rheumato

MDPI

Review

Emerging AI- and Biomarker-Driven Precision Medicine in Autoimmune Rheumatic Diseases: From Diagnostics to Therapeutic Decision-Making

Ola A. Al-Ewaidat [1] and Moawiah M. Naffaa [2,*]

[1] Department of Internal Medicine, Stanford University School of Medicine, Palo Alto, CA 94305, USA
[2] Independent Researcher, Mountain View, CA 94040, USA
* Correspondence: moawiahn@hotmail.com

Abstract

Background/Objectives: Autoimmune rheumatic diseases (AIRDs) are complex, heterogeneous, and relapsing–remitting conditions in which early diagnosis, flare prediction, and individualized therapy remain major unmet needs. This review aims to synthesize recent progress in AI-driven, biomarker-based precision medicine, integrating advances in imaging, multi-omics, and digital health to enhance diagnosis, risk stratification, and therapeutic decision-making in AIRD. **Methods:** A comprehensive synthesis of 2020–2025 literature was conducted across PubMed, Scopus, and preprint databases, focusing on studies applying artificial intelligence, machine learning, and multimodal biomarkers in rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, spondyloarthritis, and related autoimmune diseases. The review emphasizes methodological rigor (TRIPOD+AI, PROBAST+AI, CONSORT-AI/SPIRIT-AI), implementation infrastructures (ACR RISE registry, federated learning), and equity frameworks to ensure generalizable, safe, and ethically governed translation into clinical practice. **Results:** Emerging evidence demonstrates that AI-integrated imaging enables automated quantification of synovitis, erosions, and vascular inflammation; multi-omics stratification reveals interferon- and B-cell-related molecular programs predictive of therapeutic response; and digital biomarkers from wearables and smartphones extend monitoring beyond the clinic, capturing early flare signatures. Registry-based AI pipelines and federated collaboration now allow multicenter model training without compromising patient privacy. Across diseases, predictive frameworks for biologic and Janus kinase (JAK) inhibitor response show growing discriminatory performance, though prospective and equity-aware validation remain limited. **Conclusions:** AI-enabled fusion of imaging, molecular, and digital biomarkers is reshaping the diagnostic and therapeutic landscape of AIRD. Standardized validation, interoperability, and governance frameworks are essential to transition these tools from research to real-world precision rheumatology. The convergence of registries, federated learning, and transparent reporting standards marks a pivotal step toward pragmatic, equitable, and continuously learning systems of care.

Keywords: autoimmune rheumatic diseases; precision medicine; artificial intelligence; biomarkers; omics stratification; digital health

check for updates

Academic Editor: Bruce M. Rothschild

Received: 2 October 2025
Revised: 3 November 2025
Accepted: 12 November 2025
Published: 17 November 2025

Citation: Al-Ewaidat, O.A.; Naffaa, M.M. Emerging AI- and Biomarker-Driven Precision Medicine in Autoimmune Rheumatic Diseases: From Diagnostics to Therapeutic Decision-Making. *Rheumato* **2025**, *5*, 17. https://doi.org/10.3390/rheumato5040017

Copyright: © 2025 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/).

1. Background

Autoimmune rheumatic diseases (AIRDs) are clinically and biologically heterogeneous, characterized by overlapping phenotypes, fluctuating disease activity, and variable

therapeutic responses [1,2]. This complexity renders early diagnosis, prognostication, and individualized treatment particularly challenging when relying on conventional single-marker heuristics. Recent reviews emphasize that artificial intelligence (AI) is uniquely positioned to address these challenges by modeling the nonlinear, multimodal structure of AIRDs—provided that data quality, transparent governance, and rigorous validation are ensured [3,4].

