20, 2019, 944.

[39] W. Jiang, W. Ye, X. Tan, Y.J. Bao, *Network-based multi-omics integrative analysis methods in drug discovery: a systematic review*, in *BioData Min.*, 18, 2025, 27.

[40] Y. Nam, J. Kim, S.H. Jung, J. Woerner, E.H. Suh, D.G. Lee, *et al.*, *Harnessing artificial intelligence in multimodal omics data integration: Paving the path for the next frontier in precision medicine*, in *Annu Rev Biomed Data Sci*, 7, 2024, 225–250.

[41] E. Hajiramezanali, A. Hasanzadeh, N. Duffield, K.R. Narayanan, X. Qian, *BayReL: Bayesian Relational Learning for multi-omics data integration,* 2020.

[42] A. Yetgin., *op. cit.*; T. Wu, S.A. Cooper, V.H. Shah, *Omics and AI advance biomarker discovery for liver disease*, in *Nat Med*, 28, 2022, 1131–1132.

3.1. Patient Stratification and Precision Medicine

Patient stratification refers to the process of categorizing individuals into subgroups based on shared characteristics, such as genetic variants, molecular biomarkers, or clinical features. This stratification forms the foundation of precision medicine, which seeks to tailor therapeutic interventions to the unique biological and clinical profiles of individual patients. By aligning treatments with patient-specific characteristics, precision medicine offers the potential for enhanced therapeutic efficacy, minimized adverse effects, and more efficient drug development compared to traditional "one-size-fits-all" approaches. In oncology, AI and multi-omics integration have become particularly powerful in advancing patient stratification. AI-driven frameworks are capable of uncovering molecular subtypes of cancer associated with distinct prognoses and therapeutic responses. Such insights enable more accurate risk assessment and the identification of patients most likely to benefit from specific interventions. Moreover, DL architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are increasingly employed for classification tasks in high-dimensional omics datasets. These models facilitate the discovery of subtle, non-linear patterns across diverse data types and thereby support the development of highly personalized therapeutic strategies.[43]

3.2. Multi-Omics Integration and Systems-Level Insights

Multi-omics profiling is an emerging approach in which molecular phenomics data from multiple omics layers (including genomes, epigenomes, transcriptomes, proteomes, and metabolomes) are comprehensively measured, analyzed, and integrated from the same set of samples on a genome-wide scale. By capturing and quantifying diverse biological signals across complementary molecular layers, multi-omics profiling enables the exploration of intricate interconnections between biological molecules and supports the identification of system-level biomarkers that cannot be discerned from single-omics analyses alone. AI methodologies play a pivotal role in the integration of these heterogeneous datasets, providing advanced methods to model complex biological interactions. Integration strategies range from concatenation and transformation-based approaches to more sophisticated frameworks such as GNNs and multimodal DL.[44] These methodologies uncover cross-layer relationships that inform systems biology, improve mechanistic understanding, and advance precision medicine applications.

4. Discussion

Human health is shaped by a complex interplay of biological factors, social influences, healthcare access, education, economic conditions, and environmental surroundings. This intricate network of influences helps explain why the World Health Organization identifies the rising burden of non-communicable disease (NCD), or chronic conditions, as a critical public health concern especially in low- and middle-

[43] A. YETGIN., *op. cit*; X. He, X. Liu, F. Zuo, H. Shi, J. Jing, *Artificial intelligence-based multi-omics analysis fuels cancer precision medicine*, in *Semin Cancer Biol*., 88, 2023, 187–200.
[44] Y. NAM, J. KIM, S.H. JUNG, J. WOERNER, E. H. SUH, D.G. LEE, *et al., op. cit.*

income countries.[45] The majority of NCD deaths are due to cardiovascular diseases, accounting for 17.9 million annually, followed by cancers (9.3 million), chronic respiratory diseases (4.1 million) and diabetes (2 million). A growing body of evidence indicate that advances in omics technologies and omics-driven AI have untangled specific molecular mechanisms predictive of, or underlying, these diseases.[46] As an example, in cardiovascular research, genomics approaches enabled the identification of both rare pathogenic variants and common causative genetic variations. Additionally, transcriptomic profiling revealed novel molecular drivers and regulatory networks in nonfailing and failing human hearts and identified druggable targets and variability in patient therapeutic responses. Furthermore, advances in metabolic syndrome research were empowered by the integration of omics data with gut microbiome profiles, constant glucose monitoring, and dietary intake, analyzed by AI-based methodologies that for instance predict post-prandial glycemic responses. This integrative method provided a powerful framework for predicting diabetes risk.[47]

