



Review

Artificial Intelligence in Rare Diseases: Workflow-Integrated Precision Kidney Care

Charat Thongprayoon¹, Francesco Pesce^{2,3} and Wisit Cheungpasitporn^{1,*}

¹ Division of Nephrology and Hypertension, Department of Medicine, Mayo Clinic, Rochester, MN 55905, USA; charat.thongprayoon@gmail.com

² Division of Renal Medicine, Ospedale Isola Tiberina-Gemelli Isola, 00168 Rome, Italy; francesco.pesce2@unicatt.it

³ Department of Translational Medicine and Surgery, Università Cattolica del Sacro Cuore, 00168 Rome, Italy

* Correspondence: wcheungpasitporn@gmail.com

Abstract

Rare diseases affect over 300 million individuals worldwide yet remain underdiagnosed and poorly characterized due to fragmented data, small cohorts, and phenotypic heterogeneity. Advances in artificial intelligence (AI) are enabling integration of genomics, imaging, electronic health records, and patient-generated data to support diagnosis, phenotyping, prognosis, and therapeutic discovery. In kidney care, these capabilities are reflected in tools for genomic variant prioritization, AI-assisted histopathology, and integrated risk stratification models for rare and complex kidney diseases. This review synthesizes current AI applications across the rare disease continuum and proposes a clinically grounded framework to distinguish exploratory models from systems that are methodologically robust and operationally deployable. We highlight advances that address data sparsity and heterogeneity, alongside persistent challenges in validation, generalizability, equity, and workflow integration. Finally, we outline future directions, including federated learning, digital twins, and AI-driven clinical decision agents, as pathways toward precision-guided, workflow-integrated rare disease care.

Keywords: rare diseases; artificial intelligence; genomic diagnosis; clinical decision support; nephrology

1. The Rare Disease Paradox and the Diagnostic Odyssey

Rare diseases are typically defined based on low prevalence thresholds, although definitions vary by region [1–4]. The European Medicines Agency defines a rare disease as one affecting fewer than 5 in 10,000 individuals, while broader international frameworks increasingly recognize the combination of low prevalence, clinical complexity, and substantial unmet medical need [1–4]. Despite their individual rarity, rare diseases collectively affect more than 300 million individuals worldwide, representing a substantial but often underrecognized global health burden [1–4].

Patients frequently endure years of sequential referrals, misdiagnoses, and delayed genetic testing before receiving a definitive diagnosis [5–8]. This prolonged diagnostic odyssey, often spanning several years, reflects a systemic diagnostic delay marked by fragmented care pathways, repeated incorrect diagnoses, and delayed access to definitive genomic evaluation. Importantly, this delay is not merely inefficient but clinically harmful, as it postpones disease-modifying interventions and compounds psychological and financial burdens [9–12].



Academic Editors: Giovanni Costantini and Daniele Giansanti

Received: 13 April 2026

Revised: 23 May 2026

Accepted: 25 May 2026

Published: 27 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

Traditional statistical and machine-learning paradigms struggle in this setting [13,14]. Rare diseases violate core assumptions of population-level modeling due to small sample sizes, heterogeneous phenotypes, and long disease trajectories [15–19]. This creates a fundamental mismatch between conventional evidence-generation frameworks, which prioritize population-level inference, and the realities of rare disease care, which are inherently individualized. Given the inherent data sparsity and heterogeneity of rare diseases, multiple complementary AI paradigms have emerged, each addressing distinct methodological constraints (Table 1) [13,16,20–26]. Consequently, clinical decision-making in rare diseases often incorporates individualized N-of-1 reasoning. AI systems designed for this domain should therefore move beyond population-level prediction toward individualized, mechanism-informed reasoning, with a particular emphasis on compressing the diagnostic timeline described above. In this context, AI-enabled genomic diagnostics represent a key application, offering the potential to reduce selected components of diagnostic delay from months or years to shorter clinically meaningful intervals in optimized settings [27–31].

Table 1. Core Artificial Intelligence Approaches in Rare Disease Care: Mechanisms, Strengths, Limitations, and Clinical Applications. Representative references are provided for scientific transparency and illustrative context.

| Approach | Core Concept | Why It Is Relevant in Rare Diseases | Typical Data Sources | Key Strength | Major Limitation | Representative Clinical Applications |
|--------------------------------------|---|---|----------------------------------|---|--|---|
| Transfer Learning [27] | Reuse knowledge learned from large datasets | Compensates for limited rare disease sample sizes | Imaging, genomics | Strong baseline performance with limited local data | Risk of domain mismatch and reduced generalizability | Rare genetic disorders; rare cancers |
| Federated Learning [31] | Train models across institutions without sharing raw patient data | Enables multi-center learning while preserving privacy | EHRs, registries | Privacy-preserving scalability | Data heterogeneity and infrastructure complexity | International registries; pediatric rare diseases |
| Graph Neural Networks (GNNs) [30] | Learn biological relationships within networks and pathways | Captures complex molecular interactions despite small cohorts | Multi-omics, biological networks | Effective with limited sample sizes | Complex graph construction and interpretation | Gene discovery; drug repurposing |
| Few-shot/ Low-resource Learning [27] | Learn from minimal labeled examples | Designed for extreme data scarcity | Imaging, EHRs | Rapid adaptation to small datasets | Sensitivity to noisy or biased labels | Rare phenotype classification |
| Causal Inference Frameworks [28] | Model potential cause-and-effect relationships | Supports inference when randomized trials are impractical | Longitudinal EHR data | Mechanistic and clinically interpretable insights | Susceptibility to confounding and modeling assumptions | Treatment-effect estimation; drug repurposing |
| Explainable AI (XAI) [29] | Improve interpretability of model outputs | Enhances clinician trust and auditability | Multimodal | Improved transparency and regulatory readiness | Potential tradeoff with predictive performance | Diagnostics; genomics |

Table 1. Cont.

| Approach | Core Concept | Why It Is Relevant in Rare Diseases | Typical Data Sources | Key Strength | Major Limitation | Representative Clinical Applications |
|--------------------------------------|---|---|----------------------|---|---|--|
| Reinforcement Learning (RL) [28] | Optimize sequential clinical decisions over time | Models dynamic disease trajectories and treatment adaptation | Time-series EHR data | Personalized treatment strategies | Requires dense longitudinal data | Treatment sequencing |
| Foundation/Multimodal Models [30,31] | Integrate text, imaging, genomic, and clinical data | Handles fragmented and heterogeneous rare disease information | Multimodal | Cross-domain integration and generalization | Bias, interpretability, and validation challenges | Diagnostic support; phenotype extraction |

Abbreviations: AI, artificial intelligence; EHR, electronic health record; GNNs, graph neural networks; RL, reinforcement learning; XAI, explainable artificial intelligence.

This manuscript is structured as a narrative review aimed at synthesizing current and emerging applications of artificial intelligence in rare diseases, with a focus on clinical translation in kidney care. A targeted literature search was conducted using PubMed, Embase, and Google Scholar to identify relevant publications from January 2015 through January 2026. Search terms included combinations of ‘rare disease,’ ‘artificial intelligence,’ ‘machine learning,’ ‘genomics,’ ‘federated learning,’ ‘digital twins,’ ‘clinical decision support,’ and ‘precision medicine.’ Additional references were identified through citation tracking of key articles.

Study selection was guided by relevance to rare disease diagnosis, phenotyping, prognosis, and therapeutic decision-making, with emphasis on methodological rigor, translational applicability, and representation of key AI paradigms. Priority was given to peer-reviewed original studies, high-impact reviews, and representative methodological papers that illustrate clinically meaningful applications or emerging frameworks. Given the scope of the review, formal inclusion and exclusion criteria were not applied in a systematic manner; instead, selection was purposive to ensure coverage of both established and evolving approaches relevant to rare disease care.

Evidence was synthesized using a concept-driven approach, integrating findings across domains including genomics, federated learning, and digital twin modeling to construct a clinically grounded evaluative framework. Emerging areas, such as agentic AI systems, are included to provide forward-looking context and are explicitly distinguished from empirically validated applications. While this approach allows for flexible integration of a rapidly evolving literature, it may introduce selection bias and does not provide a comprehensive systematic appraisal of all available studies.

2. AI-Accelerated Genomic Diagnosis: Compressing Diagnostic Timelines

2.1. Core Concept

As outlined in Section 1, patients with rare diseases often experience prolonged diagnostic delays spanning several years, driven by fragmented clinical data, phenotypic heterogeneity, and limited access to specialized expertise. AI-enabled genomic diagnostics directly address this challenge by restructuring the diagnostic process, shifting from sequential, trial-and-error evaluation toward integrated, data-driven prioritization. Central to this approach is the use of structured phenotype representations. Among available ontologies, the Human Phenotype Ontology (HPO) is prioritized because it provides a

granular, hierarchical, and computationally tractable framework specifically designed to capture phenotypic abnormalities and enable genotype–phenotype matching [6,31–33].

Unlike broader clinical terminologies such as the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED CT), which are optimized for clinical documentation and interoperability, or curated gene–disease resources such as Online Mendelian Inheritance in Man (OMIM), which primarily catalog genetic associations, HPO is structured to support phenotype-driven diagnostic reasoning. Its standardized vocabulary and hierarchical relationships allow efficient encoding of patient-specific features and facilitate algorithmic comparison with known disease phenotypes, making it particularly well suited for AI-enabled variant prioritization in rare disease contexts.

AI-enabled variant prioritization, when coupled with deep phenotype matching using standardized ontologies such as the Human Phenotype Ontology (HPO), has the potential to improve diagnostic efficiency and shorten components of the diagnostic process. Rather than representing solely incremental efficiency gains, these approaches may support more rapid and hypothesis-driven diagnostic workflows in selected clinical settings. The clinical objective is to reduce diagnostic delays while maintaining clinical rigor and appropriate expert oversight [34].

2.2. Architecture of a High-Sensitivity Diagnostic Pipeline

A clinically viable pipeline integrates multiple AI components, each optimized for a specific task: [21,28,35–45].

1. DeepVariant for deep-learning–based variant calling with high fidelity.
2. SpliceAI for detection of cryptic and non-canonical splice variants.
3. AlphaMissense and REVEL for probabilistic assessment of missense pathogenicity.
4. Phenotype-aware ranking tools (e.g., Exomiser) that align genetic variants with patient-specific HPO profiles.

Variants are systematically filtered and ranked against the patient’s phenotype, producing a concise shortlist suitable for expert clinical review. Importantly, the clinical objective is not exhaustive variant identification but efficient prioritization of biologically plausible candidates that directly address the diagnostic delays described in Section 1. A clinically deployable architecture integrating deep learning–based variant calling with phenotype-aware prioritization is illustrated in Figure 1.

While AI-enabled genomic pipelines have demonstrated improvements in the efficiency of variant prioritization, their real-world performance remains variable and context-dependent [21,28,35–45]. Reported sensitivity and diagnostic yield are influenced by multiple factors, including sequencing quality, completeness and accuracy of phenotypic annotation, and the availability and curation of reference databases. In clinical practice, false positive burden represents a significant challenge, particularly when large numbers of candidate variants are generated and require expert review and adjudication [21,28,35–45]. In addition, model performance may be reduced in populations that are underrepresented in training datasets, limiting generalizability and raising concerns regarding bias and equity. Furthermore, variability across sequencing platforms and institutional workflows may affect reproducibility and consistency of results. Collectively, these limitations underscore the need for rigorous external validation, careful clinical integration, and continued reliance on expert interpretation when applying AI-assisted genomic pipelines in rare disease diagnostics.

