Review

A Comprehensive Review of the Impact of Machine Learning and Omics on Rare Neurological Diseases

Nofe Alganmi [1,2,3]

1   Department of Computer Science, Faculty of Computing and Information Technology,
    King Abdulaziz University, Jeddah 21589, Saudi Arabia; nalghanimi@kau.edu.sa
2   Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia
3   Centre of Artificial Intelligence in Precision Medicines, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Abstract: Background: Rare diseases, predominantly caused by genetic factors and often presenting neurological manifestations, are significantly underrepresented in research. This review addresses the urgent need for advanced research in rare neurological diseases (RNDs), which suffer from a data scarcity and diagnostic challenges. Bridging the gap in RND research is the integration of machine learning (ML) and omics technologies, offering potential insights into the genetic and molecular complexities of these conditions. **Methods:** We employed a structured search strategy, using a combination of machine learning and omics-related keywords, alongside the names and synonyms of 1840 RNDs as identified by Orphanet. Our inclusion criteria were limited to English language articles that utilized specific ML algorithms in the analysis of omics data related to RNDs. We excluded reviews and animal studies, focusing solely on studies with the clear application of ML in omics data to ensure the relevance and specificity of our research corpus. **Results:** The structured search revealed the growing use of machine learning algorithms for the discovery of biomarkers and diagnosis of rare neurological diseases (RNDs), with a primary focus on genomics and radiomics because genetic factors and imaging techniques play a crucial role in determining the severity of these diseases. With AI, we can improve diagnosis and mutation detection and develop personalized treatment plans. There are, however, several challenges, including small sample sizes, data heterogeneity, model interpretability, and the need for external validation studies. **Conclusions:** The sparse knowledge of valid biomarkers, disease pathogenesis, and treatments for rare diseases presents a significant challenge for RND research. The integration of omics and machine learning technologies, coupled with collaboration among stakeholders, is essential to develop personalized treatment plans and improve patient outcomes in this critical medical domain.

Keywords: neurological diseases; machine learning; omics; rare diseases

1. Introduction

Global health has improved, but noncommunicable diseases (NCDs) now dominate health system burdens, with one in three adults globally affected and incurring substantial healthcare costs [1–4]. In contrast, rare diseases (RDs), though individually scarce, collectively impact a significant global population [5–10]. RDs pose challenges due to their rarity, global distribution, complex pathophysiology, and high medical costs [11]. They often incur costs similar to those of common conditions like Alzheimer's, cardiovascular diseases, and cancer [12]. RDs also impact health, psychology, and social well-being, with slow research and treatment development due to diagnostic challenges, dispersed patient populations, a limited disease understanding, and inadequate funding [13–16]. "Orphan drugs" for RDs are significantly more costly than drugs for common diseases [17,18].

Most RDs are genetically rooted [19], a fact that omics technologies can exploit to accelerate diagnosis and drug discovery. DNA and RNA sequencing advancements have led

to various genomic analysis techniques, like whole-exome sequencing, whole-genome sequencing (WGS), and single-cell RNA sequencing, providing deep genomic insights [20,21]. These omics investigations unveil disease aspects previously obscured by traditional approaches. For example, WGS has identified pathogenic variants in rare epilepsies and the genetic causes of rare diseases [20]. Omics extends beyond genomic resolution, including proteomics, metabolomics, epigenomics, and lipidomics, which assess proteins, metabolites, DNA machinery, and lipids [22]. Radiomics, a novel omics field, involves high-throughput medical imaging assessments [23]. Artificial intelligence (AI) and machine learning algorithms analyze these diverse data, enabling reanalysis for research and healthcare solutions [24]. Machine learning, a key interest area, involves training algorithms on large datasets to predict unseen data. The algorithms fall into three categories: supervised learning (learning from labeled data), unsupervised learning (finding patterns in input data without targets), and reinforcement learning (action–reward-based learning) [25]. AI can assist RD research and treatment, aiding in variant classification, biomarker identification, gene interactions, and the understanding of protein and metabolite profiles [26]. It facilitates disease diagnosis and prognosis by integrating phenotype data with omics data, discovering new drug molecules, and managing patient registries and rare disease databases. This review assesses omics and AI in a combined approach to overcome the RD treatment challenges. Understanding RDs' molecular pathophysiology and drug development is crucial, especially for neurological RDs involving the nerves, muscles, and brain. AI enhances pharmaceutical development with automated processes, efficiency, and unconventional insight generation. The focus is on compiling ML applications exploring omics data in rare neurological disorders and raising the awareness of AI and omics in rare diseases. A list of some of the rare neurological disorders and a brief explanation of the algorithms and omics data described in this review are given separately in Table 1.

Table 1. Important rare neurological disorders with algorithms and omics data.

Study Reference	Study Name	AI Model	Study Findings	Study Data Source	Study Source Code
[27]	SpliceAI	32-layer deep convolutional neural network	Predicts splice junctions, identifies ASD mutations	E-MTAB-7351	https://github.com/Illumina/SpliceAI (accessed on 2 March 2024)
[28]	VEST	Random forest	Prioritizes rare, disease-causing gene variants	Human Gene Mutation Database (HGMD), https://www.hgmd.cf.ac.uk/ac/index.php (accessed on 2 March 2024)	https://www.cravat.us/CRAVAT/ (accessed on 2 March 2024)
[29]	AluScanCNV	Machine learning-based selection for CNV features	Identifies glioma copy number losses in cancer tissue	https://static-content.springer.com/esm/art:10.1186/s13336-014-0015-z/MediaObjects/13336_2014_15_MOESM2_ESM.xlsx (accessed on 2 March 2024)	https://static-content.springer.com/esm/art:10.1186/s13336-014-0015-z/MediaObjects/13336_2014_15_MOESM3_ESM.zip (accessed on 2 March 2024)
[30]	PURA Syndrome Study	Exome sequencing, AI algorithms, quantum mechanics-based molecular models	Identifies PURA syndrome mutations, AI-aided structural analysis	Not available	https://github.com/google-deepmind/alphafold (accessed on 2 March 2024)
[31]	PhenoApt	ML-based graph embedding techniques	Accelerates Mendelian diagnosis via gene identification	https://github.com/phenoapt/phenoapt (accessed on 2 March 2024)	https://github.com/phenoapt/phenoapt (accessed on 2 March 2024)
[32]	DOMINO	Linear discriminant analysis	Evaluates gene dominance in Mendelian disorders	https://www.cell.com/ajhg/fulltext/S0002-9297(17)30368-3#supplementaryMaterial (accessed on 2 March 2024)	https://domino.iob.ch/ (accessed on 2 March 2024)
[33]	Neurological Disease ML Study	Multinomial linear model	Predicts neurological disorders using clinical data	https://github.com/SimonLammmm/ukbb-ndd-ml (accessed on 2 March 2024)	https://cran.r-project.org/web/packages/nnet/index.html (accessed on 2 March 2024)
[34]	PPML	Random forest classifier	Classifies PKU phenotype based on PAH gene mutations	http://www.bioinfogenetics.info/PPML/ (accessed on 2 March 2024)	http://www.bioinfogenetics.info/PPML/ (accessed on 2 March 2024)
[35]	PIUMet	Network-based algorithm (ML, statistical analysis, network optimization)	Identified Huntington's disrupted metabolites	https://fraenkel-nsf.csbi.mit.edu/piumet2/ (accessed on 2 March 2024)	https://fraenkel-nsf.csbi.mit.edu/piumet2/ (accessed on 2 March 2024)

Table 1. *Cont.*

Study Reference	Study Name	AI Model	Study Findings	Study Data Source	Study Source Code
[36]	Trevino et al.—Human Corticogenesis	Deep learning model derived from BP-Net	Predicted ASD mutations in corticogenesis	GSE162170	https://github.com/GreenleafLab/Brain_ASD (accessed on 2 March 2024)
[37]	Wilscher et al.—Grade I–IV Gliomas	Self-organizing maps (SOMs)	Mapped gene patterns in gliomas	GSE61374 GSE129477 GSE53733	https://www.izbi.uni-leipzig.de/opossom-browser/ (accessed on 2 March 2024)
[38]	Loeffler-Wirth et al.—Brain Transcriptome and Methylome	Self-organizing maps (SOMs)	Identified aging gene sets in gliomas	GSE11512	https://bioconductor.org/packages/release/bioc/html/oposSOM.html (accessed on 2 March 2024)
[39]	Huang et al.—Arcuate Organoids	ML-based analysis on single-cell RNA sequencing	Explored arcuate nucleus dysregulation in Prader–Willi syndrome	GSE164101 GSE164102	https://bioconductor.org/packages/release/bioc/html/oposSOM.html (accessed on 2 March 2024)
[40]	Huntington's Disease Gene Expression Study	Random forest classifier	Identified early-onset genes in Huntington's disease	GSE64810 GSE129473	https://bitbucket.org/bubfnexus/asymptomatic_hd_mrnaseq (accessed on 2 March 2024)
[41]	ConTeXT Framework—Neurodevelopmental Disorders	CoNTeXT framework	Linked ASD to neurodevelopmental disorders	GSE69838	https://context.semel.ucla.edu/ (accessed on 2 March 2024)

2. Difficulties in Disease Mechanism Investigation and Biomarker Discovery

One of the prime challenges in the rare disease diagnosis domain is the lack of understanding of the disease and the mechanisms that cause it. Since the molecular pathophysiological factors of rare diseases are unknown, clinicians find it difficult to link the symptoms between different organ systems and differentiate between disorders with overlapping symptoms. The lack of valid parameters and biomarkers, as well as the low frequency of occurrence of the disease, makes it difficult to derive statistically significant and clinically relevant parameters that can assist in diagnosis. However, next-generation sequencing (NGS) technologies such as whole-genome sequencing, whole-exome sequencing, and DNA methylation techniques are now being commonly utilized in the research and diagnosis of rare diseases. One of the clear advantages of NGS is the ability to interrogate multiple targets at the same time, making it possible to uncover the molecular heterogeneity between and within rare neurological diseases. The main challenges discussed in this section are summarized in Table 2.