Similar approaches have empowered our knowledge on chronic respiratory disease biology. Omics technologies and AI enable the integration of complex molecular, genetic and clinical data, leading to novel insights in disease mechanisms, biomarkers, and risk factors. Such strategies exemplify how advanced tools uncover shared pathways and predictive features across asthma, chronic obstructive pulmonary disease, and related conditions.[48] Multi-omics approaches have also emerged as powerful tools for accelerating cancer research, allowing to unravel the complex molecular interactions and dysregulations associated with specific tumor cohorts of patients. These strategies have already led to advances regarding molecular subtyping, disease-gene association prediction and drug discovery.[49] Despite the advancements in generating and integrating omics data with AI, several limitations remain. As already mentioned, these include the small sample sizes within individual datasets, difficulties in combining data from diverse cohorts due to technical biases, and the limited availability of information that may be crucial for assuming robust conclusions. Additionally, it is worth mentioning that beyond genetic and physiologic factors, disease occurrence is driven by a combination of global demographic and social tendencies, such as population aging, environmental factors, economic transitions, as well as dietary and lifestyle modifications.[50] In this framework, omics data on NCD alone are insufficient to capture the diversi-

[45] P. Geldsetzer, S. Flores, G. Wang, B. Flores, A.B. Rogers, A. Bunker, *et al.*, *A systematic review of healthcare provider-targeted mobile applications for non-communicable diseases in low- and middle-income countries*, in *NPJ Digit Med*, 5, 2022, 99.

[46] X. He, X. Liu, F. Zuo, H. Shi, J. Jing, *op. cit.*; R.S. Wang, B.A. Maron, J. Loscalzo, *Multiomics network medicine approaches to precision medicine and therapeutics in cardiovascular diseases*, in *Arterioscler Thromb Vasc Biol*., 43, 2023, 493–503.

[47] M. Subramanian, A. Wojtusciszyn, L. Favre, S. Boughorbel, J. Shan, K. B. Letaief, *et al.*, *Precision medicine in the era of artificial intelligence: implications in chronic disease management*, in *J Transl Med*, 18, 2020; 472.

[48] E. Maiorino, S.H. Baek, F. Guo, X. Zhou, P.H. Kothari, E.K. Silverman, *et al.*, *Discovering the genes mediating the interactions between chronic respiratory diseases in the human interactome*, in *Nat Commun*, 11, 2020, 811; Y. Guo, Q. Liu, Z. Zheng, M. Qing, T. Yao, B. Wang, *et al.*, *Genetic association of inflammatory marker GlycA with lung function and respiratory diseases*, in *Nat Commun*, 15, 2024, 3751; S. Aishwarya, P. C. Siddalingaswamy, K. Chadaga, *Explainable artificial intelligence driven insights into smoking prediction using machine learning and clinical parameters,* in *Sci Rep*, 15, 2025, 24069.

[49] X. He, X. Liu, F. Zuo, H. Shi, J. Jing, *op. cit.*

[50] P. Geldsetzer, S. Flores, G. Wang, B. Flores, A.B. Rogers, A. Bunker, *et al.*, *op. cit.;* R. Ferrara, R. Nappo, F. Ansermet, P. Ricci, F. Massoni, G. Carbone, *et al.*, *The impact of dsm-5on the diagnosis of autism spectrum disorder*, in *Psychi-*