A Synergistic Workflow: From Patient Phenotype to Confirmed Mutation



This schematic illustrates the integration of deep learning-based variant calling with ontology-driven phenotype-prioritization to accelerate the identification of causal mutations

Figure 1. AI-enabled genomic diagnostic workflow from patient phenotype to confirmed causal mutation. Clinical evaluation initiates whole-exome sequencing, generating raw sequencing data (FASTQ). Deep learning-based variant calling (e.g., DeepVariant) identifies genomic variants with high sensitivity. Ontology-driven prioritization tools (e.g., Exomiser) integrate patient-specific phenotypes using structured vocabularies such as the Human Phenotype Ontology, which enables precise phenotype encoding and computational matching to known genotype-phenotype relationships. Shortlisted variants undergo expert clinical interpretation, culminating in a structured diagnostic report. This integrated pipeline demonstrates how artificial intelligence compresses the diagnostic timeline by combining high-fidelity variant detection with phenotype-aware ranking, enabling efficient identification of causal mutations in rare diseases.

Recent advances in large-scale genomics foundation models further extend this pipeline from variant prioritization toward mechanistic hypothesis generation and biological interpretation. AlphaGenome [46,47], a sequence-to-biology model developed by Google DeepMind, directly processes raw DNA sequence to predict regulatory and molecular effects across both coding and noncoding regions, including gene expression, splicing, chromatin accessibility, transcription factor binding, and long-range genomic interactions [48]. Unlike modular pipelines that rely on independent, task-specific models, AlphaGenome provides a unified representation of genomic function, enabling context-aware interpretation of variants beyond protein-coding regions. In practice, an optimized clinical workflow may pair conventional pipelines for high-sensitivity variant detection and filtering with foundation models for downstream mechanistic explanation, reframing the central clinical question from “Is this variant pathogenic?” to “How does this variant alter biological function?” This complementary integration represents a shift from probabilistic ranking toward biology-informed mechanistic interpretation, while remaining investigational and requiring expert clinical interpretation and external validation before routine clinical deployment in rare disease genomics [46,47] (Table 2).

However, these integrated models also introduce important limitations. Their deep learning architectures often reduce interpretability, making it challenging to trace specific predictions to biologically transparent mechanisms. In addition, their performance depends heavily on large-scale training datasets, which may not fully represent diverse populations or rare variant distributions, raising concerns regarding bias and equity. Furthermore, clinical generalization remains an ongoing challenge, as model performance may vary across sequencing platforms, patient populations, and healthcare settings. These limitations

underscore the need for rigorous external validation, careful clinical integration, and continued human oversight when deploying such models in rare disease diagnostics.

Table 2. Comparison of Conventional Variant-Prioritization Pipelines and Genomics Foundation Models in Rare Disease Diagnostics [46,47].

| Aspect | Conventional Variant-Prioritization Pipelines (DeepVariant, SpliceAI, AlphaMissense or REVEL, Exomiser) | AlphaGenome (Genomics Foundation Model) |
|-----------------------------------|---|--|
| Primary input | Processed sequencing variants | Raw DNA sequence |
| Modeling approach | Multiple independent task-specific models | Unified sequence-to-biology foundation model |
| Primary output | Pathogenicity scores and ranked candidate variants | Predicted regulatory, transcriptional, and molecular effects |
| Coding variant interpretation | Strong performance | Strong performance |
| Noncoding variant interpretation | Limited coverage | Major strength |
| Splicing analysis | Separate specialized models (e.g., SpliceAI) | Integrated within unified regulatory predictions |
| Regulatory element modeling | Limited or indirect | Explicitly modeled |
| Long-range genomic interactions | Generally absent | Included |
| Clinical interpretation focus | “Is this variant likely pathogenic?” | “How might this variant alter biological function?” |
| Typical role in clinical workflow | Variant filtering and phenotype-guided prioritization | Mechanistic interpretation and biological hypothesis generation |
| Major limitation | Fragmented architecture and limited noncoding interpretation | Reduced interpretability and ongoing need for external clinical validation |

Abbreviations: DNA, deoxyribonucleic acid; REVEL, Rare Exome Variant Ensemble Learner.

2.3. Validation Metrics That Matter

In rare disease diagnostics, high sensitivity is particularly important. The clinical cost of missing a pathogenic variant far outweighs the burden of reviewing false positives. Accordingly, pipelines should target: [44,45,49–54].

- Sensitivity approaching or exceeding 90 percent for known pathogenic variants
- Top-3 or Top-5 recall rates exceeding 80 percent in phenotype-driven prioritization tasks
- Median end-to-end turnaround times of ≤ 7 days in clinically optimized settings

These thresholds should be interpreted as aspirational and context-dependent benchmarks rather than universally established standards, as performance varies across datasets, sequencing platforms, patient populations, and implementation environments.

Precision optimization should occur only after sensitivity thresholds are met. This hierarchy of metrics reflects clinical reality rather than abstract model performance.

2.4. Real-World Nephrogenetics Implementation

Published nephrogenetics workflows integrating genomic sequencing with phenotype-aware prioritization have demonstrated high diagnostic yield and the potential to improve diagnostic efficiency in selected clinical settings [55–59]. Across reported cohorts, AI-assisted approaches that combine variant calling, splice prediction, and phenotype-driven ranking with expert clinical interpretation have improved the efficiency of variant prioritization and diagnostic workflows. However, these improvements are context-dependent and have been most consistently observed in specialized centers with curated datasets and

established genomic infrastructure. Reported sensitivity, diagnostic yield, and turnaround times vary across studies and are influenced by cohort characteristics, sequencing strategies, and validation frameworks, and may not be generalizable to broader clinical environments.

Rather than representing a single standardized pipeline, these approaches reflect a converging model of care in which artificial intelligence augments, but does not replace, expert clinical judgment. In practice, the clinical impact arises from improved triaging of candidate variants, reduction in manual review burden, and earlier identification of plausible diagnoses within existing laboratory and clinical infrastructures [44,45].

For example, a young adult presenting with unexplained chronic kidney disease, persistent microscopic hematuria, and a family history of renal disease may undergo whole-exome sequencing. In a conventional workflow, interpretation of variants of uncertain significance can be time-intensive and inconclusive. Using AI-assisted variant prioritization integrated with structured phenotype encoding (e.g., Human Phenotype Ontology terms), candidate variants in genes such as *COL4A3* or *COL4A5* may be prioritized based on phenotype–genotype concordance. This can support earlier diagnostic consideration of Alport syndrome, facilitating timely initiation of renin–angiotensin system blockade and appropriate family screening. While such workflows require expert validation, they illustrate how artificial intelligence can enhance efficiency and support clinical decision-making in nephrogenetics.

Overall, current evidence supports the feasibility of AI-augmented genomic diagnosis in rare nephropathies and suggests that, when implemented within supervised clinical workflows, such systems can accelerate specific components of the diagnostic process [49,50,54,55]. However, these gains are not uniform, and broader generalizability and reproducibility remain dependent on external validation, prospective evaluation, and integration into routine nephrology practice [60,61]. As such, AI-assisted diagnostic support should be interpreted as a promising but evolving capability rather than an established standard of care. The critical next step is translation into routine nephrology workflows, where AI-enabled genomic pipelines can be systematically embedded into clinical decision-making processes; several high-yield, workflow-integrated use cases are outlined in Table 3.

Table 3. Practical Workflow-Integrated AI Applications for Nephrologists Managing Rare Diseases [60,61].

| Clinical Scenario | Core Clinical Challenge | AI-Assisted Workflow Application | Potential Clinical Impact |
|--|--|---|---|
| Phenotype summarization at initial evaluation | Rare disease features are dispersed across years of clinical documentation | Summarize longitudinal EHR data into structured phenotype profiles and extract key symptom patterns before genetic evaluation | Shortens diagnostic delays; improves genetics referrals; reduces premature labeling as “idiopathic CKD” |
| Rare disease detection during routine CKD care | Early disease manifestations are often nonspecific | Identify pattern-based clinical red flags such as hematuria with family history or CKD with extra-renal manifestations | Earlier genetic evaluation; fewer delayed diagnoses at ESKD |
| Variant interpretation support | Variants of uncertain significance complicate clinical decision-making | Summarize variant annotations and identify clinical features that strengthen or weaken pathogenicity assessment | Supports clinically actionable monitoring and multidisciplinary discussion |

Table 3. Cont.

| Clinical Scenario | Core Clinical Challenge | AI-Assisted Workflow Application | Potential Clinical Impact |
|---|--|---|---|
| Differential diagnosis expansion | Anchoring bias toward common glomerular diseases | Expand differential diagnoses for atypical CKD presentations or discordant biopsy–phenotype findings | Improves rare disease recognition and diagnostic breadth |
| Clinical trial and expanded-access matching | Dynamic eligibility criteria contribute to missed opportunities | Match evolving phenotypes and laboratory trends to clinical trials or expanded-access programs | Improves trial access and referral equity |
| Longitudinal disease monitoring | Disease progression may occur gradually over years | Track longitudinal trajectories including eGFR decline, proteinuria trends, and evolving symptom patterns | Earlier intervention and more individualized treatment escalation |
| Patient communication and education | Rare disease counseling is often inconsistent and time-intensive | Generate patient-specific educational summaries aligned with diagnosis, laboratory findings, and treatment plans | Improves adherence, communication efficiency, and shared decision-making |
| Multidisciplinary case preparation | Multidisciplinary reviews are labor-intensive | Prepare structured case summaries including phenotype data, genomic findings, unresolved questions, and management considerations | More efficient multidisciplinary discussions and standardized documentation |

Abbreviations: AI, artificial intelligence; CKD, chronic kidney disease; EHR, electronic health record; ESKD, end-stage kidney disease; eGFR, estimated glomerular filtration rate.

3. Beyond Single Centers: Federated Learning for Rare Disease Discovery

3.1. The Limits of Local Excellence

Even the most advanced single-institution pipelines cannot overcome the fundamental limitation of ultra-rare disease prevalence [62–65]. Discovery of novel gene–disease associations, progression models, and therapeutic biomarkers requires learning across institutions [66,67].

3.2. Federated Learning as a Privacy-Preserving Solution

Federated learning enables distributed model training without sharing patient-level data [66,68–71]. Each institution trains a local model on harmonized data, and model updates or parameters, often combined with privacy-enhancing safeguards, are aggregated centrally rather than transferring raw patient-level data. This paradigm preserves privacy while enabling collective intelligence across institutions [66]. However, reproducibility and real-world performance remain important considerations, as model outputs may vary across sites due to differences in data structure, coding practices, and patient populations, even when harmonized training protocols are applied.

As an illustrative example, consider the development of a predictive model for progression of a rare glomerular disease across multiple institutions, each with limited patient numbers. Using a federated learning framework, participating centers can locally train models on harmonized datasets while sharing only model parameters. The aggregated model benefits from a larger, more diverse dataset without requiring transfer of patient-level data. In practice, such an approach may improve risk stratification and enable earlier

identification of patients at risk of rapid progression, although performance remains dependent on data harmonization and consistent implementation across sites.