Table 2. Summary of the main challenges discussed in Section 2.

Challenge Name	Challenge Description	Impact on RNDs	Available Solutions
Mutation Detection	Difficulty in detecting pathogenic deep intronic variants using traditional sequencing methods.	Delays in diagnosing RNDs, affecting patient treatment and care.	ML tools like SpliceAI for splice junction prediction are enhancing mutation detection accuracy in RNDs.
AI in Tumor Identification	Detecting genetic mutations and CNVs that contribute to tumor progression.	Improves accuracy in tumor classification, diagnosis, and understanding of tumor biology, leading to better-targeted treatments for neurological conditions.	Tools like AluScanCNV for efficient CNV calling in glioma samples. Use of one-dimensional convolutional neural networks to analyze CNVs from NGS data, aiding in detection of genetic abnormalities like 1p/19q co-deletion. Employment of advanced techniques such as exome sequencing, Alpha Fold, and QM-MM analyses to identify and understand mutations at molecular level.
Genotype–Phenotype Integration	Integrating genomic data with phenotypic and clinical features to predict outcomes and identify biomarkers, with challenges in gene prioritization and mutation identification.	Enhances understanding of disease heritability, aids in accurate diagnosis, and informs treatment strategies by linking genetic mutations to clinical manifestations.	ML tools like PhenoApt for gene prioritization using data from HPO, OMIM, and Orphanet.
Omics Data Integration for Disease Characterization	Challenges in omics data integration and how to differentiate pathological and genetic features in diseases with overlapping symptoms.	Accurate metabolite identification and disease characterization are crucial in understanding biochemical activity and developing targeted treatments.	Implement network-based algorithms like PIUMet for metabolite identification from LC-MS data, integrating protein–metabolite interactions. Use ML models like LASSO and random forest to analyze RNA seq data, highlighting biologically relevant genes and pathways in diseases like ALS and FTD.

Table 2. *Cont.*

Challenge Name	Challenge Description	Impact on RNDs	Available Solutions
Disease Mechanisms and Research Models	Utilizing ML and omics to link gene variants, phenotypes, and clinical features for understanding of pathogenesis and development of models.	Enhances understanding of disease mechanisms, facilitates development of experimental models, and aids in early diagnosis and prognosis.	Apply deep learning models, like BP-Net, for gene-regulatory circuit mapping in corticogenesis, revealing ASD-related mutations. Use self-organizing maps (SOMs) for detailed molecular mapping in gliomas and brain development studies, identifying potential biomarkers. Develop organoids, such as ARCOs, to model neurodevelopmental disorder pathogenesis, aiding in understanding disease mechanisms and identifying therapeutic targets.

2.1. Mutation Detection or Prediction

Detecting pathogenic variants in genomes is crucial for diagnosis and in guiding precision medicine. Deep intronic variants, often challenging to detect via whole-exome sequencing, play a role in multiple disorders [42]. Machine learning tools like SpliceAI and SpliceRover are revolutionizing this area of detection. SpliceAI, using a 32-layer deep convolutional neural network, predicts splice junctions from pre-mRNA, identifying cryptic splicing variants [43]. It has successfully identified de novo mutations in conditions like intellectual disability and autism spectrum disorder (ASD), with observed enrichment in these disorders. SpliceRover employs convolutional neural networks to identify splice sites, offering a more nuanced analysis compared to traditional probabilistic methods [27]. It detected a significant cryptic exon in Joubert syndrome [44]. Additionally, tools like the Variant Effect Scoring Tool (VEST) use algorithms like random forest to prioritize gene variants for diseases like Freeman–Sheldon syndrome and Miller syndrome, outperforming other tools in missense variant prioritization [28]. These advancements highlight the growing role of machine learning in understanding complex genetic variations and rare neurological conditions.

2.2. AI in Tumor Identification

The application of bioinformatics and AI in tumor studies, particularly for brain tumors classified by growth rate and recurrence, is advancing tumor diagnosis and treatment. Glioma tumors, originating from mutations in glial cells, are sub-classified as astrocytomas, oligodendrogliomas, or ependymomas and graded based on their aggressiveness [45]. Pediatric and adult brain tumors exhibit copy number alterations (CNAs), contributing to genomic instability and tumor progression [46–48]. CNV calling from sequencing data, particularly AluScan, is complex, but AluScanCNV has been developed for efficient CNV calling, distinguishing non-cancerous and cancerous tissues in glioma samples [29].

Molecular testing, crucial for the diagnosis of oligodendroglial tumors, requires the detection of IDH gene mutations and 1p/19q co-deletion [49]. A one-dimensional convolutional neural network analyzed CNVs from NGS data to detect 1p/19q co-deletion in 61 tumors, validated against 427 low-grade glial tumors from The Cancer Genome Atlas [49]. In PURA syndrome research, exome sequencing and AI algorithms identified a de novo mutation, c.697-699del p.Phe233del in the PURA gene, with structural analysis using Alpha Fold and hybrid quantum mechanics–molecular mechanics (QM-MM) analyses [30,50]. This study marks a significant advancement in understanding the functional impact of mutations at an atomic level, laying the groundwork for future functional analyses.

2.3. Genotype–Phenotype Integration

Integrating genomic data with phenotype and clinical features enhances models that predict phenotypic traits and outcomes, revealing biomarkers and insights into the

heritability of complex traits. PhenoApt, using ML-based graph embedding techniques, prioritizes genes for Mendelian disorder diagnoses by mapping data from HPO, OMIM, and Orphanet [31]. It assigns scores based on phenotype–gene vector representations, aiding in gene prioritization.

DOMINO, another tool, focuses on identifying dominant mutations in Mendelian disorders, a challenge due to frequent non-pathogenic heterozygous variants [32]. It uses linear discriminant analysis on genomic data, protein interactions, and structures, trained on 985 genes with known Mendelian inheritance patterns. In epilepsy and intellectual disability cases, DOMINO accurately identified known genes and predicted new candidates. An ML study on genotyping and clinical data from neurological disease patients developed a multinomial linear model that accurately identified 88% of disease samples, emphasizing the importance of age and cognition [33]. This analysis also found common SNPs across neurological diseases, linking MND to RBBP5 and TNF, and MG to oncogenes and brain-related genes. In phenylketonuria (PKU), the PPML machine learning framework predicts the PKU phenotype based on nucleotide mutations and amino acid changes [34]. Using a random forest classifier, it accurately classified PKU into classical, mild, and mild hyperphenylalaninemia categories, enhancing the genotype-to-phenotype linkage, crucial for treatment strategy and prognosis prediction.

2.4. Omics Data Integration for Disease Characterization

To develop effective disease treatments, understanding the affected molecules and their interactions is key, with metabolites playing a crucial role as they reflect the biochemical activity in cells. Liquid chromatography–mass spectrometry (LC-MS) is commonly used to globally measure metabolites, but identifying them can be challenging due to multiple metabolites matching a single peak [35]. Pirhaji et al. developed PIUMet, a network-based algorithm integrating protein and metabolite interactions to identify metabolites from LC-MS peaks. Utilizing ML, statistical analysis, and network optimization, PIUMet infers putative metabolites and dysregulated pathways. Applied to Huntington's disease (HD) data, it identified disrupted features like the sphingolipid subnetwork and steroid metabolism.

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) share pathological, clinical, and genetic features, including C9orf72 repeat expansion [51]. Dickson et al. analyzed RNA seq data from the frontal cortex tissue of FTLD and FTLD/MND patients to understand their clinical variability. Although the initial regression models did not yield significant genes post-adjustment, ML models like LASSO and random forest regression with leave-one-out cross-validation highlighted biologically relevant genes consistently associated with outcomes. Genes such as VEGFA, CDKL1, EEF2K, and SGSM3 were promising, with VEGFA linked to the disease onset age.

2.5. Disease Mechanisms and Research Models

ML and omics technologies are instrumental in understanding rare neurological diseases, revealing the connections between gene variants, phenotypes, and clinical features. These technologies have advanced the knowledge of disease pathogenesis and aided in developing experimental models. Trevino et al. used single-cell methods to map the gene-regulatory circuit in human corticogenesis, employing a deep learning model derived from BP-Net [36]. This model predicted genetic variants' impacts on epigenomic elements and highlighted ASD-related mutations. Wilscher et al. utilized self-organizing maps (SOMs) for detailed mapping from the transcriptomics and DNA methylation data of gliomas [37]. Their high-resolution molecular map revealed connections between gene expression, methylome changes, and the tumor microenvironment, offering insights into glioma subtypes and prognosis. Loeffler-Wirth et al. implemented SOMs to analyze the transcriptome and methylome in developing and aging brains [38]. Their maps showed gene expression and methylation changes over the lifespan, identifying gene sets impacting gliomas and providing potential biomarkers.