Among AIRDs, rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) represent prototypical yet contrasting disease archetypes. RA is characterized by chronic synovial inflammation driven by autoreactive B and T cells, pro-inflammatory cytokines such as TNF, IL-6, and IL-1, and progressive joint destruction mediated by osteoclast activation. SLE, in contrast, arises from loss of immune tolerance, immune-complex deposition, complement activation, and type I interferon overproduction that drives multi-organ involvement [5,6]. Despite advances in diagnostics—including serologic markers such as RF, ACPA, ANA, and anti-dsDNA—both disorders exhibit substantial clinical overlap, fluctuating phenotypes, and seronegative or atypical presentations that delay definitive diagnosis [7,8]. Current therapies range from conventional and biologic DMARDs to JAK inhibitors in RA and B-cell- or interferon-targeted biologics in SLE, yet variable treatment response and relapse remain major unmet needs [9,10]. These challenges underscore the necessity of multidimensional approaches capable of integrating molecular, imaging, and behavioral data to refine disease classification and personalize therapy, an area where AI and biomarker-driven strategies are beginning to demonstrate transformative potential.

Three converging developments have markedly advanced the feasibility of AI-driven precision medicine in AIRD. First, deep learning-based image analysis now enables automated, quantitative assessment of musculoskeletal ultrasound and MRI, producing standardized measures of synovitis, erosions, and joint space narrowing that improve reproducibility across centers [11]. Second, advances in multi-omics profiling have revealed interferon-driven and B cell-enriched molecular programs in SLE and related AIRDs, refining disease subtypes and predicting differential responses to targeted therapies such as anifrolumab or B cell-directed agents [8,12]. Third, continuous digital phenotyping through smartphones and wearable devices allows longitudinal tracking of mobility, sleep, and symptom trajectories, augmenting traditional clinic-based indices and enabling earlier detection of flare risk in rheumatoid arthritis and other immune-mediated inflammatory diseases [13,14].

The AI tasks with greatest translational potential in AIRD include diagnostic support and triage, disease-activity and flare prediction, and treatment-response modeling for biologic DMARDs and JAK inhibitors. Current best practices employ interpretable ensemble learners such as gradient boosting for registry and Electronic Health Record (HER) data; convolutional neural networks and transformers for imaging and time-series data; and multimodal fusion frameworks for integrating -omics, imaging, and digital phenotyping streams. To ensure clinical reliability, these applications must be accompanied by robust calibration, uncertainty quantification, and external or temporal validation [15,16].

Equally critical to progress are the infrastructures that enable implementation. The ACR RISE registry, a large-scale, EHR-enabled quality registry encompassing millions of patients, has become a pivotal substrate for model development, deployment, and post-deployment monitoring. Recent analyses demonstrate that active engagement with RISE dashboards is associated with measurable improvements in clinical quality metrics, underscoring its role as an implementation backbone for precision rheumatology [17,18]. To enable generalization across institutions without centralizing sensitive patient data, federated learning (FL) approaches are increasingly adopted. These methods are supported by governance frameworks that formalize data-sharing agreements, auditability standards, and model-card reporting to ensure transparency and safety in clinical use [19,20].

Translation of AI tools from research to practice further depends on adherence to rigorous reporting and bias-assessment frameworks. The TRIPOD+AI guideline establishes minimum requirements for transparent reporting of AI-based prediction models [21,22]. Complementarily, PROBAST+AI provides structured tools for assessing risk of bias and applicability [21], while CONSORT-AI and SPIRIT-AI extend standards for trial reporting and protocol design [23,24]. Together, these instruments form the methodological foundation for the trustworthy evaluation of precision tools in rheumatology. Importantly, given the well-documented influence of ancestry, sex, and socioeconomic context on disease biology, phenotype expression, and care access, equity and subgroup generalizability must be treated as first-order design principles. Continuous recalibration and drift monitoring within registries such as RISE are therefore essential to maintain validity across heterogeneous populations and evolving care environments [25].

Bringing these strands together, the field now stands at a critical inflection point: multimodal biomarkers, advanced AI methodologies, and robust implementation infrastructures are converging to enable pragmatic precision rheumatology [26]. This review synthesizes recent advances, highlights validated exemplars, and delineates the standards and governance practices required to move AIRD care from conceptual promise to routine, clinic-embedded decision support.