While centralized aggregation remains the most widely implemented approach, alternative architectures are increasingly being explored to address scalability, robustness, and governance considerations. Decentralized approaches, such as peer-to-peer model exchange, enable institutions to share updates directly without reliance on a single coordinating server, potentially reducing single points of failure and enhancing resilience. Hybrid models combine centralized orchestration with hierarchical or regional aggregation layers, allowing for flexible coordination across institutional networks while preserving local control. These emerging modalities expand the design space for federated learning systems and may be particularly relevant in rare disease contexts, where data are distributed across geographically and organizationally diverse centers.

In practice, operationalizing federated learning requires coordinated computational infrastructure, including secure local environments for on-site model training, reliable communication of encrypted updates, and standardized aggregation procedures. Variability in hardware capacity, network stability, and institutional technical expertise can affect training efficiency, reproducibility, and scalability. In addition, differences in data harmonization, feature representation, and clinical workflow integration may further influence model performance across sites. These considerations highlight that, beyond conceptual advantages, federated learning systems require substantial operational coordination to achieve consistent and clinically meaningful performance across healthcare settings.

3.3. *The Role of OMOP Common Data Model*

Interoperability is the prerequisite for federated AI. The Observational Medical Outcomes Partnership (OMOP) Common Data Model provides standardized tables and vocabularies for demographics, diagnoses, medications, laboratory measurements [72], procedures, and encounters [72–74]. Mapping local data to OMOP allows identical training scripts to run across sites, ensuring methodological consistency.

Beyond structural standardization, effective data harmonization requires alignment of coding systems, phenotype definitions, and temporal data representation. This includes mapping local terminologies to standardized vocabularies, normalizing laboratory units, and ensuring consistent representation of longitudinal clinical trajectories. Such harmonization enables models trained across institutions to learn from comparable signals despite differences in local practice patterns and data capture.

3.4. *Governance and Operational Considerations*

Successful multi-institutional AI deployment also depends on robust governance frameworks. These include data use agreements that define permissible data access and model outputs, clear accountability structures across participating institutions, and mechanisms for ongoing model monitoring and validation. Governance must also address issues of bias, fairness, and equitable representation, particularly given the heterogeneity of rare disease populations. In addition, regulatory alignment and auditability are essential to ensure that federated models remain transparent and clinically trustworthy over time [66,67,75,76].

3.5. *Proof Points Across Domains*

Federated learning has already demonstrated feasibility and clinical relevance: [66,67,75,76].

- Multi-institutional imaging and clinical prediction models with performance approaching centralized models while preserving data locality and privacy
- Federated pathogenicity classification models in genomics

- Rare hematologic disease platforms achieving clinically meaningful discrimination
- Federated acute kidney injury prediction across intensive care units

These examples provide direct analogs for rare nephropathies and genetic kidney diseases.

4. From Prediction to Individualized Causal Reasoning: Digital Twins in Rare Disease

4.1. Biology-Informed Modeling Beyond Statistical Association

An evolving conceptual transition in clinical artificial intelligence increasingly emphasizes biologically informed and individualized modeling approaches rather than systems optimized solely for statistical association [13,14,77]. This perspective is supported by growing literature highlighting the limitations of purely correlation-based models in complex and heterogeneous disease states, particularly in rare diseases where mechanistic priors are strong and data are inherently sparse [78]. Patient-specific digital twin models represent one emerging approach that integrates biological knowledge with patient-specific data to support simulation of disease trajectories and therapeutic responses [13,14].

In contrast to models that rely primarily on statistical pattern recognition, biology-informed approaches incorporate molecular pathways, genotype–phenotype relationships, and physiological interactions, thereby enabling more interpretable and clinically meaningful predictions [13,14]. This shift from statistical association to mechanism-informed inference is particularly relevant in rare disease settings, where individualized reasoning and causal understanding are essential for accurate diagnosis and treatment planning [13,14]. Figure 2 provides an integrated view of how upstream AI components, including phenotype extraction, triage, and genomic interpretation, combine to support digital twin–based simulation and actionable clinical outputs [78–80].

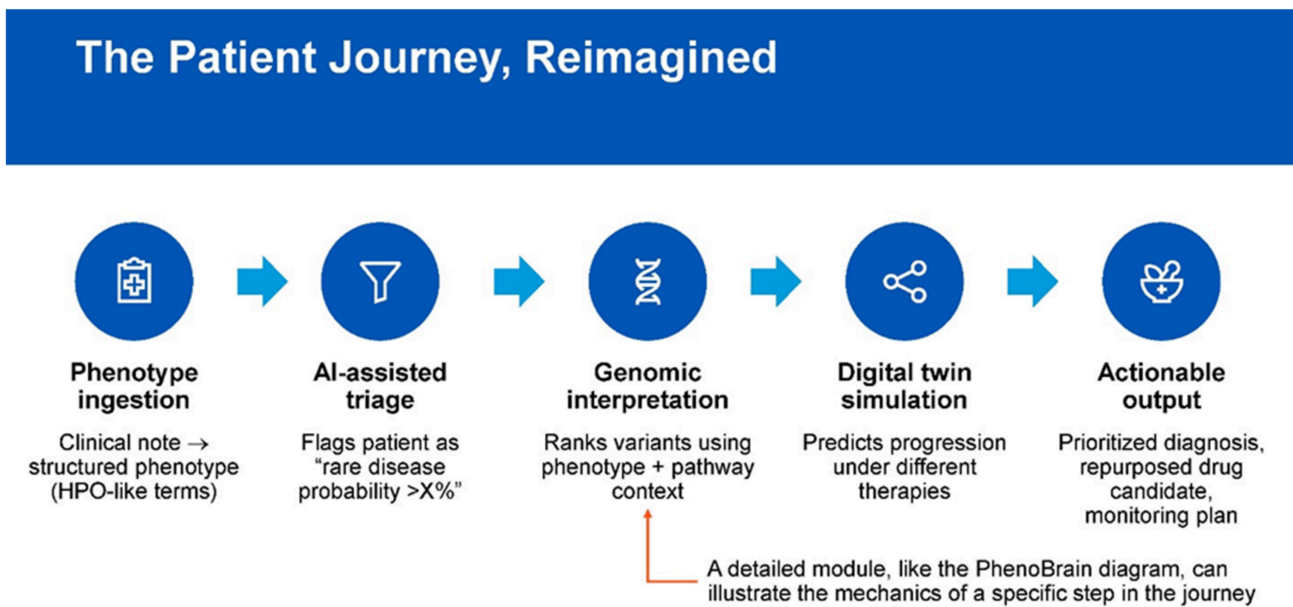


Figure 2. Integrated AI-driven rare disease patient journey. Clinical data are first transformed into structured phenotypes, enabling early AI-assisted triage to identify patients at risk of rare disease. Genomic interpretation integrates phenotype and pathway context to prioritize candidate variants. These outputs feed into patient-specific digital twin models that simulate disease trajectories under alternative therapeutic strategies. The pipeline culminates in actionable clinical outputs, including prioritized diagnoses, therapeutic recommendations, and longitudinal monitoring plans. This figure synthesizes how multiple AI paradigms converge into a workflow-native system for individualized, data-driven care in rare diseases.

A conceptual clinical application of digital twins can be illustrated in a patient with a rare hereditary nephropathy and variable disease progression. A digital twin model integrating longitudinal laboratory data (e.g., estimated glomerular filtration rate trajectory), genetic findings, and treatment history could simulate disease progression under different therapeutic strategies, such as early versus delayed initiation of renin–angiotensin system blockade. While current implementations remain limited, such models could eventually support individualized treatment planning by providing scenario-based projections, contingent on further validation and clinical integration. This example highlights how digital twin frameworks translate biological and longitudinal data into clinically meaningful simulations, bridging conceptual modeling with potential decision support in rare disease care.

Digital twin models represent an emerging framework for integrating patient-specific data to simulate disease trajectories and therapeutic responses [79,81,82]. While early studies have demonstrated feasibility in selected domains, clinical implementation in rare disease care remains limited [83,84]. Most current applications are based on conceptual models, retrospective analyses, or small-scale studies, and robust prospective or multi-center validation is lacking [83,84]. In addition, variability in data quality, model assumptions, and implementation context further constrains generalizability across clinical settings [77,83]. Accordingly, digital twins should be considered a forward-looking approach with potential for future clinical integration rather than a near-term standard of care, and their adoption will require rigorous validation, standardized evaluation frameworks, and careful integration into clinical workflows.

4.2. Architecture of a Rare Disease Digital Twin

A digital twin comprises four integrated layers: [66,77–79,85,86].

1. Phenotypic Twin: Structured, time-aware representation of clinical features using standardized ontologies.
2. Molecular and Pathway Twin: Genomic and proteomic representations linking variants to biological mechanisms.
3. Physiology and Organ-System Twin: Functional modeling of organ interactions and disease progression.
4. Intervention Simulation Twin: Counterfactual testing of therapeutic strategies.

Such systems evolve longitudinally as new data accrue, enabling detection of phenotypic drift before clinical thresholds are crossed.

Despite their conceptual appeal, digital twin models face substantial limitations in current practice. Their performance depends on the availability of high-quality, longitudinal, and multimodal data, which are often incomplete, inconsistently captured, or not readily interoperable in real-world clinical settings. In addition, the underlying biological representations may rely on simplified or partially specified models of complex disease processes, potentially limiting predictive accuracy and mechanistic validity. Validation is frequently restricted to small cohorts, retrospective analyses, or simulated environments, and standardized frameworks for evaluating clinical performance remain underdeveloped. These constraints introduce risks of overfitting, misinterpretation, and limited generalizability across patient populations and care settings. Furthermore, the computational complexity and infrastructure requirements of digital twin systems may limit scalability beyond specialized centers. Collectively, these limitations underscore the need for cautious interpretation, rigorous prospective validation, and careful integration into clinical workflows before routine clinical deployment.

4.3. Generative AI and Synthetic Patients

Generative models offer a potential solution to data scarcity by synthesizing realistic patient trajectories while preserving statistical properties of real-world data [20,87–92]. When rigorously validated, synthetic cohorts may support hypothesis generation, trial design, and algorithm stress-testing in ultra-rare conditions.

Agentic AI systems, which are capable of autonomous task planning and multi-step reasoning, represent a rapidly evolving area of research [93]. By design, these systems can interact with external tools and generate adaptive outputs, increasing both their potential utility and their risk profile. Although early demonstrations in controlled experimental settings are promising, their application in clinical rare disease care remains largely exploratory. Current limitations include susceptibility to reasoning errors, variability in performance across tasks, and limited transparency in decision-making processes. In clinical contexts, these issues raise concerns regarding safety, accountability, and reproducibility. Furthermore, robust validation frameworks for agentic systems remain underdeveloped, particularly in rare disease settings where standardized benchmarks are lacking. Accordingly, these systems should be considered investigational and, if used, deployed within tightly controlled environments with clear human oversight, rather than as independent decision-makers in routine clinical practice.