Huang et al. created arcuate organoids (ARCOs) from human iPSCs to model hypothalamic arcuate nucleus development in neurodevelopmental disorders [39]. They found abnormal differentiation and transcriptomic dysregulation in ARCOs from Prader–Willi syndrome patients, demonstrating their value in studying early human arcuate development in these diseases. For Huntington's disease, gene expression profiles from the caudate nuclei of asymptomatic HD+ individuals were compared with those of symptomatic HD individuals and healthy controls [40]. A random forest classifier identified genes potentially involved in the early onset of the disease.

In neuropsychiatric research, the SH-SY5Y neuroblastoma cell line is commonly used. The CoNTeXT framework, an ML algorithm, estimates the developmental stage and regional identity of transcriptomic signatures [41]. This study found significant gene overlaps in ASD, Fragile X Syndrome, intellectual disability, and schizophrenia, highlighting pathways specific to each disorder during early neurodevelopment. The landscape of the diagnosis of rare neurological diseases is evolving rapidly, transitioning from traditional heuristic approaches to more advanced and precise methodologies. Traditional methods, which relied heavily on clinical experience and medical literature, often resulted in a long, uncertain journey towards diagnosis for many patients. In contrast, recent advancements in genomics and data analysis are providing new pathways to understand these complex conditions.

3. Diagnosis

Gene panels, microarrays, and exome sequencing have become pivotal in uncovering the molecular basis of many previously undiagnosed and rare diseases. These techniques, when coupled with long-read technology, transcriptomics, metabolomics, proteomics, and methylome data, are enhancing the precision and speed of diagnosis. The integration of artificial intelligence (AI) with these methods is pushing the boundaries further, allowing for more comprehensive and nuanced analysis. For instance, Choi et al. [52] conducted a systematic evaluation of machine learning algorithms and feature selection methods to classify neuromuscular diseases with remarkable accuracy. Their study utilized support vector machines (SVM) and directed acyclic graphs, achieving a 100% success rate. This breakthrough is significant, as it demonstrates the potential of AI in identifying diseases with complex genetic backgrounds. Similarly, Caputo et al. [53] developed a machine learning-driven protocol for the classification of Facioscapulohumeral Muscular Dystrophy (FSHD) based on DNA methylation patterns, showcasing the ability of these technologies to discern subtle genetic variations. Their methodology, which incorporated various machine learning models, was able to differentiate FSHD patients from controls with high precision. In the realm of idiopathic inflammatory myopathies (IIM), a study analyzed the plasma and urine metabolomes of patients, employing machine learning algorithms to identify specific biomarkers [53]. This approach is essential in diseases like IIM, where subtypes exhibit overlapping symptoms but require distinct treatments.

Moreover, neural networks are proving invaluable in distinguishing conditions such as sporadic Creutzfeldt–Jakob disease (sCJD) from healthy states [54]. By analyzing differentially methylated CpG loci, these models can effectively differentiate sCJD patients with notable accuracy. Advancements in DNA methylation studies are also facilitating the better understanding and diagnosis of conditions like malformations of cortical development (MCDs). Jabari et al. [55] applied machine learning and deep learning to decipher the DNA methylation pattern in MCD, achieving high accuracy and predictive value.

Machine learning models have also been instrumental in pre-diagnostic risk assessments for diseases like amyotrophic lateral sclerosis (ALS) [56,57]. Although challenges remain in accurately differentiating ALS patients from healthy individuals, these studies have identified metabolic dysregulation that occurs years before disease onset. The novel computational method CTD [58] exemplifies the integration of untargeted metabolomics with genomic data, offering a more refined approach to diagnosing inborn errors of metabolism

(IEMs). This method connects metabolite perturbations to disease-specific networks, improving clinical decision-making.

In oncology, AI models like those developed by Zhao et al. [59] and Capper et al. [60] are differentiating between types of brain tumors based on multi-omics data and DNA methylation profiles, demonstrating the potential of AI in precision medicine. The field of radiomics is another area where AI is making significant strides. By extracting quantitative features from radiographic images, machine learning algorithms are enhancing the diagnosis of rare neurological diseases and tumors [61–63]. These models not only compete with but, in some cases, outperform human experts in diagnosing conditions like high-grade gliomas [64,65].

In summary, the integration of AI with genomic and omics technologies is revolutionizing the diagnosis and understanding of rare neurological diseases. By enabling more precise, efficient, and early diagnostics, these advancements hold great promise for patients who have long struggled with undiagnosed conditions. However, this evolving landscape also presents new challenges and opportunities for future research and clinical application.

4. Prognosis

Early diagnosis and optimal care for rare diseases are pivotal, especially for underserved populations. Advances in medical bioinformatics, artificial intelligence (AI), and machine learning (ML) have enabled the identification of disease patterns, the prediction of disease progression, and the assessment of treatment responses. The random forest algorithm, applied to multi-omics data, identified 111 genes linked to survival outcomes in astrocytoma and oligodendroglioma, serving as diagnostic biomarkers [65]. Neural networks have been crucial in identifying prognosis-related genes in neuroblastoma, a common childhood extracranial solid tumor, with 84% sensitivity and 90% specificity for poor-outcome patients [66]. Deep neural networks outperformed support vector machines and random forest in predicting neuroblastoma outcomes from omics data [67].

Linear support vector machines and random forest, trained on various omics data, have been used for predictive classification in neuroblastoma [68]. Integrative network fusion improved prognosis prediction by integrating microarray and aCGH datasets. DNA methylation alterations in neuroblastoma, analyzed using random forest and XGBoost, highlighted distinct methylation patterns as indicators of disease progression [69,70]. Network-based methods have been evaluated for the integration of multi-omics data to predict clinical outcomes in neuroblastoma, achieving 65-80% accuracy [71].

ML algorithms have also been applied to assess genotyping data from medulloblastoma patients, identifying genetic predictors of intellectual functioning post-treatment [72]. In medulloblastoma, logistic regression used mRNA expression and DNA methylation signatures to guide prognosis. A novel framework by Mihaylov et al. integrated gene expression and clinical data from neuroblastoma and breast cancer patients to predict the survival time [73]. Bratulic et al. explored metabolomic profiles for early cancer detection, finding that glycosaminoglycan profiles could be used to detect cancer types with good sensitivity [74]. In ALS, Sparse Canonical Correlation Analysis explored the role of genes in cognitive dysfunction using whole-genome sequencing [75,76]. For epilepsy, random forest and XGBoost identified co-expressed genes linked to the cardiac event risk [77]. The machine learning-driven metabolomic profiling of aneurysmal subarachnoid hemorrhage patients uncovered biomarkers for functional outcomes [78].

Radiomics studies, enhanced by ML, have improved the glioblastoma biopsy guidance and differentiated brain metastases from glioblastoma [79,80] Convolutional neural networks have been used to detect fatty infiltration in neuromuscular diseases, with HRNet being the most effective [81]. Finally, ML regression models have been employed to predict the muscle fat fraction in FSHD, aiding in disease progression assessment [82]. These advancements in AI and ML are transforming the landscape of diagnosis, prognosis, and treatment in rare diseases.

5. Therapeutic Approach

Precision medicine, particularly in the context of rare diseases, is transforming healthcare through a holistic approach that includes diagnosis, treatment, and follow-up tailored to an individual's genetic makeup. Artificial intelligence (AI) and machine learning (ML) are pivotal in this transformation, analyzing diverse data types like clinical features, multiomics data, and medical images and incorporating phenotype, pharmacogenomic, and pharmacokinetic factors.

Gene therapy, especially CRISPR-based tools, is revolutionizing the treatment of rare neurological diseases. Shen et al. developed inDelphi, an ML algorithm, to predict Cas9-induced insertions and deletions with high accuracy, aiding template-free DNA editing for diseases like Hermansky–Pudlak syndrome and Menkes disease [83]. In Duchenne muscular dystrophy (DMD), Nishida et al. and Malueka et al. explored exon-skipping therapies, using AI to identify cryptic exons and classify dystrophin gene exons for potential therapeutic targets [84].

For adamantinomatous craniopharyngioma (ACP), Lin et al. utilized random forest and LASSO regression to identify diagnostic markers S100A2 and SDC1 from gene expression profiles, pinpointing potential drug targets like Pentostatin and Wortmannin [85]. In medulloblastoma (MB), an ML model was used to discover gene expression-based stemness indices and DNA methylation-based stemness indices, leading to the identification of 96 compounds targeting MB pathways [86].

Gilard et al.'s study on glioblastoma used random forest classifiers to differentiate between diseased and control samples based on metabolomic profiles, highlighting phosphatidylcholine (PC aa C36:6) as a key biomarker [87]. De Jong et al. applied various ML models in precision medicine for rare epilepsy conditions, with the XGBoost trees classifier demonstrating notable effectiveness in predicting the drug response [88]. DNA methylation profiling in temporal lobe epilepsy (TLE) patients identified potential biomarkers for the drug response, utilizing ML for accurate prediction [88].

Dahlin et al.'s research on the gut microbiota in drug-resistant epilepsy children revealed the potential benefits of a ketogenic diet, employing ML to analyze the gut microbiome's role in epilepsy [89]. Kurkiewicz et al. approached myotonic dystrophy type 1 (DM1) as a spectrum disorder, using ML models to predict the modal allele length of the DM1 CTG expansion, a crucial factor in disease progression and the treatment response [90].

In summary, AI and ML are pivotal in advancing precision medicine, especially in rare diseases, by enabling personalized treatment strategies based on genetic and molecular profiles.