This review adopts a narrative approach designed to synthesize recent translational advances in AI-driven biomarker discovery and precision medicine within autoimmune rheumatic diseases. Relevant literature was identified through PubMed and Scopus searches spanning 2015–2025, restricted to articles published in English. Search terms incorporated combinations of MeSH descriptors including "artificial intelligence," "machine learning," "biomarkers," "multi-omics," "autoimmune rheumatic diseases," "rheumatoid arthritis," and "systemic lupus erythematosus." Eligible studies comprised mechanistic, translational, and clinical investigations applying AI or advanced analytical methods to diagnostic, prognostic, or therapeutic modeling in AIRD. Excluded materials encompassed anecdotal case reports, and purely genetic-association studies without clinical modeling components. This narrative rather than systematic format allows for critical integration of heterogeneous evidence while maintaining focus on clinically actionable insights and methodological innovation.

2. Biomarker Evolution in Autoimmune Rheumatic Diseases

2.1. Classic Biomarkers—Autoantibodies and Inflammatory Markers: Limitations and Drift

Classic serologic biomarkers—including rheumatoid factor (RF), anti-citrullinated protein antibodies (ACPA), antinuclear antibodies (ANA), and inflammatory markers such as erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP)—remain fundamental to AIRD diagnosis and activity monitoring (Figure 1). Elevated RF titers are consistently associated with more severe RA phenotypes, including extra-articular involvement and reduced responsiveness to TNF-inhibitor therapy, underscoring RF's continued clinical relevance decades after its discovery [27,28]. Additional isotypes, particularly RF-IgA and ACPA-IgA, have been linked to poorer TNF-inhibitor outcomes and enhanced neutrophil extracellular trap formation, suggesting more profound engagement in RA pathogenesis [29,30]. Nevertheless, RF and ACPA alone fail to capture the full mechanistic heterogeneity of disease or predict therapeutic response with high fidelity.