5. Accountability, Regulation, and Adaptive Governance for Rare Disease AI

As rare disease AI systems transition from static models to continuously learning, federated, and agentic architectures, questions of accountability and regulatory oversight become central rather than peripheral. Federated learning environments refer to distributed computational settings in which multiple institutions collaboratively train models using local data, while sharing only model parameters or updates rather than raw patient-level data [94–97]. In such environments, model performance may drift over time as local data distributions evolve, institutions join or leave networks, or clinical practices change [97,98]. Determining accountability in such settings requires explicit governance structures that define responsibility across stakeholders, including local deploying institutions, central coordinating bodies, and model developers. Without clearly assigned ownership for monitoring, auditing, and corrective action, federated systems risk becoming clinically opaque despite their privacy-preserving design [99,100].

Generative models and synthetic cohorts further complicate this governance landscape. Generative approaches offer a potential solution to data scarcity by synthesizing realistic patient trajectories while preserving statistical properties of real-world data [87,101–105]. Several validated implementations of synthetic clinical data have been reported, demonstrating that appropriately generated datasets can reproduce key statistical distributions and support downstream analytical tasks, including model development and validation [87,106,107]. For example, prior studies have shown that synthetic datasets can approximate real-world clinical quality measures and population-level characteristics, supporting their use in methodological development and benchmarking [87,92,108–111].

However, these implementations remain context-dependent, and validation is often limited to specific use cases, such as common diseases or structured clinical datasets [87,101–105]. In rare disease settings, where phenotypic heterogeneity is high and sample sizes are extremely limited, the generation of reliable synthetic cohorts remains more challenging and less well established [13,14]. Accordingly, while synthetic data hold promise for hypothesis generation, trial design, and algorithm stress-testing, their clinical applicability requires careful validation to ensure fidelity, generalizability, and absence of bias [87,101–105]. These considerations introduce additional layers of accountability,

including the need to define responsibility for synthetic data generation, validation, and downstream use in clinical or research settings.

In rare disease contexts, clinical utility may not be adequately captured by conventional performance metrics such as overall accuracy or area under the curve. Instead, effective systems must demonstrate sensitivity to individual pathogenic signals, defined as the ability to detect clinically meaningful, patient-specific abnormalities, including rare variants or atypical phenotype combinations that may be diluted in population-level analyses. This differs fundamentally from traditional evaluation approaches, which emphasize average performance across large cohorts rather than detection of rare but high-impact signals [20,112,113]. In rare disease settings, leaderboard performance on static benchmark datasets is therefore a limited proxy for clinical utility, which is more appropriately defined by sensitivity to individual pathogenic signals, robustness to data sparsity, and integration into real-world diagnostic workflows [16,45,114–116].

The proposed evaluation framework extends beyond conventional metrics by incorporating domains that can be operationalized in practice, including sensitivity to individual pathogenic signals, biological plausibility, and longitudinal consistency. Sensitivity to individual pathogenic signals may be evaluated by assessing the model’s ability to correctly prioritize known pathogenic variants or clinically confirmed diagnoses in patient-specific cases. Biological plausibility can be assessed through concordance with established genotype–phenotype relationships and known disease mechanisms. Longitudinal validation may be operationalized by tracking agreement between predicted and observed disease trajectories over time within the same patient. These domains complement traditional performance measures such as sensitivity, specificity, and area under the curve, which remain important but may not fully capture clinical utility in rare disease contexts. Accordingly, alternative validation frameworks that do not rely solely on randomized controlled trials are required; key validation pillars applicable to rare disease AI are summarized in Table 4 [20,112,113]. These approaches are intended to complement, rather than replace, conventional clinical validation paradigms where feasible.

Table 4. Validation Approaches for AI Systems in Rare Disease Care Beyond Conventional Randomized Trials.

| Validation Domain | Core Clinical Question | What Is Evaluated | Relevance in Rare Diseases | Comparison with Conventional Validation |
|--------------------------------|--|--|---|--|
| Causal plausibility | Does the model output align with known disease biology? | Concordance between AI predictions and established genotype–phenotype or pathway relationships | Rare diseases often have strong mechanistic foundations that support biologically informed validation | Complements population-level statistical inference |
| Counterfactual simulation | Could an alternative intervention plausibly alter patient outcomes? | Simulation of disease trajectories under different therapeutic strategies within patient-specific models | Supports individualized causal reasoning when randomized trials are impractical or infeasible | Complements parallel control-arm comparisons |
| Longitudinal N-of-1 validation | Does the prediction remain consistent over time within the same patient? | Agreement between predicted and observed longitudinal trajectories including laboratory values, imaging findings, and symptoms | Particularly relevant in slowly progressive and highly individualized rare diseases | Complements large cohort reproducibility assessments |

Abbreviations: N-of-1, single-patient (within-individual) longitudinal evaluation.

Figure 3 provides a workflow-integrated evaluative framework that synthesizes these validation principles with upstream data inputs, AI modeling approaches, and downstream clinical integration. The figure illustrates how rare disease AI systems should be assessed not only on model performance but also on biological plausibility, longitudinal consistency, and real-world deployability within clinical workflows.

Workflow-Integrated Evaluative Framework for AI in Rare Disease Care

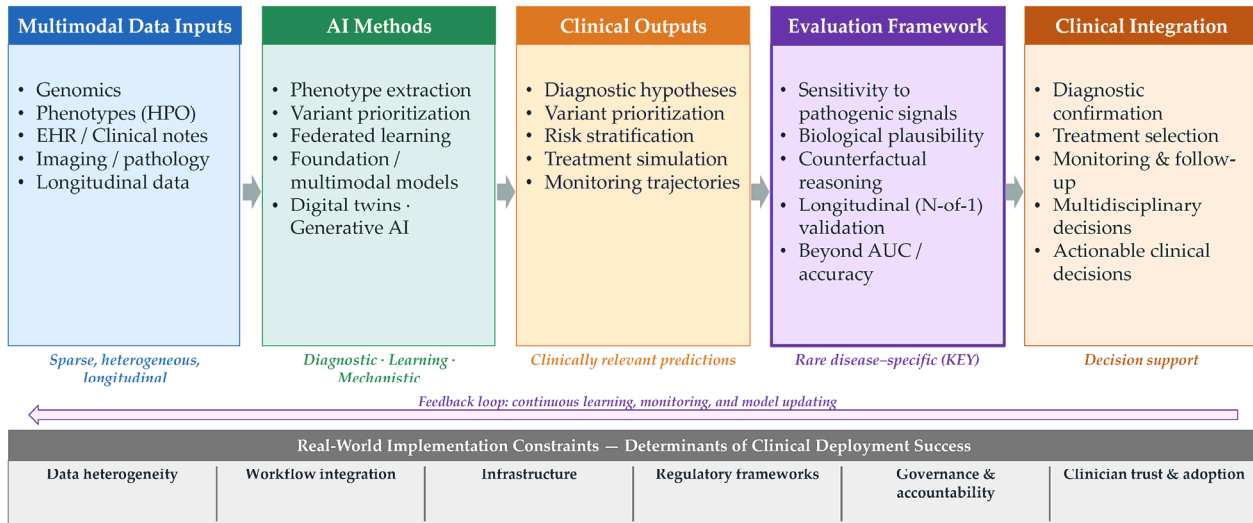


Figure 3. Workflow-integrated evaluative framework for AI in rare disease care. Multimodal data inputs—including genomics, clinical phenotypes, imaging, and electronic health records—are processed through complementary AI approaches such as phenotype extraction, genomic variant prioritization, federated learning, and digital twin modeling. These components converge to generate clinically relevant outputs, including diagnostic hypotheses, risk stratification, and therapeutic insights. Evaluation is structured across multiple domains: sensitivity to individual pathogenic signals, biological plausibility, counterfactual reasoning, and longitudinal validation. The framework also incorporates real-world constraints, including workflow integration, infrastructure requirements, governance, and regulatory oversight, highlighting the pathway from model development to clinically actionable, trustworthy systems. Abbreviations: HPO, Human Phenotype Ontology; AI, artificial intelligence; AUC, area under the curve; N-of-1, within-individual longitudinal evaluation.

Across AI approaches in rare disease care, several cross-cutting limitations warrant careful consideration. Failure modes include false positive prioritization of variants, reduced sensitivity in atypical or incomplete phenotypic presentations, and performance degradation due to data drift over time. Bias may arise from underrepresentation of certain populations in training datasets, resulting in unequal performance across demographic groups and potential disparities in diagnostic accuracy. In addition, models optimized in controlled research environments may not maintain performance when deployed in heterogeneous clinical settings with variable data quality and workflow structures. These challenges underscore the importance of continuous performance monitoring, periodic recalibration, and transparent reporting of model limitations to ensure reliable and equitable use in real-world clinical practice.

Regulatory evaluation of continuously learning AI systems poses additional challenges [117–119]. Traditional approval paradigms, which assume fixed algorithms and static performance characteristics, are poorly suited to models that update iteratively or adapt across sites. Emerging regulatory concepts, such as predetermined change control plans and lifecycle-based oversight, offer a potential pathway but remain underdeveloped for federated and rare disease contexts [119–121]. In practice, post-deployment surveillance can be implemented through continuous performance monitoring, detection of data drift,

and periodic recalibration using local data. Auditability requires maintaining detailed logs of model inputs, outputs, and updates, enabling retrospective evaluation of clinical decisions, while transparency may be supported through model documentation, explainability tools, and clear communication of model limitations to clinicians and patients.

Rare diseases may warrant adaptive regulatory pathways distinct from those applied to common conditions [122,123]. In this context, carefully governed conditional deployment, coupled with rigorous real-world performance monitoring and mandatory reassessment, provides a pragmatic balance between innovation and patient safety [123]. Such frameworks parallel adaptive trial designs, accelerated approval pathways, and regulatory approaches for software as a medical device, offering early access under controlled conditions while ensuring ongoing validation [123–125].

Translating AI systems from controlled research environments into routine clinical practice presents substantial challenges. Data heterogeneity across institutions, including differences in documentation practices, coding systems, and patient populations, can significantly affect model performance and generalizability. Integration into clinical workflows requires interoperability with electronic health record systems, user-centered interface design, and alignment with clinician decision-making processes. In addition, implementation demands technical infrastructure, computational resources, and ongoing maintenance that may not be uniformly available. Regulatory uncertainty, particularly for adaptive and continuously learning systems, further complicates deployment. Finally, clinician trust and adoption remain critical barriers, as effective use depends on transparency, interpretability, and demonstrated clinical benefit.

Feasibility of AI deployment varies substantially across healthcare environments [61]. While large academic centers may have access to advanced computational infrastructure, curated datasets, and specialized expertise, smaller or resource-constrained settings may face significant barriers related to data availability, technical capacity, and workforce training. As a result, models developed and validated in highly resourced environments may not directly translate to broader clinical contexts without adaptation. In addition, differences in data standards, interoperability, and clinical workflow integration may further affect performance and usability across settings. Addressing these disparities will require scalable implementation strategies, standardized data frameworks, and targeted efforts to ensure equitable access to AI-enabled care.