6. Methods

To identify scientific articles that described the application of artificial intelligence to omics data about rare neurological diseases, the Medline database and PubMed were used, with additional searches in Scopus and Web of Science to ensure the thorough coverage of the biomedical literature. A triple combination of keywords related to machine learning ("machine learning", "artificial intelligence"), omics ("genomics", "proteomics", "multiomics"), and rare neurological diseases ("rare neurological disease", "rare neurological disorder") was used to create the search string. Additionally, the names and synonyms for 1840 specific rare neurological diseases were searched in combination with the general terms/keywords of machine learning and omics. These specific rare neurological diseases were identified with the help of Orphanet [91]. Orphanet is a comprehensive database that provides information on rare diseases and orphan drugs to improve the diagnosis, care, and treatment of patients with RDs. It addresses the scarcity and fragmentation of knowledge on RDs by providing multiple levels of classification and nomenclature [92,93]. Only diseases with known point prevalence ("1-5/10,000", "1-9/100,000", "1-9/1,000,000", "1/1,000,000") were included in the search. For most diseases, Orphanet provides PubMed search strings, which were used to construct the search term (for example, "Aneurysm* (subarachnoid hemorrhage [ti] OR subarachnoid hemorrhage[ti] OR subarachnoid hemorrhage[mh]) OR

aneurysmal SAH[tw]" for the disorder acquired aneurysmal subarachnoid hemorrhage). For diseases where no search terms were available from Orphanet, the disorder name was used instead. The inclusion/exclusion criteria were as follows: (1) manuscripts written in English and that included a title and abstract were selected; (2) Orphanet classification was used and only rare neurological diseases with Orpha codes were included for further study; (3) manuscripts involving the use of at least one concrete AI/ML algorithm to handle/explore omics data related to rare neurological diseases were included; (4) reviews and studies on animal models were excluded from the results. The literature search covered articles published from January 2000 to December 2023, allowing us to capture the evolution of AI and omics technologies in the context of rare neurological diseases over the past two decades. We aimed to minimize potential bias by conducting a comprehensive search across multiple databases, including studies with varying outcomes and methodologies, and using systematic and transparent selection criteria. The list of rare neurological disorders, as well as the details of the algorithm and omics data described in this review, can be found in the Supplementary Materials.

7. Discussion and Conclusions

In this review, the scientific literature on ML and omics methods was assessed to explore which artificial intelligence techniques are being utilized to advance the understanding of rare neurological diseases (RNDs) as well as how they are being applied. The most commonly used algorithms were random forest, support vector machines, and artificial neural networks. The most common applications were with regard to biomarker discovery and the diagnosis of rare neurological diseases based on omics data. The majority of the studies gathered in the review were found to be focused on genomics and radiomics. This was expected given that genetic factors are the leading cause of rare diseases and magnetic resonance imaging is the most frequently utilized clinical tool in neuroimaging. The integration of various omics technologies to enhance our understanding of RNDs is illustrated in Figure 1. The random forest algorithm is advantageous as it uses an ensemble of decision trees to lower the variance and reduce overfitting. It is also robust to outliers and requires no feature scaling. Support vector machines are useful in cases where the number of features is more than the number of samples, and the kernel functions associated with SVM can be customized to enhance classification. Artificial neural networks are able to learn and model non-linear complex relationships and they capture new features in the hidden layers that can be instrumental to understanding the molecular details of diseases. Indeed, images derived from medical imaging techniques can be best standardized and processed by deep neural networks.

With the rise of "big data", there is an increasing need to automate tasks that currently require human intervention. In the field of biomedicine, artificial intelligence (AI) techniques have been developed to analyze a wide range of data, from individual omics data and clinical phenotypes to large-scale health databases and multiparametric studies involving large cohorts of patients. Over the past 20 years, machine learning has become a well-established and highly useful discipline. Although there are several learning paradigms available today, machine learning has been successful in various applications, including life sciences and medical research. However, the clinical use of machine learning methods is still relatively rare. AI algorithms have the potential to enhance the diagnosis and understanding of rare neurological diseases by performing mutation detection, prediction, classification, and the identification of disease biomarkers. This can lead to an increase in the number of diagnosed cases and uncover new disease mechanisms and therapeutic targets. However, there is still a need to improve the rate of research and development for rare neurological diseases. While AI has made significant progress in diagnosis, progress in therapy development has been modest. It is known that machine learning plays a significant role in improving treatment by accelerating drug development, predicting the drug's efficacy, optimizing the dosages, and repurposing existing drugs for other diseases [94,95]. With the ever-evolving AI frameworks, it can be assumed that the

promising results obtained so far will soon change the current scenario in the treatment of rare neurological diseases. To diagnose and characterize rare neurological disease patients, AI-based multi-omics integrative approaches are being adopted as genomic data alone are often insufficient. Additionally, novel applications of AI are being explored to develop new research models for RNDs [96]. However, there is still room for improvement in AI-mediated diagnosis, particularly in designing and training models for rare diseases. This is due to various confounding and detrimental factors, such as small patient cohorts and differences in patient ethnicity and gender. The most significant limitation in building predictive models for RNDs is the data collection process [97]. Applying machine learning models to unstructured, poorly standardized, and low-quality control data can adversely affect the model's performance [98]. This is because noise, incompleteness, and sparsity can lead to model overfitting, resulting in high prediction accuracy on training data but low prediction accuracy on new evaluation data. Regarding the limitation of small sample sizes, it must be noted that deep learning models generally require thousands of samples to generalize over the data and achieve robust solutions, while shallow models may still need at least a few hundred samples to build reasonably high-performing models. However, there are several ways to deal with this issue of small sample sizes in machine learning for RNDs. One method is to learn from data from other disorders that are related to the disease being studied or at least share overlapping features. If the heterogeneity is accounted for, one can take into account data regarding the same disease but derived from diverse sources, such as 'multi-omics' data; medical imaging; clinical features; patient registries; open-source databases on genes, proteins, mutations, and drug interactions; and phenotypic data. Data augmentation or the enhancement of the existing data strategy with simulated samples can be considered as well [92,99]. Transfer learning is another option, wherein one can use the knowledge learned by other similar models and fine-tune it to suit the studied domain. Deep learning paradigms have been successful in big data scenarios with large sample sizes, but they often produce models that are difficult to interpret. To enable clinicians to understand the meaning of the classification results, it is necessary to use less complex but explainable models [43,93]. Interpretations of the data derived from explainable models must be uniform across multiple learning algorithms and within the domain or disease being studied. This is possible only when feature extraction and weighting is a stable process and captures the biologically relevant data patterns. Such efforts may strengthen the clinical decision in the small sample regime of RNDs. Additionally, features must be assessed from a biological and statistical standpoint, and robust error analysis must be conducted. Sometimes, routine diagnostic techniques may be insufficient in providing a feature set that can be analyzed by AI to generate results relevant to disease pathogenesis; this may warrant slight modifications in the diagnostic tools. In this report, Dionnet et al. developed a 'minigene' functional assay to identify aberrant splicing in CAPN3, the gene responsible for limb girdle muscular dystrophy [22]. Whole-exome sequencing followed by analysis with AI tools failed to predict the splicing impact for the majority of the deep exonic variants. However, a change in the functional assay that specifically targeted the CAPN3 gene helped to identify 24 variants with AI techniques and seven were clinically important. Although AI predictions can help to solve medical challenges in RNDs, all results must be experimentally validated to confirm their biomedical relevance. One issue that needs to be addressed is the lack of external validation studies for AI models in clinical practice. These studies are crucial in assessing the generalizability and reliability of the algorithm and determining its potential use in clinical settings. However, only a few studies have conducted this type of validation, partly due to the difficulty in obtaining large and diverse datasets and the lack of standardized methods for data collection and analysis. Without adequate validation, there is a risk that the algorithms will produce unreliable or inaccurate results when applied to new datasets or patient cohorts. To solve this problem, collaboration among researchers, clinicians, and data scientists is necessary to develop standardized methods and share data and algorithms to facilitate external validation studies. Furthermore, it is important to note that AI-based applications must be tailored to