ty across different populations and the gathering of additional data beyond omics, is an imperative need toward both precision and social medicine purposes.[51] Hence, the inclusion of data from heterogeneous resources, including clinical records, socioeconomic indicators, geospatial exposures, and environmental monitoring systems may aid omics-driven AI models to unravel hidden causative relationships. Cumulative measures of environmental exposure can be therefore combined with omics data to uncover how factors such as pollution, diet, accessibility to healthcare, among others, shape NCD trajectories. AI offers a transformative framework for this integration with ML algorithms that are increasingly applied to harmonize multi-layered datasets. Furthermore, LLMs may be employed to facilitate the communication of these insights to diverse stakeholders, ranging from researchers with diverse backgrounds, clinicians, bioinformaticians, data scientists, epidemiologists, ethicists, and policymakers as well as patients.[52] These approaches may promote a more inclusive understanding of health within a broader social and ecological ecosystem that comprise multidisciplinary collaborations and targets. Likewise, important ethical and governance challenges arise regarding privacy protection and fair access to data. To this aim, AI models must ensure transparency and explainability.[53] All these factors combined together will contribute to design frameworks that advance both precision medicine and precision public health, helping to tackle the combined challenge of molecular complexity and health inequities, especially in NCD.

5. Conclusions

The integration of omics technologies and AI offers unique chances toward the advancement of both personalized and social medicine, particularly in addressing the complexity of NCD. Future progress will depend on the ability to integrate molecular data with social, environmental, and lifestyle determinants, while ensuring methodological robustness and reducing technical and cohort-related biases.

At the same time, ethical determinants remain central. Transparency, fairness, and explainability of AI systems must be prioritized to ensure equitable benefits across populations.[54] Importantly, patients should be regarded as active partners in this process: their autonomy, informed consent, and freedom of choice must guide the clinical implementation of AI-driven approaches.

---

atric Annals, 51, 1, 2020, 38-46; T. Archer, S. Ricci, F. Massoni, L. Ricci, M. Rapp-Ricciardi, *Cognitive benefits of exercise intervention*, in *Clinica Terapeutica*, 167, 6, 2016, 180-85; M.V. Rosati, A. Sancini, F. Tomei, C. Sacco, V. Traversini, A. De Vita, D.P. De Cesare, G. Giammichele, F. De Marco, F. Pagliara, F. Massoni, L. Ricci, G. Tomei, S. Ricci, *Correlation between benzene and testosterone in workers exposed to urban pollution*, in *Clinica terapeutica*, 168, 6, 2017, e380-87; P. Ricci, F. Massoni, L. Ricci, E. Onofri, G. Donato, S. Ricci, *Quality of life in dementia sufferers: The role of diet and exercise*, in *Current Alzheimer Research*, 15, 5, 2018, 400-07; R. Businaro, M. Corsi, G. Azzara, T. Di Raimo, G. Laviola, E. Romano, L. Ricci, M. Maccarrone, E. Aronica, A. Fuso, S. Ricci, *Interleukin-18 modulation in autism spectrum disorders*, in *Journal of Neuroinflammation*, 13, 1, 2016, 1-13.

[51] M. Pentecost, V. Adams, R. Baru, C. Caduff, J.A. Greene, H. Hansen, *et al.*, *Revitalising global social medicine*, in *Lancet*, 398, 2021, 573–574.

[52] C. Lin, C.F. Kuo, *Roles and potential of large Language Models in healthcare: A comprehensive review*, in *Biomed J*, 2025, 100868.

[53] L. Ricci, B. Di Nicolò, P. Ricci, F. Massoni, S. Ricci, *The exercise of rights beyond therapy: on Human Enhancement*, in *BioLaw Journal*, 1, 2019, 497-512; S. Ricci, A. Miglino, *Medicina e Società: dalla tutela dell'integrità fisica al diritto alla salute*, Roma, 2005, 1-99.

[54] F.M. Damato, P. Ricci, R. Rinaldi, *Informed consent and compulsory treatment on individuals with severe eating disorders: a bio-ethical and juridical problem*, in *Clinica Terapeutica*, 174, 4, 2023, 365-69.

In this perspective, the transformative potential of omics-informed AI lies both in predictive accuracy and therapeutic innovation, as well as in fostering a more participatory and socially-centered model of medicine.