In clinical practice, AI-supported decision tools may be used during multidisciplinary case discussions for patients with complex or atypical presentations. For instance, in a patient with discordant biopsy findings and clinical features, AI-assisted summarization of longitudinal data and expansion of differential diagnoses may help identify overlooked rare conditions. Such tools function as decision support systems rather than replacements for clinical judgment, and their effectiveness depends on appropriate integration into existing workflows and clinician oversight.

Taken together, effective governance of rare disease AI will require alignment between technical design, institutional accountability, and regulatory innovation. In practice, this coordination may be operationalized through multidisciplinary oversight committees, standardized reporting frameworks, shared performance benchmarks, and clearly defined stakeholder responsibilities across development, deployment, and monitoring phases. Without such coordination, even technically robust systems risk failing at the point of clinical trust and adoption.

6. Conclusions

Current evidence suggests that AI applications may improve diagnostic efficiency, support privacy-preserving multi-institutional discovery, and facilitate the development of

patient-specific modeling approaches such as digital twins in rare disease care. Privacy-preserving discovery refers to approaches including federated learning and distributed model training, in which data remain within local institutions while models learn from aggregated, anonymized parameters, thereby facilitating collaborative knowledge generation without direct data sharing. Current evidence suggests that clinically meaningful AI systems in rare disease care should prioritize sensitivity to individual pathogenic signals, interoperability, biological plausibility, and integration into existing clinical workflows rather than focusing solely on incremental gains in benchmark performance metrics. Collectively, these developments position rare disease care as an important translational environment for evaluating clinically integrated AI applications.

However, these advances must be interpreted within the context of important methodological and practical limitations. If validation frameworks remain centered on static datasets, retrospective benchmarks, and population-level metrics, many systems are unlikely to translate effectively into real-world settings, particularly in rare diseases, where delayed or missed diagnoses carry disproportionate clinical consequences. In addition, models optimized primarily for average population-level performance may inadequately capture individualized disease trajectories and heterogeneous phenotypic presentations encountered in rare disease care. Models optimized for average performance rather than individualized reasoning risk reinforcing diagnostic inertia, defined as the persistence of delayed or incomplete diagnostic processes driven by reliance on conventional workflows, anchoring bias, and insufficient integration of longitudinal patient-specific data.

The need for individualized reasoning in rare disease care reflects the inherent heterogeneity and low prevalence of these conditions. In practice, this can be supported by AI systems that integrate structured phenotypic data, genomic information, and longitudinal clinical trajectories to enable N-of-1 inference, phenotype-driven prioritization, and patient-specific risk modeling. Such approaches allow dynamic updating of diagnostic hypotheses and therapeutic strategies as new data become available, aligning AI outputs with real-world clinical decision-making. At the same time, progress in this field will depend not only on algorithmic innovation but also on rigorous validation, careful evaluation of failure modes, and thoughtful integration into clinical workflows. As discussed throughout this review, many emerging approaches, including digital twins, federated learning architectures, and agentic AI systems, remain at varying stages of maturity and clinical validation. Accordingly, their future role in routine practice will depend on prospective evaluation, reproducibility across healthcare environments, and demonstration of clinically meaningful benefit.

While rare disease settings offer a compelling environment for advancing individualized and adaptive AI systems, translation into routine practice requires robust governance, continuous monitoring, and clear delineation of the role of AI as decision support rather than autonomous clinical authority. Overall, the current literature supports cautious optimism regarding the potential of AI-enabled rare disease care while underscoring the continued importance of clinician oversight, implementation science, and real-world translational evaluation. Taken together, a balanced approach that combines innovation with critical appraisal is essential to ensure that artificial intelligence systems deliver reliable, equitable, and clinically meaningful benefits.

Author Contributions: Conceptualization, C.T., F.P. and W.C.; Writing—original draft preparation, W.C.; Writing—review and editing, C.T. and F.P.; Visualization, C.T., F.P. and W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-5, OpenAI, San Francisco, CA, USA) for the purpose of language editing. AI tools were not used for literature search, data analysis, or generation of scientific content. All outputs were reviewed and edited by the authors, who take full responsibility for the accuracy, integrity, and interpretation of the work.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-------|---|
| AI | Artificial intelligence |
| ACEi | Angiotensin-converting enzyme inhibitor |
| CKD | Chronic kidney disease |
| COL4A | Type IV collagen alpha chain genes |
| EHR | Electronic health record |
| ESKD | End-stage kidney disease |
| GNNs | Graph neural networks |
| GBM | Glomerular basement membrane |
| HPO | Human Phenotype Ontology |
| MDT | Multidisciplinary Team |
| REVEL | Rare Exome Variant Ensemble Learner |
| RL | Reinforcement learning |
| RCTs | Randomized controlled trials |
| VUS | Variant of Uncertain Significance |
| XAI | Explainable artificial intelligence |

References

1. 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](#)]
2. Baynam, G.; Hartman, A.L.; Letinturier, M.C.V.; Bolz-Johnson, M.; Carrion, P.; Grady, A.C.; Dong, X.; Dooms, M.; Dreyer, L.; Graessner, H.; et al. Global health for rare diseases through primary care. *Lancet Glob. Health* **2024**, *12*, e1192–e1199. [[CrossRef](#)] [[PubMed](#)]
3. Klimova, B.; Storek, M.; Valis, M.; Kuca, K. Global View on Rare Diseases: A Mini Review. *Curr. Med. Chem.* **2017**, *24*, 3153–3158. [[CrossRef](#)]
4. Groft, S.C.; Posada, M.; Taruscio, D. Progress, challenges and global approaches to rare diseases. *Acta Paediatr.* **2021**, *110*, 2711–2716. [[CrossRef](#)]
5. Faye, F.; Crocione, C.; Anido de Peña, R.; Bellagambi, S.; Escati Peñalosa, L.; Hunter, A.; Jensen, L.; Oosterwijk, C.; Schoeters, E.; de Vicente, D.; et al. Time to diagnosis and determinants of diagnostic delays of people living with a rare disease: Results of a Rare Barometer retrospective patient survey. *Eur. J. Hum. Genet.* **2024**, *32*, 1116–1126. [[CrossRef](#)]
6. Wise, A.L.; Manolio, T.A.; Mensah, G.A.; Peterson, J.F.; Roden, D.M.; Tamburro, C.; Williams, M.S.; Green, E.D. Genomic medicine for undiagnosed diseases. *Lancet* **2019**, *394*, 533–540. [[CrossRef](#)] [[PubMed](#)]
7. Benito-Lozano, J.; López-Villalba, B.; Arias-Merino, G.; Posada de la Paz, M.; Alonso-Ferreira, V. Diagnostic delay in rare diseases: Data from the Spanish rare diseases patient registry. *Orphanet J. Rare Dis.* **2022**, *17*, 418. [[CrossRef](#)]
8. Benito-Lozano, J.; Arias-Merino, G.; Gómez-Martínez, M.; Ancochea-Díaz, A.; Aparicio-García, A.; Posada de la Paz, M.; Alonso-Ferreira, V. Diagnostic Process in Rare Diseases: Determinants Associated with Diagnostic Delay. *Int. J. Environ. Res. Public Health* **2022**, *19*, 6456. [[CrossRef](#)]
9. Zurynski, Y.; Deverell, M.; Dalkeith, T.; Johnson, S.; Christodoulou, J.; Leonard, H.; Elliott, E.J. Australian children living with rare diseases: Experiences of diagnosis and perceived consequences of diagnostic delays. *Orphanet J. Rare Dis.* **2017**, *12*, 68. [[CrossRef](#)]
10. Sussman, J.H.; Dagli, M.M.; Wald, S.L.; Akhtar, S.S.; Natan, K.; Xu, J.A.; Li, A.; Dawson, B.; Welch, W.C. The Experiences of Patients With Rare Diseases in Pennsylvania: A Community-Based Study. *Public Health Rep.* **2025**, *140*, 605–616. [[CrossRef](#)]