the biomedical issue. Biomedical data and the associated challenges are complex, and numerous AI-based algorithms and methods are being improved. Technical limitations and data management and protection must also be carefully considered when designing an AI approach in the medical context. Artificial intelligence and machine learning models show great promise in the identification, diagnosis, treatment, and follow-up of rare neurological diseases. With vast amounts of heterogeneous data now available, ML algorithms can identify patterns and rapidly analyze such data, which would otherwise be incomprehensible to human analysts. While omics-based classifiers assist in the diagnosis of RNDs and help to distinguish between disease mimics, predictive modeling techniques can help to monitor disease progression, thereby allowing for earlier interventions and better treatment planning. From a precision medicine perspective, by identifying biomarkers associated with a particular rare disease, AI algorithms can help to develop personalized treatment plans, helping to improve patient outcomes and reduce the risk of side effects. Rare neurological diseases pose specific challenges such as a limited understanding of the molecular pathophysiology of the disease, small patient groups, and a big data regime—specifically, 'omics' data. AI models should be designed to overcome these challenges and need to be validated through clinical trials and real-world evidence.
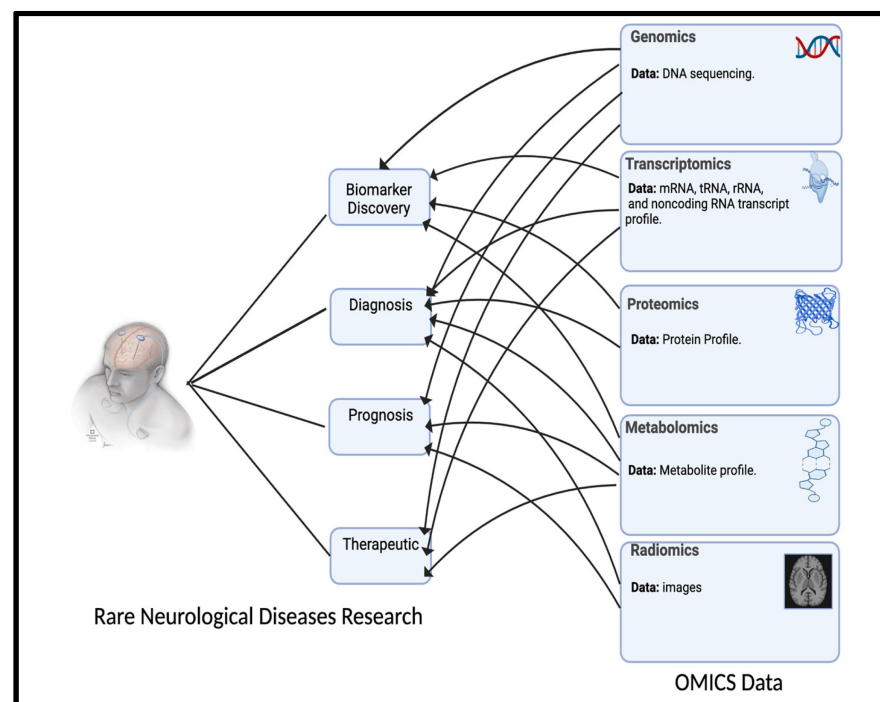

Figure 1. Omics technology for rare neurological disease (RND) research. The figure was drawn using BioRender.com.

The use of AI and omics data in rare disease research raises significant ethical and privacy concerns, including the challenges of obtaining valid consent, protecting confidentiality, and navigating privacy, data protection, and copyright issues [100–102]. Patients and caregivers generally support the use of AI in healthcare research, highlighting the need for transparency and disclosure [103]. Privacy laws like the GDPR in the EU and HIPAA in the US are crucial for patient data protection, yet their application can pose threats to the progress of rare disease research [103,104]. The tension between the potential benefits and risks of AI in healthcare, including privacy concerns, has been underscored [105]. Ethical frameworks have been suggested to address the use and sharing of clinical data for AI applications, advocating for data stewards and the protection of patient privacy [106]. However, the need for further research into these ethical implications, especially in low- and middle-income countries, is paramount [107]. Regulations and governance approaches

need refinement to tackle the ethical challenges posed by AI in rare disease research effectively [104,107]. The issue of equity is also pivotal, with an emphasis on ensuring that AI and omics advancements benefit all populations and do not exacerbate health disparities [107]. The integration of privacy, trust, accountability, responsibility, and bias into the research framework is essential to navigate the complex landscape of AI and omics data in rare disease research [104,107,108].

A major opportunity for further exploration exists in the future of this research, particularly in relation to the emerging role of Large Language Models (LLMs) and knowledge bases in enhancing omics and machine learning research in RNDs. A key objective of future research for rare neurological disorders should be to leverage the full potential of LLMs and knowledge bases through the strategic integration of these two tools in omics and machine learning research. Realizing the transformative impact of these technologies will require the development of robust frameworks for their ethical and effective application.

Finally, artificial intelligence techniques strongly rooted in clinical understanding, the appropriate ethical principles, and sound computational frameworks can help to address the knowledge gap in rare neurological diseases and benefit patients and their families.

References

1. Vos, T.; Lim, S.S.; Abbafati, C.; Abbas, K.M.; Abbasi, M.; Abbasifard, M.; Abbasi-Kangevari, M.; Abbastabar, H.; Abd-Allah, F.; Abdelalim, A.; et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *Lancet* **2020**, *396*, 1204–1222. [CrossRef] [PubMed]
2. Hajat, C.; Stein, E. The global burden of multiple chronic conditions: A narrative review. *Prev. Med. Rep.* **2018**, *12*, 284–293. [CrossRef]
3. Haque, M.; Islam, T.; A Rahman, N.A.; McKimm, J.; Abdullah, A.; Dhingra, S. Strengthening primary health-care services to help prevent and control long-term (chronic) non-communicable diseases in low- and middle-income countries. *Risk Manag. Health Policy* **2020**, *13*, 409–426. [CrossRef]
4. CDC. Health and Economic Costs of Chronic Diseases. Available online: https://www.cdc.gov/chronicdisease/about/costs/index.htm#ref1C (accessed on 6 December 2023).
5. Slebodnik, M. Orphanet: The portal for rare diseases and orphan drugs. *Ref. Rev.* **2009**, *23*, 45–46. [CrossRef]
6. U.S. Food & Drug Administration. Rare Diseases at FDA. Available online: https://www.fda.gov/patients/rare-diseases-fda (accessed on 2 March 2024).
7. Medicines for Rare Diseases—Orphan Drugs. Available online: https://eur-lex.europa.eu/EN/legal-content/summary/medicines-for-rare-diseases-orphan-drugs.html (accessed on 1 December 2023).
8. Richter, T.; Nestler-Parr, S.; Babela, R.; Khan, Z.M.; Tesoro, T.; Molsen, E.; Hughes, D.A. Rare disease terminology and definitions—A systematic global review: Report of the ISPOR rare disease special interest group. *Value Health* **2015**, *18*, 906–914. [CrossRef] [PubMed]

9. Hsu, J.C.; Wu, H.-C.; Feng, W.-C.; Chou, C.-H.; Lai, E.C.-C.; Lu, C.Y. Disease and economic burden for rare diseases in Taiwan: A longitudinal study using Taiwan's national health insurance research database. *PLoS ONE* **2018**, *13*, e0204206. [CrossRef] [PubMed]

10. Nguengang Wakap, S.; Lambert, D.M.; Olry, A.; Rodwell, C.; Gueydan, C.; Lanneau, V.; Murphy, D.; Le Cam, Y.; Rath, A. Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *Eur. J. Hum. Genet.* **2020**, *28*, 165–173. [CrossRef] [PubMed]

11. Yang, G.; Cintina, I.; Pariser, A.; Oehrlein, E.; Sullivan, J.; Kennedy, A. The national economic burden of rare disease in the united states in 2019. *Orphanet J. Rare Dis.* **2022**, *17*, 163. [CrossRef] [PubMed]

12. Tisdale, A.; Cutillo, C.M.; Nathan, R.; Russo, P.; Laraway, B.; Haendel, M.; Nowak, D.; Hasche, C.; Chan, C.-H.; Griese, E.; et al. The IDeaS initiative: Pilot study to assess the impact of rare diseases on patients and healthcare systems. *Orphanet J. Rare Dis.* **2021**, *16*, 429. [CrossRef] [PubMed]

13. Nestler-Parr, S.; Korchagina, D.; Toumi, M.; Pashos, C.L.; Blanchette, C.; Molsen, E.; Morel, T.; Simoens, S.; Kaló, Z.; Gatermann, R.; et al. Challenges in research and health technology assessment of rare disease technologies: Report of the ispor rare disease special interest group. *Value Health* **2018**, *21*, 493–500. [CrossRef]

14. Stoller, J.K. The Challenge of Rare Diseases. *Chest* **2018**, *153*, 1309–1314. [CrossRef] [PubMed]

15. NORD Rare Insights. *Barriers to Rare Disease Diagnosis, Care and Treatment in the US: A 30-Year Comparative Analysis*; RareInsights: Washington, DC, USA, 2020.

16. Ahmed, M.A.; Okour, M.; Brundage, R.; Kartha, R.V. Orphan drug development: The increasing role of clinical pharmacology. *J. Pharmacokinet. Pharmacodyn.* **2019**, *46*, 395–409. [CrossRef] [PubMed]

17. Handfield, R.; Feldstein, J. Insurance companies' perspectives on the orphan drug pipeline. *Am. Health Drug Benefits* **2013**, *6*, 589–598. [PubMed]

18. Althobaiti, H.; Seoane-Vazquez, E.; Brown, L.M.; Fleming, M.L.; Rodriguez-Monguio, R. Disentangling the cost of orphan drugs marketed in the united states. *Healthcare* **2023**, *11*, 558. [CrossRef] [PubMed]

19. Amberger, J.S.; Bocchini, C.A.; Schiettecatte, F.; Scott, A.F.; Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **2015**, *43*, D789–D798. [CrossRef] [PubMed]

20. Qaiser, F.; Sadoway, T.; Yin, Y.; Ali, Q.Z.; Nguyen, C.M.; Shum, N.; Backstrom, I.; Marques, P.T.; Tabarestani, S.; Munhoz, R.P.; et al. Genome sequencing identifies rare tandem repeat expansions and copy number variants in Lennox–Gastaut syndrome. *Brain Commun.* **2021**, *3*, fcab207. [CrossRef] [PubMed]

21. Bao, X.; Li, Q.; Chen, J.; Chen, D.; Ye, C.; Dai, X.; Wang, Y.; Li, X.; Rong, X.; Cheng, F.; et al. Molecular subgroups of intrahepatic cholangiocarcinoma discovered by single-cell RNA sequencing–assisted multiomics analysis. *Cancer Immunol. Res.* **2022**, *10*, 811–828. [CrossRef] [PubMed]

22. Dionnet, E.; Defour, A.; Da Silva, N.; Salvi, A.; Lévy, N.; Krahn, M.; Bartoli, M.; Puppo, F.; Gorokhova, S. Splicing impact of deep exonic missense variants in capn3 explored systematically by minigene functional assay. *Hum. Mutat.* **2020**, *41*, 1797–1810. [CrossRef] [PubMed]

23. Joshi, G.; Jain, A.; Araveeti, S.R.; Adhikari, S.; Garg, H.; Bhandari, M. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape. *Electronics* **2024**, *13*, 498. [CrossRef]

24. Misra, B.B.; Langefeld, C.; Olivier, M.; Cox, L.A. Integrated omics: Tools, advances and future approaches. *J. Mol. Endocrinol.* **2019**, *62*, 21–45. [CrossRef] [PubMed]

25. Sahu, M.; Gupta, R.; Ambasta, R.K.; Kumar, P. Artificial intelligence and machine learning in precision medicine: A paradigm shift in big data analysis. *Prog. Mol. Biol. Transl. Sci.* **2022**, *190*, 57–100. [PubMed]

26. Morales, E.F.; Escalante, H.J. A brief introduction to supervised, unsupervised, and reinforcement learning. In *Biosignal Processing and Classification Using Computational Learning and Intelligence*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 111–129.

27. Jaganathan, K.; Panagiotopoulou, S.K.; McRae, J.F.; Darbandi, S.F.; Knowles, D.; Li, Y.I.; Kosmicki, J.A.; Arbelaez, J.; Cui, W.; Schwartz, G.B.; et al. Predicting splicing from primary sequence with deep learning. *Cell* **2019**, *176*, 535–548. [CrossRef] [PubMed]

28. Carter, H.; Douville, C.; Stenson, P.D.; Cooper, D.N.; Karchin, R. Identifying mendelian disease genes with the variant effect scoring tool. *BMC Genom.* **2013**, *14*, S3. [CrossRef] [PubMed]

29. Yang, J.-F.; Ding, X.-F.; Chen, L.; Mat, W.-K.; Xu, M.Z.; Chen, J.-F.; Wang, J.-M.; Xu, L.; Poon, W.-S.; Kwong, A.; et al. Copy number variation analysis based on AluScan sequences. *J. Clin. Bioinform.* **2014**, *4*, 15. [CrossRef] [PubMed]

30. López-Rivera, J.J.; Rodríguez-Salazar, L.; Soto-Ospina, A.; Estrada-Serrato, C.; Serrano, D.; Chaparro-Solano, H.M.; Londoño, O.; Rueda, P.A.; Ardila, G.; Villegas-Lanau, A.; et al. Structural protein effects underpinning cognitive developmental delay of the pura p. phe233del mutation modelled by artificial intelligence and the hybrid quantum mechanics–molecular mechanics framework. *Brain Sci.* **2022**, *12*, 871. [CrossRef] [PubMed]

31. Chen, Z.; Zheng, Y.; Yang, Y.; Huang, Y.; Zhao, S.; Zhao, H.; Yu, C.; Dong, X.; Zhang, Y.; Wang, L.; et al. PhenoApt leverages clinical expertise to prioritize candidate genes via machine learning. *Am. J. Hum. Genet.* **2022**, *109*, 270–281. [CrossRef] [PubMed]

32. Quinodoz, M.; Royer-Bertrand, B.; Cisarova, K.; Di Gioia, S.A.; Superti-Furga, A.; Rivolta, C. DOMINO: Using machine learning to predict genes associated with dominant disorders. *Am. J. Hum. Genet.* **2017**, *101*, 623–629. [CrossRef] [PubMed]

33. Lam, S.; Arif, M.; Song, X.; Uhlén, M.; Mardinoglu, A. Machine learning analysis reveals biomarkers for the detection of neurological diseases. *Front. Mol. Neurosci.* **2022**, *15*, 889728. [CrossRef] [PubMed]

34. Fang, Y.; Gao, J.; Guo, Y.; Li, X.; Yuan, E.; Yuan, E.; Song, L.; Shi, Q.; Yu, H.; Zhao, D.; et al. Allelic phenotype prediction of phenylketonuria based on the machine learning method. *Hum. Genom.* **2023**, *17*, 34. [CrossRef] [PubMed]

35. Pirhaji, L.; Milani, P.; Leidl, M.; Curran, T.; Avila-Pacheco, J.; Clish, C.B.; White, F.M.; Saghatelian, A.; Fraenkel, E. Revealing disease-associated pathways by network integration of untargeted metabolomics. *Nat. Methods* **2016**, *13*, 770–776. [CrossRef] [PubMed]

36. Trevino, A.E.; Müller, F.; Andersen, J.; Sundaram, L.; Kathiria, A.; Shcherbina, A.; Farh, K.; Chang, H.Y.; Pașca, A.M.; Kundaje, A.; et al. Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **2021**, *184*, 5053–5069. [CrossRef] [PubMed]

37. Willscher, E.; Hopp, L.; Kreuz, M.; Schmidt, M.; Hakobyan, S.; Arakelyan, A.; Hentschel, B.; Jones, D.T.W.; Pfister, S.M.; Loeffler, M.; et al. High-resolution cartography of the transcriptome and methylome landscapes of diffuse gliomas. *Cancers* **2021**, *13*, 3198. [CrossRef] [PubMed]

38. Loeffler-Wirth, H.; Hopp, L.; Schmidt, M.; Zakharyan, R.; Arakelyan, A.; Binder, H. The transcriptome and methylome of the developing and aging brain and their relations to gliomas and psychological disorders. *Cells* **2022**, *11*, 362. [CrossRef] [PubMed]

39. Huang, W.-K.; Wong, S.Z.H.; Pather, S.R.; Nguyen, P.T.; Zhang, F.; Zhang, D.Y.; Zhang, Z.; Lu, L.; Fang, W.; Chen, L.; et al. Generation of hypothalamic arcuate organoids from human induced pluripotent stem cells. *Cell Stem Cell* **2021**, *28*, 1657–1670.e10. [CrossRef] [PubMed]

40. Agus, F.; Crespo, D.; Myers, R.H.; Labadorf, A. The caudate nucleus undergoes dramatic and unique transcriptional changes in human prodromal Huntington's disease brain. *BMC Med. Genom.* **2019**, *12*, 137. [CrossRef] [PubMed]

41. Chiocchetti, A.; Haslinger, D.; Stein, J.; La Torre-Ubieta, L.; Cocchi, E.; Rothämel, T.; Lindlar, S.; Waltes, R.; Fulda, S.; Geschwind, D.; et al. Transcriptomic signatures of neuronal differentiation and their association with risk genes for autism spectrum and related neuropsychiatric disorders. *Transl. Psychiatry* **2016**, *6*, 864. [CrossRef] [PubMed]

42. Vaz-Drago, R.; Custódio, N.; Carmo-Fonseca, M. Deep intronic mutations and human disease. *Hum. Genet.* **2017**, *136*, 1093–1111. [CrossRef] [PubMed]

43. Zuallaert, J.; Godin, F.; Kim, M.; Soete, A.; Saeys, Y.; De Neve, W. SpliceRover: Interpretable convolutional neural networks for improved splice site prediction. *Bioinformatics* **2018**, *34*, 4180–4188. [CrossRef] [PubMed]

44. Hiraide, T.; Shimizu, K.; Okumura, Y.; Miyamoto, S.; Nakashima, M.; Ogata, T.; Saitsu, H. A deep intronic *TCTN2* variant activating a cryptic exon predicted by SpliceRover in a patient with Joubert syndrome. *J. Hum. Genet.* **2023**, *68*, 499–505. [CrossRef] [PubMed]

45. Louis, D.N.; Ohgaki, H.; Wiestler, O.D.; Cavenee, W.K.; Burger, P.C.; Jouvet, A.; Scheithauer, B.W.; Kleihues, P. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* **2007**, *114*, 97–109. [CrossRef]

46. Giunti, L.; Pantaleo, M.; Sardi, I.; Provenzano, A.; Magi, A.; Cardellicchio, S.; Castiglione, F.; Tattini, L.; Novara, F.; Buccoliero, A.M.; et al. Genome-wide copy number analysis in pediatric glioblastoma multiforme. *Am. J. Cancer Res.* **2014**, *4*, 293–303. [PubMed]

47. Ma, J.; Hong, Y.; Chen, W.; Li, D.; Tian, K.; Wang, K.; Yang, Y.; Zhang, Y.; Chen, Y.; Song, L.; et al. High copy-number variation burdens in cranial meningiomas from patients with diverse clinical phenotypes characterized by hot genomic structure changes. *Front. Oncol.* **2020**, *10*, 1382. [CrossRef]

48. Mirchia, K.; Sathe, A.A.; Walker, J.M.; Fudym, Y.; Galbraith, K.; Viapiano, M.S.; Corona, R.J.; Snuderl, M.; Xing, C.; Hatanpaa, K.J.; et al. Total copy number variation as a prognostic factor in adult astrocytoma subtypes. *Acta Neuropathol. Commun.* **2019**, *7*, 92. [CrossRef] [PubMed]

49. Park, H.; Chun, S.-M.; Shim, J.; Oh, J.-H.; Cho, E.J.; Hwang, H.S.; Lee, J.-Y.; Kim, D.; Jang, S.J.; Nam, S.J.; et al. Detection of chromosome structural variation by targeted next-generation sequencing and a deep learning application. *Sci. Rep.* **2019**, *9*, 3644. [CrossRef] [PubMed]

50. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with alphafold. *Nature* **2021**, *596*, 583–589. [CrossRef]

51. Dickson, D.W.; Baker, M.C.; Jackson, J.L.; DeJesus-Hernandez, M.; Finch, N.A.; Tian, S.; Heckman, M.G.; Pottier, C.; Gendron, T.F.; Murray, M.E.; et al. Extensive transcriptomic study emphasizes importance of vesicular transport in *C9orf72* expansion carriers. *Acta Neuropathol. Commun.* **2019**, *7*, 150. [CrossRef] [PubMed]

52. Choi, S.B.; Park, J.S.; Chung, J.W.; Yoo, T.K.; Kim, D.W. Multicategory classification of 11 neuromuscular diseases based on microarray data using support vector machine. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 3460–3463.