11. Benito-Lozano, J.; Arias-Merino, G.; Gómez-Martínez, M.; Arconada-López, B.; Ruiz-García, B.; Posada de la Paz, M.; Alonso-Ferreira, V. Psychosocial impact at the time of a rare disease diagnosis. *PLoS ONE* **2023**, *18*, e0288875. [[CrossRef](#)]
12. Chazal, P.E.; Aymé, S. An Objective Approach to Identify Priority Rare Diseases for the Development of Solutions Reducing the Diagnostic Delay Based on French Data. *Front. Pharmacol.* **2021**, *12*, 734601. [[CrossRef](#)]
13. Banerjee, J.; Taroni, J.N.; Allaway, R.J.; Prasad, D.V.; Guinney, J.; Greene, C. Machine learning in rare disease. *Nat. Methods* **2023**, *20*, 803–814. [[CrossRef](#)]
14. Decherchi, S.; Pedrini, E.; Mordenti, M.; Cavalli, A.; Sangiorgi, L. Opportunities and Challenges for Machine Learning in Rare Diseases. *Front. Med.* **2021**, *8*, 747612. [[CrossRef](#)]
15. Liu, J.; Barrett, J.S.; Leonardi, E.T.; Lee, L.; Roychoudhury, S.; Chen, Y.; Trifillis, P. Natural History and Real-World Data in Rare Diseases: Applications, Limitations, and Future Perspectives. *J. Clin. Pharmacol.* **2022**, *62*, S38–S55. [[CrossRef](#)]
16. Schaefer, J.; Lehne, M.; Schepers, J.; Prasser, F.; Thun, S. The use of machine learning in rare diseases: A scoping review. *Orphanet J. Rare Dis.* **2020**, *15*, 145. [[CrossRef](#)] [[PubMed](#)]
17. Partington, G.; Cro, S.; Mason, A.; Phillips, R.; Cornelius, V. Design and analysis features used in small population and rare disease trials: A targeted review. *J. Clin. Epidemiol.* **2022**, *144*, 93–101. [[CrossRef](#)]
18. Dingemans, A.J.M.; Hinne, M.; Truijien, K.M.G.; Goltstein, L.; van Reeuwijk, J.; de Leeuw, N.; Schuurs-Hoeijmakers, J.; Pfundt, R.; Diets, I.J.; den Hoed, J.; et al. PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. *Nat. Genet.* **2023**, *55*, 1598–1607. [[CrossRef](#)] [[PubMed](#)]
19. Kidwell, K.M.; Roychoudhury, S.; Wendelberger, B.; Scott, J.; Moroz, T.; Yin, S.; Majumder, M.; Zhong, J.; Huml, R.A.; Miller, V. Application of Bayesian methods to accelerate rare disease drug development: Scopes and hurdles. *Orphanet J. Rare Dis.* **2022**, *17*, 186. [[CrossRef](#)]
20. Amorim, A.M.B.; Orzeł, U.; Caniceiro, A.B.; Rosário-Ferreira, N.; Moreira, I.S. Artificial intelligence in rare diseases: Toward clinical impact. *Trends Pharmacol. Sci.* **2025**, *46*, 1241–1268. [[CrossRef](#)] [[PubMed](#)]
21. Gomes, B.; Ashley, E.A. Artificial Intelligence in Molecular Medicine. *N. Engl. J. Med.* **2023**, *388*, 2456–2465. [[CrossRef](#)]
22. Yoo, T.K.; Choi, J.Y.; Kim, H.K. Feasibility study to improve deep learning in OCT diagnosis of rare retinal diseases with few-shot classification. *Med. Biol. Eng. Comput.* **2021**, *59*, 401–415. [[CrossRef](#)] [[PubMed](#)]
23. Alsentzer, E.; Li, M.M.; Kobren, S.N.; Noori, A.; Kohane, I.S.; Zitnik, M. Few shot learning for phenotype-driven diagnosis of patients with rare genetic diseases. *npj Digit. Med.* **2025**, *8*, 380. [[CrossRef](#)]
24. Chen, Y.; Guo, X.; Pan, Y.; Xia, Y.; Yuan, Y. Dynamic feature splicing for few-shot rare disease diagnosis. *Med. Image Anal.* **2023**, *90*, 102959. [[CrossRef](#)]
25. Chen, Y.; Guo, X.; Xia, Y.; Yuan, Y. Disentangle Then Calibrate With Gradient Guidance: A Unified Framework for Common and Rare Disease Diagnosis. *IEEE Trans. Med. Imaging* **2024**, *43*, 1816–1827. [[CrossRef](#)]
26. Lee, J.; Liu, C.; Kim, J.; Chen, Z.; Sun, Y.; Rogers, J.R.; Chung, W.K.; Weng, C. Deep learning for rare disease: A scoping review. *J. Biomed. Inform.* **2022**, *135*, 104227. [[CrossRef](#)]
27. Defelippe, V.M.; van Thiel, G.J.; Otte, W.M.; Schutgens, R.E.; Stunnenberg, B.; Cross, H.J.; O’Callaghan, F.; De Giorgis, V.; Jansen, F.E.; Perucca, E.; et al. Toward responsible clinical n-of-1 strategies for rare diseases. *Drug Discov. Today* **2023**, *28*, 103688. [[CrossRef](#)] [[PubMed](#)]
28. Davidson, K.W.; Silverstein, M.; Cheung, K.; Paluch, R.A.; Epstein, L.H. Experimental Designs to Optimize Treatments for Individuals: Personalized N-of-1 Trials. *JAMA Pediatr.* **2021**, *175*, 404–409. [[CrossRef](#)] [[PubMed](#)]
29. Hurvitz, N.; Azmanov, H.; Kesler, A.; Ilan, Y. Establishing a second-generation artificial intelligence-based system for improving diagnosis, treatment, and monitoring of patients with rare diseases. *Eur. J. Hum. Genet.* **2021**, *29*, 1485–1490. [[CrossRef](#)]
30. Singh, K.; Usman, S.; Zeeshan, S.; Yanamala, N.; Nichols, M.; Bhise, V.; Bokhari, S.; Sengupta, P.P.; Ahmed, Z. Bioinformatics and AI/ML approaches using multi-omics data to accelerate diagnosis and delivery of precision care for patients with rare diseases. *Methods Cell Biol.* **2026**, *204*, 109–139. [[CrossRef](#)]
31. Zhao, W.; Wu, C.; Fan, Y.; Qiu, P.; Zhang, X.; Sun, Y.; Zhou, X.; Zhang, S.; Peng, Y.; Wang, Y.; et al. An agentic system for rare disease diagnosis with traceable reasoning. *Nature* **2026**, *651*, 775–784. [[CrossRef](#)] [[PubMed](#)]
32. De La Vega, F.M.; Chowdhury, S.; Moore, B.; Frise, E.; McCarthy, J.; Hernandez, E.J.; Wong, T.; James, K.; Guidugli, L.; Agrawal, P.B.; et al. Artificial intelligence enables comprehensive genome interpretation and nomination of candidate diagnoses for rare genetic diseases. *Genome Med.* **2021**, *13*, 153. [[CrossRef](#)] [[PubMed](#)]
33. Haendel, M.A.; Chute, C.G.; Robinson, P.N. Classification, Ontology, and Precision Medicine. *N. Engl. J. Med.* **2018**, *379*, 1452–1462. [[CrossRef](#)]
34. Mao, X.; Huang, Y.; Jin, Y.; Wang, L.; Chen, X.; Liu, H.; Yang, X.; Xu, H.; Luan, X.; Xiao, Y.; et al. A phenotype-based AI pipeline outperforms human experts in differentially diagnosing rare diseases using EHRs. *npj Digit. Med.* **2025**, *8*, 68. [[CrossRef](#)] [[PubMed](#)]
35. Poplin, R.; Chang, P.C.; Alexander, D.; Schwartz, S.; Colthurst, T.; Ku, A.; Newburger, D.; Dijamco, J.; Nguyen, N.; Afshar, P.T.; et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat. Biotechnol.* **2018**, *36*, 983–987. [[CrossRef](#)]

36. Cheng, J.; Novati, G.; Pan, J.; Bycroft, C.; Žemgulytė, A.; Applebaum, T.; Pritzel, A.; Wong, L.H.; Zielinski, M.; Sargeant, T.; et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **2023**, *381*, eadg7492. [[CrossRef](#)]
37. Ioannidis, N.M.; Rothstein, J.H.; Pejaver, V.; Middha, S.; McDonnell, S.K.; Baheti, S.; Musolf, A.; Li, Q.; Holzinger, E.; Karyadi, D.; et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* **2016**, *99*, 877–885. [[CrossRef](#)]
38. Cipriani, V.; Pontikos, N.; Arno, G.; Sergouniotis, P.I.; Lenassi, E.; Thawong, P.; Danis, D.; Michaelides, M.; Webster, A.R.; Moore, A.T.; et al. An Improved Phenotype-Driven Tool for Rare Mendelian Variant Prioritization: Benchmarking Exomiser on Real Patient Whole-Exome Data. *Genes* **2020**, *11*, 460. [[CrossRef](#)]
39. Lin, Y.L.; Chang, P.C.; Hsu, C.; Hung, M.Z.; Chien, Y.H.; Hwu, W.L.; Lai, F.; Lee, N.C. Comparison of GATK and DeepVariant by trio sequencing. *Sci. Rep.* **2022**, *12*, 1809. [[CrossRef](#)]
40. Pinto, V.; Sousa, L.; Silva, C. Variant calling in genomics: A comparative performance analysis and decision guide. *PLoS ONE* **2026**, *21*, e0339891. [[CrossRef](#)]
41. Rowlands, C.; Thomas, H.B.; Lord, J.; Wai, H.A.; Arno, G.; Beaman, G.; Sergouniotis, P.; Gomes-Silva, B.; Campbell, C.; Gossan, N.; et al. Comparison of in silico strategies to prioritize rare genomic variants impacting RNA splicing for the diagnosis of genomic disorders. *Sci. Rep.* **2021**, *11*, 20607. [[CrossRef](#)]
42. Moles-Fernández, A.; Domènech-Vivó, J.; Tenés, A.; Balmaña, J.; Diez, O.; Gutiérrez-Enríquez, S. Role of Splicing Regulatory Elements and In Silico Tools Usage in the Identification of Deep Intronic Splicing Variants in Hereditary Breast/Ovarian Cancer Genes. *Cancers* **2021**, *13*, 3341. [[CrossRef](#)]
43. Tian, Y.; Pesaran, T.; Chamberlin, A.; Fenwick, R.B.; Li, S.; Gau, C.L.; Chao, E.C.; Lu, H.M.; Black, M.H.; Qian, D. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. *Sci. Rep.* **2019**, *9*, 12752. [[CrossRef](#)]
44. Cooperstein, I.B.; Marwaha, S.; Ward, A.; Kobren, S.N.; Carter, J.N.; Wheeler, M.T.; Marth, G.T. An optimized variant prioritization process for rare disease diagnostics: Recommendations for Exomiser and Genomiser. *Genome Med.* **2025**, *17*, 127. [[CrossRef](#)] [[PubMed](#)]
45. Smedley, D.; Smith, K.R.; Martin, A.; Thomas, E.A.; McDonagh, E.M.; Cipriani, V.; Ellingford, J.M.; Arno, G.; Tucci, A.; Vandrovцова, J.; et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care—Preliminary Report. *N. Engl. J. Med.* **2021**, *385*, 1868–1880. [[CrossRef](#)] [[PubMed](#)]
46. Avsec, Ž.; Latysheva, N.; Cheng, J.; Novati, G.; Taylor, K.R.; Ward, T.; Bycroft, C.; Nicolaisen, L.; Arvaniti, E.; Pan, J.; et al. Advancing regulatory variant effect prediction with AlphaGenome. *Nature* **2026**, *649*, 1206–1218. [[CrossRef](#)]
47. García-González, J.; Gogolewski, K. AlphaGenome, a Swiss-army knife for exploring non-coding DNA. *Trends Genet.* **2026**, *42*, 4–6. [[CrossRef](#)]
48. Dolgin, E. Bringing AI to Cancer’s Regulatory Genome. *Cancer Discov.* **2026**, *16*, Of1. [[CrossRef](#)]
49. Lunke, S.; Bouffler, S.E.; Patel, C.V.; Sandaradura, S.A.; Wilson, M.; Pinner, J.; Hunter, M.F.; Barnett, C.P.; Wallis, M.; Kamien, B.; et al. Integrated multi-omics for rapid rare disease diagnosis on a national scale. *Nat. Med.* **2023**, *29*, 1681–1691. [[CrossRef](#)]
50. Keefe, A.C.; Scott, A.A.; Kruidenier, L.; Conta, J.; Sternen, D.; Clowes Candadai, S.; Stasi, S.M.; Parish-Morris, J.; Sikes, M.; Adam, M.P.; et al. Implementation of First-Line Rapid Genome Sequencing in Non-Critical Care Pediatric Wards. *J. Pediatr.* **2025**, *286*, 114699. [[CrossRef](#)] [[PubMed](#)]
51. Stenton, S.L.; O’Leary, M.C.; Lemire, G.; VanNoy, G.E.; DiTroia, S.; Ganesh, V.S.; Groopman, E.; O’Heir, E.; Mangilog, B.; Osei-Owusu, I.; et al. Critical assessment of variant prioritization methods for rare disease diagnosis within the rare genomes project. *Hum. Genom.* **2024**, *18*, 44. [[CrossRef](#)]
52. Tosco-Herrera, E.; Muñoz-Barrera, A.; Jáspez, D.; Rubio-Rodríguez, L.A.; Mendoza-Alvarez, A.; Rodriguez-Perez, H.; Jou, J.; Iñigo-Campos, A.; Corrales, A.; Ciuffreda, L.; et al. Evaluation of a whole-exome sequencing pipeline and benchmarking of causal germline variant prioritizers. *Hum. Mutat.* **2022**, *43*, 2010–2020. [[CrossRef](#)]
53. Ostrer, H. When Rare Is Common—Rethinking Pediatric Genetic Care. *JAMA* **2026**, *335*, 1293–1294. [[CrossRef](#)]
54. Wojcik, M.H.; Larkin, K.; Cipicchio, M.; Doupnik, A.; Zhao, C.; Cech, C.; Lopez, D.; Chandrasekar, J.; Leadbetter, J.; Mannion, J.; et al. Toward Same-Day Genome Sequencing in the Critical Care Setting. *N. Engl. J. Med.* **2025**, *393*, 2063–2065. [[CrossRef](#)]
55. Becherucci, F.; Landini, S.; Palazzo, V.; Cirillo, L.; Raglianti, V.; Lugli, G.; Tiberi, L.; Dirupo, E.; Bellelli, S.; Mazzierli, T.; et al. A Clinical Workflow for Cost-Saving High-Rate Diagnosis of Genetic Kidney Diseases. *J. Am. Soc. Nephrol.* **2023**, *34*, 706–720. [[CrossRef](#)] [[PubMed](#)]
56. Groopman, E.E.; Marasa, M.; Cameron-Christie, S.; Petrovski, S.; Aggarwal, V.S.; Milo-Rasouly, H.; Li, Y.; Zhang, J.; Nestor, J.; Krithivasan, P.; et al. Diagnostic Utility of Exome Sequencing for Kidney Disease. *N. Engl. J. Med.* **2019**, *380*, 142–151. [[CrossRef](#)]
57. Amr, K.S.; El-Bassyouni, H.T.; Rabie, E.; Fadel, F.I.; Teama, N.M.; Selim, A.; Hamed, K.; Helmy, R.; El-Salam, M.A.; Ahmed, H.A. Diagnostic utility of whole exome sequencing in pediatric and adult patients with suspected monogenic chronic kidney disease. *Mol. Biol. Rep.* **2025**, *53*, 96. [[CrossRef](#)]
58. Vivante, A. Genetics of Chronic Kidney Disease. *N. Engl. J. Med.* **2024**, *391*, 627–639. [[CrossRef](#)] [[PubMed](#)]

59. de Haan, A.; Eijgelsheim, M.; Vogt, L.; Hoorn, E.J.; Rotmans, J.I.; Fortrie, G.; Marsman, R.F.J.; Rothuizen, T.C.; Spijker, H.S.; Claus, L.R.; et al. Genetic testing in a national cohort of adults with chronic kidney disease of unknown origin. *Nephrol. Dial. Transplant.* **2025**, *40*, 1225–1233. [[CrossRef](#)] [[PubMed](#)]
60. Chen, X.; Burgun, A.; Boyer, O.; Knebelmann, B.; Garcelon, N. Artificial intelligence and perspective for rare genetic kidney diseases. *Kidney Int.* **2025**, *108*, 388–393. [[CrossRef](#)]
61. Tangri, N.; Cheungpasitporn, W.; Crittenden, S.D.; Fornoni, A.; Peralta, C.A.; Singh, K.; Usvyat, L.A.; Waterman, A.D. Responsible Use of Artificial Intelligence to Improve Kidney Care: A Statement from the American Society of Nephrology. *J. Am. Soc. Nephrol.* **2026**, *37*, 881–890. [[CrossRef](#)]
62. Thompson, R.; Spendiff, S.; Roos, A.; Bourque, P.R.; Warman Chardon, J.; Kirschner, J.; Horvath, R.; Lochmüller, H. Advances in the diagnosis of inherited neuromuscular diseases and implications for therapy development. *Lancet Neurol.* **2020**, *19*, 522–532. [[CrossRef](#)]
63. Rexach, J.; Lee, H.; Martinez-Agosto, J.A.; Németh, A.H.; Fogel, B.L. Clinical application of next-generation sequencing to the practice of neurology. *Lancet Neurol.* **2019**, *18*, 492–503. [[CrossRef](#)]
64. Watson, M.S.; Epstein, C.; Howell, R.R.; Jones, M.C.; Korf, B.R.; McCabe, E.R.; Simpson, J.L. Developing a national collaborative study system for rare genetic diseases. *Genet. Med.* **2008**, *10*, 325–329. [[CrossRef](#)]
65. Calvino, G.; Peconi, C.; Strafella, C.; Trastulli, G.; Megalizzi, D.; Andreucci, S.; Cascella, R.; Caltagirone, C.; Zampatti, S.; Giardina, E. Federated Learning: Breaking Down Barriers in Global Genomic Research. *Genes* **2024**, *15*, 1650. [[CrossRef](#)]
66. Montalvo, N.; Requena, F.; Capriotti, E.; Rausell, A. Federated learning for the pathogenicity annotation of genetic variants in multi-site clinical settings. *Bioinformatics* **2025**, *41*, btaf523. [[CrossRef](#)]
67. Pati, S.; Baid, U.; Edwards, B.; Sheller, M.; Wang, S.H.; Reina, G.A.; Foley, P.; Gruzdev, A.; Karkada, D.; Davatzikos, C.; et al. Federated learning enables big data for rare cancer boundary detection. *Nat. Commun.* **2022**, *13*, 7346. [[CrossRef](#)]
68. Sheller, M.J.; Edwards, B.; Reina, G.A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.R.; et al. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Sci. Rep.* **2020**, *10*, 12598. [[CrossRef](#)]
69. Gu, X.; Sabrina, F.; Fan, Z.; Sohail, S. A Review of Privacy Enhancement Methods for Federated Learning in Healthcare Systems. *Int. J. Environ. Res. Public Health* **2023**, *20*, 6539. [[CrossRef](#)]
70. Yurdem, B.; Kuzlu, M.; Gullu, M.K.; Catak, F.O.; Tabassum, M. Federated learning: Overview, strategies, applications, tools and future directions. *Heliyon* **2024**, *10*, e38137. [[CrossRef](#)]
71. Brauneck, A.; Schmalhorst, L.; Kazemi Majdabadi, M.M.; Bakhtiari, M.; Völker, U.; Baumbach, J.; Baumbach, L.; Buchholtz, G. Federated Machine Learning, Privacy-Enhancing Technologies, and Data Protection Laws in Medical Research: Scoping Review. *J. Med. Internet Res.* **2023**, *25*, e41588. [[CrossRef](#)]
72. Sarrat-González, D.; Escribà-Montagut, X.; Houghtaling, J.; González, J.R. dsOMOP: Bridging OMOP CDM and DataSHIELD for secure federated analysis of standardized clinical data. *Bioinformatics* **2025**, *41*, btaf286. [[CrossRef](#)]
73. Mateus, P.; Moonen, J.; Beran, M.; Jaarsma, E.; van der Landen, S.M.; Heuvelink, J.; Birhanu, M.; Harms, A.G.J.; Bron, E.; Wolters, F.J.; et al. Data harmonization and federated learning for multi-cohort dementia research using the OMOP common data model: A Netherlands consortium of dementia cohorts case study. *J. Biomed. Inform.* **2024**, *155*, 104661. [[CrossRef](#)]
74. Reich, C.; Ostropolets, A.; Ryan, P.; Rijnbeek, P.; Schuemie, M.; Davydov, A.; Dymshyts, D.; Hripcsak, G. OHDSI Standardized Vocabularies—a large-scale centralized reference ontology for international data harmonization. *J. Am. Med. Inform. Assoc.* **2024**, *31*, 583–590. [[CrossRef](#)]
75. van Baalen, V.; Didden, E.M.; Rosenberg, D.; Bardenheuer, K.; van Speybroeck, M.; Brand, M. Increase transparency and reproducibility of real-world evidence in rare diseases through disease-specific Federated Data Networks. *Pharmacoepidemiol. Drug Saf.* **2024**, *33*, e5778. [[CrossRef](#)]
76. Cremonesi, F.; Planat, V.; Kalokyri, V.; Kondylakis, H.; Sanavia, T.; Miguel Mateos Resinas, V.; Singh, B.; Uribe, S. The need for multimodal health data modeling: A practical approach for a federated-learning healthcare platform. *J. Biomed. Inform.* **2023**, *141*, 104338. [[CrossRef](#)]
77. Dette, C.M.U.; Alberg, V.; Rüdeshheim, S.; Selzer, D.; Marok, F.Z.; Bragazzi, N.L.; Fuhr, L.M.; Brunak, S.; Pearson, E.R.; Zahn, T.; et al. Advancing rare disease therapeutics through digital twins: Opportunities in drug development and precision dosing. *Comput. Struct. Biotechnol. J.* **2025**, *28*, 592–608. [[CrossRef](#)]
78. Drummond, D.; Gonsard, A. Definitions and Characteristics of Patient Digital Twins Being Developed for Clinical Use: Scoping Review. *J. Med. Internet Res.* **2024**, *26*, e58504. [[CrossRef](#)]
79. Vallée, A. Digital Twins for Personalized Medicine Require Epidemiological Data and Mathematical Modeling: Viewpoint. *J. Med. Internet Res.* **2025**, *27*, e72411. [[CrossRef](#)]
80. Masison, J.; Beezley, J.; Mei, Y.; Ribeiro, H.; Knapp, A.C.; Sordo Vieira, L.; Adhikari, B.; Scindia, Y.; Grauer, M.; Helba, B.; et al. A modular computational framework for medical digital twins. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2024287118. [[CrossRef](#)]