53. Caputo, V.; Megalizzi, D.; Fabrizio, C.; Termine, A.; Colantoni, L.; Bax, C.; Gimenez, J.; Monforte, M.; Tasca, G.; Ricci, E.; et al. D4z4 methylation levels combined with a machine learning pipeline highlight single CpG sites as discriminating biomarkers for fshd patients. *Cells* **2022**, *11*, 4114. [CrossRef] [PubMed]

54. Dabin, L.C.; Guntoro, F.; Campbell, T.; Bélicard, T.; Smith, A.R.; Smith, R.G.; Raybould, R.; Schott, J.M.; Lunnon, K.; Sarkies, P.; et al. Altered DNA methylation profiles in blood from patients with sporadic creutzfeldt–jakob disease. *Acta Neuropathol.* **2020**, *140*, 863–879. [CrossRef] [PubMed]

55. Jabari, S.; Kobow, K.; Pieper, T.; Hartlieb, T.; Kudernatsch, M.; Polster, T.; Bien, C.G.; Kalbhenn, T.; Simon, M.; Hamer, H.; et al. DNA methylation-based classification of malformations of cortical development in the human brain. *Acta Neuropathol.* **2022**, *143*, 93–104. [CrossRef]

56. Bjornevik, K.; Zhang, Z.; O'Reilly, É.J.; Berry, J.D.; Clish, C.B.; Deik, A.; Jeanfavre, S.; Kato, I.; Kelly, R.S.; Kolonel, L.N.; et al. Prediagnostic plasma metabolomics and the risk of amyotrophic lateral sclerosis. *Neurology* **2019**, *92*, 2089–2100. [CrossRef] [PubMed]

57. Lawton, K.A.; Brown, M.V.; Alexander, D.; Li, Z.; Wulff, J.E.; Lawson, R.; Jaffa, M.; Milburn, M.V.; Ryals, J.A.; Bowser, R.; et al. Plasma metabolomic biomarker panel to distinguish patients with amyotrophic lateral sclerosis from disease mimics. *Amyotroph. Lateral Scler. Front. Degener.* **2014**, *15*, 362–370. [CrossRef] [PubMed]

58. Thistlethwaite, L.R.; Li, X.; Burrage, L.C.; Riehle, K.; Hacia, J.G.; Braverman, N.; Wangler, M.F.; Miller, M.J.; Elsea, S.H.; Milosavljevic, A. Clinical diagnosis of metabolic disorders using untargeted metabolomic profiling and disease-specific networks learned from profiling data. *Sci. Rep.* **2022**, *12*, 6556. [CrossRef] [PubMed]

59. Zhao, B.; Wang, Y.; Ma, W. Molecular landscape of IDH-mutant astrocytoma and oligodendroglioma grade 2 indicate tumor purity as an underlying genomic factor. *Mol. Med.* **2022**, *28*, 34. [CrossRef] [PubMed]

60. Capper, D.; Jones, D.T.W.; Sill, M.; Hovestadt, V.; Schrimpf, D.; Sturm, D.; Koelsche, C.; Sahm, F.; Chavez, L.; Reuss, D.E.; et al. DNA methylation-based classification of central nervous system tumours. *Nature* **2018**, *555*, 469–474. [CrossRef] [PubMed]

61. Ranjith, G.; Parvathy, R.; Vikas, V.; Chandrasekharan, K.; Nair, S. Machine learning methods for the classification of gliomas: Initial results using features extracted from MR spectroscopy. *Neuroradiol. J.* **2015**, *28*, 106–111. [CrossRef] [PubMed]

62. Chen, B.; Chen, C.; Zhang, Y.; Huang, Z.; Wang, H.; Li, R.; Xu, J. Differentiation between germinoma and craniopharyngioma using radiomics-based machine learning. *J. Pers. Med.* **2022**, *12*, 45. [CrossRef]

63. Wang, C.; You, L.; Zhang, X.; Zhu, Y.; Zheng, L.; Huang, W.; Guo, D.; Dong, Y. A radiomics-based study for differentiating parasellar cavernous hemangiomas from meningiomas. *Sci. Rep.* **2022**, *12*, 15509. [CrossRef] [PubMed]

64. Zhang, B.; Chang, K.; Ramkissoon, S.; Tanguturi, S.; Bi, W.L.; Reardon, D.A.; Ligon, K.L.; Alexander, B.M.; Wen, P.Y.; Huang, R.Y. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. *Neuro-Oncology* **2017**, *19*, 109–117. [CrossRef]

65. Kandalgaonkar, P.; Sahu, A.; Saju, A.C.; Joshi, A.; Mahajan, A.; Thakur, M.; Sahay, A.; Epari, S.; Sinha, S.; Dasgupta, A.; et al. Predicting IDH sub-type of grade 4 astrocytoma and glioblastoma from tumor radiomic patterns extracted from multiparametric magnetic resonance images using a machine learning approach. *Front. Oncol.* **2022**, *12*, 879376. [CrossRef] [PubMed]

66. Wei, J.S.; Greer, B.T.; Westermann, F.; Steinberg, S.M.; Son, C.-G.; Chen, Q.-R.; Whiteford, C.C.; Bilke, S.; Krasnoselsky, A.L.; Cenacchi, N.; et al. Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. *Cancer Res.* **2004**, *64*, 6883–6891. [CrossRef]

67. Tranchevent, L.-C.; Azuaje, F.; Rajapakse, J.C. A deep neural network approach to predicting clinical outcomes of neuroblastoma patients. *BMC Med. Genom.* **2019**, *12*, 178. [CrossRef] [PubMed]

68. Francescatto, M.; Chierici, M.; Rezvan Dezfooli, S.; Zandonà, A.; Jurman, G.; Furlanello, C. Multi-omics integration for neuroblastoma clinical endpoint prediction. *Biol. Direct* **2018**, *13*, 5. [CrossRef] [PubMed]

69. Sugino, R.P.; Ohira, M.; Mansai, S.P.; Kamijo, T. Comparative epigenomics by machine learning approach for neuroblastoma. *BMC Genom.* **2022**, *23*, 852. [CrossRef] [PubMed]

70. Giwa, A.; Rossouw, S.C.; Fatai, A.; Gamieldien, J.; Christoffels, A.; Bendou, H. Predicting amplification of MYCN using CpG methylation biomarkers in neuroblastoma. *Future Oncol.* **2021**, *17*, 4769–4783. [CrossRef] [PubMed]

71. Wang, C.; Lue, W.; Kaalia, R.; Kumar, P.; Rajapakse, J.C. Network-based integration of multi-omics data for clinical outcome prediction in neuroblastoma. *Sci. Rep.* **2022**, *12*, 15425. [CrossRef] [PubMed]

72. Oyefiade, A.; Erdman, L.; Goldenberg, A.; Malkin, D.; Bouffet, E.; Taylor, M.D.; Ramaswamy, V.; Scantlebury, N.; Law, N.; Mabbott, D.J. PPAR and GST polymorphisms may predict changes in intellectual functioning in medulloblastoma survivors. *J. Neuro-Oncol.* **2019**, *142*, 39–48. [CrossRef] [PubMed]

73. Mihaylov, I.; Kańduła, M.; Krachunov, M.; Vassilev, D. A novel framework for horizontal and vertical data integration in cancer studies with application to survival time prediction models. *Biol. Direct* **2019**, *14*, 22. [CrossRef]

74. Bratulic, S.; Limeta, A.; Dabestani, S.; Birgisson, H.; Enblad, G.; Stålberg, K.; Hesselager, G.; Häggman, M.; Höglund, M.; Simonson, O.E.; et al. Noninvasive detection of any-stage cancer using free glycosaminoglycans. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2115328119. [CrossRef] [PubMed]

75. Turner, M.R.; Al-Chalabi, A.; Chio, A.; Hardiman, O.; Kiernan, M.C.; Rohrer, J.D.; Rowe, J.; Seeley, W.; Talbot, K. Genetic screening in sporadic ALS and FTD. *J. Neurol. Neurosurg. Psychiatry* **2017**, *88*, 1042–1044. [CrossRef] [PubMed]

76. Placek, K.; Benatar, M.; Wuu, J.; Rampersaud, E.; Hennessy, L.; Van Deerlin, V.M.; Grossman, M.; Irwin, D.J.; Elman, L.; McCluskey, L.; et al. Machine learning suggests polygenic risk for cognitive dysfunction in amyotrophic lateral sclerosis. *EMBO Mol. Med.* **2021**, *13*, e12595. [CrossRef] [PubMed]