81. Venkatesh, K.P.; Brito, G.; Kamel Boulos, M.N. Health Digital Twins in Life Science and Health Care Innovation. *Annu. Rev. Pharmacol. Toxicol.* **2024**, *64*, 159–170. [[CrossRef](#)]
82. Colwill, M.; Honap, S.; Young, A.M.; Magro, F.; Jairath, V.; Danese, S.; Peyrin-Biroulet, L. Digital twins in IBD-bridging data, biology and trial innovation. *Gut* **2026**. [[CrossRef](#)]
83. Ringeval, M.; Etindele Sosso, F.A.; Cousineau, M.; Paré, G. Advancing Health Care With Digital Twins: Meta-Review of Applications and Implementation Challenges. *J. Med. Internet Res.* **2025**, *27*, e69544. [[CrossRef](#)]
84. Spreafico, A.M.; Neri, L.; Angel, K.; Bellettato, C.M.; Scarpa, M. Artificial Intelligence, Connected Care, and Enabling Digital Health Technologies in Rare Diseases With a Focus on Lysosomal Storage Disorders: Scoping Review. *J. Med. Internet Res.* **2026**, *28*, e73612. [[CrossRef](#)] [[PubMed](#)]
85. Li, X.; Loscalzo, J.; Mahmud, A.; Aly, D.M.; Rzhetsky, A.; Zitnik, M.; Benson, M. Digital twins as global learning health and disease models for preventive and personalized medicine. *Genome Med.* **2025**, *17*, 11. [[CrossRef](#)]
86. Venkatesh, K.P.; Raza, M.M.; Kvedar, J.C. Health digital twins as tools for precision medicine: Considerations for computation, implementation, and regulation. *npj Digit. Med.* **2022**, *5*, 150. [[CrossRef](#)]
87. Eckardt, J.N.; Hahn, W.; Prelaj, A.; Bornhäuser, M.; Middeke, J.M.; Kather, J.N. Artificial intelligence-generated synthetic data for cancer research and clinical trials. *Nat. Rev. Cancer* **2026**, *26*, 351–363. [[CrossRef](#)] [[PubMed](#)]
88. Trabassi, D.; Castiglia, S.F.; Bini, F.; Marinozzi, F.; Ajoudani, A.; Lorenzini, M.; Chini, G.; Varrecchia, T.; Ranavolo, A.; De Icco, R.; et al. Optimizing Rare Disease Gait Classification through Data Balancing and Generative AI: Insights from Hereditary Cerebellar Ataxia. *Sensors* **2024**, *24*, 3613. [[CrossRef](#)]
89. Sun, C.; van Soest, J.; Dumontier, M. Generating synthetic personal health data using conditional generative adversarial networks combining with differential privacy. *J. Biomed. Inform.* **2023**, *143*, 104404. [[CrossRef](#)] [[PubMed](#)]
90. Al-Dhamari, I.; Abu Attieh, H.; Prasser, F. Synthetic datasets for open software development in rare disease research. *Orphanet J. Rare Dis.* **2024**, *19*, 265. [[CrossRef](#)]
91. Khosravi, B.; Purkayastha, S.; Erickson, B.J.; Trivedi, H.M.; Gichoya, J.W. Exploring the potential of generative artificial intelligence in medical image synthesis: Opportunities, challenges, and future directions. *Lancet Digit. Health* **2025**, *7*, 100890. [[CrossRef](#)]
92. Chen, J.; Chun, D.; Patel, M.; Chiang, E.; James, J. The validity of synthetic clinical data: A validation study of a leading synthetic data generator (Synthea) using clinical quality measures. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 44. [[CrossRef](#)]
93. Thongprayoon, C.; Pesce, F.; Cheungpasitporn, W. Clinical Artificial Intelligence Agents in Nephrology: From Prediction to Action Through Workflow-Native Intelligence-A Roadmap for Workflow-Integrated Care. *J. Clin. Med.* **2026**, *15*, 2576. [[CrossRef](#)]
94. Solomonides, A.E.; Koski, E.; Atabaki, S.M.; Weinberg, S.; McGreevey, J.D.; Kannry, J.L.; Petersen, C.; Lehmann, C.U. Defining AMIA's artificial intelligence principles. *J. Am. Med. Inform. Assoc.* **2022**, *29*, 585–591. [[CrossRef](#)]
95. Palmieri, S.; Robertson, C.T.; Cohen, I.G. New Guidance on Responsible Use of AI. *JAMA* **2026**, *335*, 207–208. [[CrossRef](#)]
96. Subasri, V.; Krishnan, A.; Kore, A.; Dhalla, A.; Pandya, D.; Wang, B.; Malkin, D.; Razak, F.; Verma, A.A.; Goldenberg, A.; et al. Detecting and Remediating Harmful Data Shifts for the Responsible Deployment of Clinical AI Models. *JAMA Netw. Open* **2025**, *8*, e2513685. [[CrossRef](#)] [[PubMed](#)]
97. Guan, H.; Bates, D.; Zhou, L. Keeping Medical AI Healthy and Trustworthy: A Review of Detection and Correction Methods for System Degradation. *IEEE Trans. Biomed. Eng.* **2025**. [[CrossRef](#)] [[PubMed](#)]
98. Guo, L.L.; Pfohl, S.R.; Fries, J.; Johnson, A.E.W.; Posada, J.; Aftandilian, C.; Shah, N.; Sung, L. Evaluation of domain generalization and adaptation on improving model robustness to temporal dataset shift in clinical medicine. *Sci. Rep.* **2022**, *12*, 2726. [[CrossRef](#)] [[PubMed](#)]
99. Eden, R.; Chukwudi, I.; Bain, C.; Barbieri, S.; Callaway, L.; de Jersey, S.; George, Y.; Gorse, A.D.; Lawley, M.; Marendy, P.; et al. A scoping review of the governance of federated learning in healthcare. *npj Digit. Med.* **2025**, *8*, 427. [[CrossRef](#)]
100. Kommusaar, J.; Elunurm, S.; Chomutare, T.; Kangasniemi, M.; Salanterä, S.; Peltonen, L.M. A roadmap for federated learning projects using health data to guide sustainable artificial intelligence development in the European Union. *Int. J. Med. Inform.* **2026**, *208*, 106242. [[CrossRef](#)]
101. Nisevic, M.; Milojevic, D.; Spajic, D. Synthetic data in medicine: Legal and ethical considerations for patient profiling. *Comput. Struct. Biotechnol. J.* **2025**, *28*, 190–198. [[CrossRef](#)]
102. Giuffrè, M.; Shung, D.L. Harnessing the power of synthetic data in healthcare: Innovation, application, and privacy. *npj Digit. Med.* **2023**, *6*, 186. [[CrossRef](#)] [[PubMed](#)]
103. Susser, D.; Schiff, D.S.; Gerke, S.; Cabrera, L.Y.; Cohen, I.G.; Doerr, M.; Harrod, J.; Kostick-Quenet, K.; McNealy, J.; Meyer, M.N.; et al. Synthetic Health Data: Real Ethical Promise and Peril. *Hast. Cent. Rep.* **2024**, *54*, 8–13. [[CrossRef](#)]
104. Pezoulas, V.C.; Zaridis, D.I.; Mylona, E.; Androutsos, C.; Apostolidis, K.; Tachos, N.S.; Fotiadis, D.I. Synthetic data generation methods in healthcare: A review on open-source tools and methods. *Comput. Struct. Biotechnol. J.* **2024**, *23*, 2892–2910. [[CrossRef](#)]
105. Theodorou, B.; Xiao, C.; Sun, J. Synthesize high-dimensional longitudinal electronic health records via hierarchical autoregressive language model. *Nat. Commun.* **2023**, *14*, 5305. [[CrossRef](#)]

106. Pillay, T.S.; van Deventer, B.S.; Gwiliza, S.; Subramoney, E.L.; van Niekerk, C. Synthetic data in the clinical laboratory: Methods, applications, and future prospects. *Clin. Chim. Acta* **2026**, *585*, 120878. [[CrossRef](#)] [[PubMed](#)]
107. Kuo, N.I.; Garcia, F.; Sönnnerborg, A.; Böhm, M.; Kaiser, R.; Zazzi, M.; Polizzotto, M.; Jorm, L.; Barbieri, S. Generating synthetic clinical data that capture class imbalanced distributions with generative adversarial networks: Example using antiretroviral therapy for HIV. *J. Biomed. Inform.* **2023**, *144*, 104436. [[CrossRef](#)]
108. El Emam, K.; Mosquera, L.; Fang, X.; El-Hussuna, A. An evaluation of the replicability of analyses using synthetic health data. *Sci. Rep.* **2024**, *14*, 6978. [[CrossRef](#)]
109. Chen, X.; Wu, Z.; Shi, X.; Cho, H.; Mukherjee, B. Generating synthetic electronic health record data: A methodological scoping review with benchmarking on phenotype data and open-source software. *J. Am. Med. Inform. Assoc.* **2025**, *32*, 1227–1240. [[CrossRef](#)]
110. Kaur, D.; Sobiesk, M.; Patil, S.; Liu, J.; Bhagat, P.; Gupta, A.; Markuzon, N. Application of Bayesian networks to generate synthetic health data. *J. Am. Med. Inform. Assoc.* **2021**, *28*, 801–811. [[CrossRef](#)] [[PubMed](#)]
111. Tucker, A.; Wang, Z.; Rotalinti, Y.; Myles, P. Generating high-fidelity synthetic patient data for assessing machine learning healthcare software. *npj Digit. Med.* **2020**, *3*, 147. [[CrossRef](#)]
112. Blagec, K.; Kraiger, J.; Frühwirt, W.; Samwald, M. Benchmark datasets driving artificial intelligence development fail to capture the needs of medical professionals. *J. Biomed. Inform.* **2023**, *137*, 104274. [[CrossRef](#)]
113. Odone, A.; Barbati, C.; Amadasi, S.; Schultz, T.; Resnik, D.B. Artificial intelligence and infectious diseases: An evidence-driven conceptual framework for research, public health, and clinical practice. *Lancet Infect. Dis.* **2026**, *26*, e152–e167. [[CrossRef](#)] [[PubMed](#)]
114. Radjasandirane, R.; Diharce, J.; Gelly, J.C.; de Brevern, A.G. Insights for variant clinical interpretation based on a benchmark of 65 variant effect predictors. *Genomics* **2025**, *117*, 111036. [[CrossRef](#)]
115. Özçetin, E.; Baş, S.E.; Özpays, F. Architectural and translational perspectives on clinical decision support systems for rare disease diagnosis: A scoping review. *Int. J. Med. Inform.* **2026**, *215*, 106442. [[CrossRef](#)]
116. Norén, G.N.; Meldau, E.L.; Ellenius, J. Critical Appraisal of Artificial Intelligence for Rare-Event Recognition: Principles and Pharmacovigilance Case Studies. *Drug Saf.* **2026**, 1–14. [[CrossRef](#)]
117. Potnis, K.C.; Ross, J.S.; Aneja, S.; Gross, C.P.; Richman, I.B. Artificial Intelligence in Breast Cancer Screening: Evaluation of FDA Device Regulation and Future Recommendations. *JAMA Intern. Med.* **2022**, *182*, 1306–1312. [[CrossRef](#)]
118. Gilbert, S.; Fenech, M.; Hirsch, M.; Upadhyay, S.; Biasiucci, A.; Starlinger, J. Algorithm Change Protocols in the Regulation of Adaptive Machine Learning-Based Medical Devices. *J. Med. Internet Res.* **2021**, *23*, e30545. [[CrossRef](#)] [[PubMed](#)]
119. Clark, P.; Kim, J.; Aphinyanaphongs, Y. Marketing and US Food and Drug Administration Clearance of Artificial Intelligence and Machine Learning Enabled Software in and as Medical Devices: A Systematic Review. *JAMA Netw. Open* **2023**, *6*, e2321792. [[CrossRef](#)]
120. DuPreez, J.A.; McDermott, O. The use of predetermined change control plans to enable the release of new versions of software as a medical device. *Expert. Rev. Med. Devices* **2025**, *22*, 261–275. [[CrossRef](#)]
121. Shanmugam, U.; Rajendran, M.K.; Natarajan, J.; Karri, V. Clinical Trial Design and Regulatory Requirements for Artificial Intelligence as a Medical Device: A PRISMA-ScR-Guided Scoping Review of Global Guidance and Evidence (2017–2025). *J. Clin. Med.* **2026**, *15*, 1937. [[CrossRef](#)]
122. Nugent, B.M.; Ramamoorthy, A.; Pippins, J.R.; Pepe, S.; Doi, M.; Thomas, A.; Bagheri, B.; Welsh, C.; Madabushi, R.; Lee, K.J. Confirmatory Evidence Used in Non-Oncologic Rare Disease New Molecular Entity Marketing Applications Approved by FDA, 2020–2023. *Clin. Pharmacol. Ther.* **2025**, *117*, 1627–1631. [[CrossRef](#)]
123. Wu, J.; Wang, C.; Toh, S.; Pisa, F.E.; Bauer, L. Use of real-world evidence in regulatory decisions for rare diseases in the United States-Current status and future directions. *Pharmacoepidemiol. Drug Saf.* **2020**, *29*, 1213–1218. [[CrossRef](#)]
124. Kim, M.; Hong, E. Analysis of Real-World Data Utilization in the Orphan Drug Approval Process: Focusing on New Drug Marketing Applications Submitted to the FDA. *Clin. Pharmacol. Ther.* **2026**, *119*, 1340–1349. [[CrossRef](#)]
125. Akodad, S.; Haon, B.; Rochaix, L.; Stevens, H. Orphan medical devices: Addressing the regulatory and access gaps in the EU and US. *Front. Public Health* **2025**, *13*, 1729821. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.