77. Ji, X.; Pei, Q.; Zhang, J.; Lin, P.; Li, B.; Yin, H.; Sun, J.; Su, D.; Qu, X.; Yin, D. Single-cell sequencing combined with machine learning reveals the mechanism of interaction between epilepsy and stress cardiomyopathy. *Front. Immunol.* **2023**, *14*, 1078731. [CrossRef]

78. Stapleton, C.J.; Acharjee, A.; Irvine, H.J.; Wolcott, Z.C.; Patel, A.B.; Kimberly, W.T. High-throughput metabolite profiling: Identification of plasma taurine as a potential biomarker of functional outcome after aneurysmal subarachnoid hemorrhage. *J. Neurosurg.* **2019**, *133*, 1842–1849. [CrossRef] [PubMed]

79. Hu, L.S.; Ning, S.; Eschbacher, J.M.; Gaw, N.; Dueck, A.C.; Smith, K.A.; Nakaji, P.; Plasencia, J.; Ranjbar, S.; Price, S.J.; et al. Multi-parametric mri and texture analysis to visualize spatial histologic heterogeneity and tumor extent in glioblastoma. *PLoS ONE* **2015**, *10*, e0141506. [CrossRef] [PubMed]

80. Bijari, S.; Jahanbakhshi, A.; Hajishafiezahramini, P.; Abdolmaleki, P. Differentiating glioblastoma multiforme from brain metastases using multidimensional radiomics features derived from MRI and multiple machine learning models. *BioMed Res. Int.* **2022**, *2022*, 2016006. [CrossRef] [PubMed]

81. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.

82. Hostin, M.-A.; Ogier, A.C.; Michel, C.P.; Le Fur, Y.; Guye, M.; Attarian, S.; Fortanier, E.; Bellemare, M.-E.; Bendahan, D. The impact of fatty infiltration on MRI segmentation of lower limb muscles in neuromuscular diseases: A comparative study of deep learning approaches. *J. Magn. Reson. Imaging* **2023**, *58*, 1826–1835. [CrossRef] [PubMed]

83. Shen, M.W.; Arbab, M.; Hsu, J.Y.; Worstell, D.; Culbertson, S.J.; Krabbe, O.; Cassa, C.A.; Liu, D.R.; Gifford, D.K.; Sherwood, R.I. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **2018**, *563*, 646–651. [CrossRef]

84. Nishida, A.; Kataoka, N.; Takeshima, Y.; Yagi, M.; Awano, H.; Ota, M.; Itoh, K.; Hagiwara, M.; Matsuo, M. Chemical treatment enhances skipping of a mutated exon in the dystrophin gene. *Nat. Commun.* **2011**, *2*, 308. [CrossRef] [PubMed]

85. Lin, D.; Zhao, W.; Yang, J.; Wang, H.; Zhang, H. Integrative Analysis of Biomarkers and Mechanisms in Adamantinomatous Craniopharyngioma. *Front. Genet.* **2022**, *13*, 830793. [CrossRef] [PubMed]

86. Lian, H.; Han, Y.-P.; Zhang, Y.-C.; Zhao, Y.; Yan, S.; Li, Q.-F.; Wang, B.-C.; Wang, J.-J.; Meng, W.; Yang, J.; et al. Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma. *Mol. Oncol.* **2019**, *13*, 2227–2245. [CrossRef]

87. Gilard, V.; Ferey, J.; Marguet, F.; Fontanilles, M.; Ducatez, F.; Pilon, C.; Lesueur, C.; Pereira, T.; Basset, C.; Schmitz-Afonso, I.; et al. Integrative metabolomics reveals deep tissue and systemic metabolic remodeling in glioblastoma. *Cancers* **2021**, *13*, 5157. [CrossRef] [PubMed]

88. De Jong, J.; Cutcutache, I.; Page, M.; Elmoufti, S.; Dilley, C.; Fröhlich, H.; Armstrong, M. Towards realizing the vision of precision medicine: AI based prediction of clinical drug response. *Brain A J. Neurol.* **2021**, *144*, 1738–1750. [CrossRef] [PubMed]

89. Dahlin, M.; Singleton, S.S.; David, J.A.; Basuchoudhary, A.; Wickström, R.; Mazumder, R.; Prast-Nielsen, S. Higher levels of bifidobacteria and tumor necrosis factor in children with drug-resistant epilepsy are associated with anti-seizure response to the ketogenic diet. *EBioMedicine* **2022**, *80*, 104061. [CrossRef]

90. Kurkiewicz, A.; Cooper, A.; McIlwaine, E.; Cumming, S.A.; Adam, B.; Krahe, R.; Puymirat, J.; Schoser, B.; Timchenko, L.; Ashizawa, T.; et al. Towards development of a statistical framework to evaluate myotonic dystrophy type 1 mRNA biomarkers in the context of a clinical trial. *PLoS ONE* **2020**, *15*, e0231000. [CrossRef] [PubMed]

91. Weinreich, S.S.; Mangon, R.; Sikkens, J.J.; Teeuw, M.E.E.; Cornel, M.C. [Orphanet: A European database for rare diseases]. *Ned. Tijdschr. Geneeskd.* **2008**, *152*, 518–519. [PubMed]

92. Kebaili, A.; Lapuyade-Lahorgue, J.; Ruan, S. Deep learning approaches for data augmentation in medical imaging: A review. *J. Imaging* **2023**, *9*, 81. [CrossRef] [PubMed]

93. Garbulowski, M.; Smolinska, K.; Diamanti, K.; Pan, G.; Maqbool, K.; Feuk, L.; Komorowski, J. Interpretable machine learning reveals dissimilarities between subtypes of autism spectrum disorder. *Front. Genet.* **2021**, *12*, 618277. [CrossRef] [PubMed]

94. Dara, S.; Dhamercherla, S.; Jadav, S.S.; Babu, C.H.M.; Ahsan, M.J. Machine Learning in Drug Discovery: A Review. *Artif. Intell. Rev.* **2022**, *55*, 1947–1999. [CrossRef] [PubMed]

95. Patel, L.; Shukla, T.; Huang, X.; Ussery, D.W.; Wang, S. Machine Learning Methods in Drug Discovery. *Molecules* **2020**, *25*, 5277. [CrossRef]

96. Vatansever, S.; Schlessinger, A.; Wacker, D.; Kaniskan, H.Ü.; Jin, J.; Zhou, M.-M.; Zhang, B. Artificial intelligence and machine learning-aided drug discov- ery in central nervous system diseases: State-of-the-arts and future directions. *Med. Res. Rev.* **2021**, *41*, 1427–1473. [CrossRef] [PubMed]

97. Mitani, A.A.; Haneuse, S. Small data challenges of studying rare diseases. *JAMA Netw. Open* **2020**, *3*, e201965. [CrossRef] [PubMed]

98. Li, R.; Li, L.; Xu, Y.; Yang, J. Machine learning meets omics: Applications and perspectives. *Brief. Bioinform.* **2022**, *23*, bbab460. [CrossRef]

99. Marouf, M.; Machart, P.; Bansal, V.; Kilian, C.; Magruder, D.S.; Krebs, C.F.; Bonn, S. Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks. *Nat. Commun.* **2020**, *11*, 166. [CrossRef] [PubMed]

100. Hallowell, N.; Parker, M.; Nellåker, C. Big data phenotyping in rare diseases: Some ethical issues. *Anesth. Analg.* **2019**, *21*, 272–274. [CrossRef] [PubMed]

101. Austin, C.P.; Cutillo, C.M.; Lau, L.P.; Jonker, A.H.; Rath, A.; Julkowska, D.; Thomson, D.; Terry, S.F.; de Montleau, B.; Ardigò, D.; et al. Future of rare diseases research 2017–2027: An IRDiRC perspective. *Clin. Transl. Sci.* **2018**, *11*, 21–27. [CrossRef] [PubMed]
102. Price, W.N.; Cohen, I.G. Privacy in the age of medical big data. *Nat. Med.* **2019**, *25*, 37–43. [CrossRef]
103. McCradden, M.D.; Baba, A.; Saha, A.; Ahmad, S.; Boparai, K.; Fadaiefard, P.; Cusimano, M.D. Ethical concerns around use of artificial intelligence in health care research from the perspective of patients with meningioma, caregivers and health care providers: A qualitative study. *Can. Med. Assoc. Open Access J.* **2020**, *8*, E90–E95. [CrossRef] [PubMed]
104. Blasimme, A.; Vayena, E. The ethics of AI in biomedical research, patient care, and public health. In *The Oxford Handbook of Ethics of AI*; Oxford Academic: Oxford, UK, 2020.
105. Williamson, S.M.; Prybutok, V. Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. *Appl. Sci.* **2024**, *14*, 675. [CrossRef]
106. Larson, D.B.; Magnus, D.C.; Lungren, M.P.; Shah, N.H.; Langlotz, C.P. Ethics of using and sharing clinical imaging data for artificial intelligence: A proposed framework. *Radiology* **2020**, *295*, 675–682. [CrossRef]
107. Murphy, K.; Di Ruggiero, E.; Upshur, R.; Willison, D.J.; Malhotra, N.; Cai, J.C.; Malhotra, N.; Lui, V.; Gibson, J. Artificial intelligence for good health: A scoping review of the ethics literature. *BMC Med. Ethics* **2021**, *22*, 14. [CrossRef] [PubMed]
108. Bartoletti, I. AI in healthcare: Ethical and privacy challenges. In *Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, 26–29 June 2019*; Proceedings 17; Springer International Publishing: Berlin/Heidelberg, Germany, 2019.