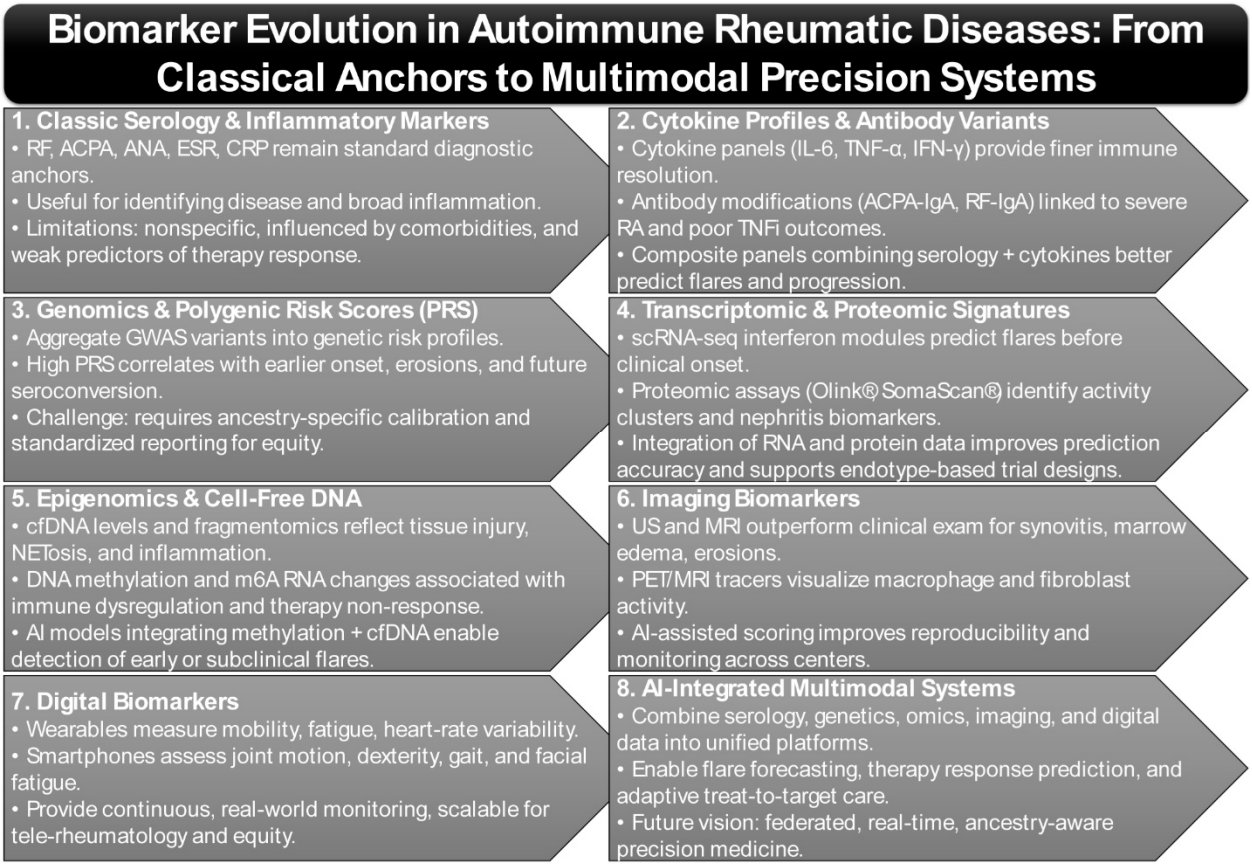


Figure 1. Biomarker Landscape Across AIRDs: From Classical Anchors to Multimodal Precision in Autoimmune Rheumatic Diseases.

Recent studies emphasize that these classical markers, while indispensable, act more as "broad-spectrum indicators" rather than precise stratifiers of disease biology. For instance, longitudinal cohort analyses demonstrate that while ACPA positivity predicts erosive progression, it cannot reliably discriminate between patients who will remain refractory versus those who will achieve remission on biologic therapy. Similarly, CRP and ESR, although widely used as inflammatory proxies, are nonspecific and subject to modulation by comorbidities such as infection, obesity, and cardiovascular disease, thus limiting their interpretive precision [31–33].

Emerging data now advocates for a layered biomarker strategy: combining serologic markers with molecular correlates such as cytokine signatures (e.g., IL-6, TNF-α, IFN-γ), autoantibody glycosylation patterns, and proteomic fingerprints derived from high-throughput assays (Figure 1) [34,35]. This composite approach has demonstrated superior predictive power for flare risk, radiographic progression, and biologic drug discontinuation.

Recent syntheses emphasize that composite biomarker panels, which integrate serology with immune-complex quantification and additional molecular indicators, outperform single-marker approaches for forecasting disease activity and remission [36,37]. Such multiplexed approaches are increasingly being embedded into AI-driven algorithms, where serologic markers are not discarded but rather contextualized as part of broader multimodal inputs. This shift reframes classical biomarkers not as outdated relics but as essential "anchors" that, when fused with omics and digital data, yield precision-grade stratification tools for clinical decision-making.

2.2. Genomics & Polygenic Risk

Polygenic risk scores (PRSs) aggregate the cumulative burden of GWAS-identified variants to quantify genetic susceptibility in AIRDs. A recent multi-ancestry optimization demonstrated improved predictive capacity of RA PRS but underscored the necessity of ensuring equity and interpretability across diverse ancestry groups (Figure 1) [38]. Similarly, a Taiwanese population study showed that individuals in the highest PRS quartile were significantly more likely to be RF- and ACPA-positive, display bone erosions, and require advanced therapies, thereby directly linking PRS to disease severity and structural damage [39].

Beyond susceptibility, PRSs are increasingly being explored as predictors of disease course and therapeutic response. For example, recent analyses suggest that higher RA-PRS correlates with earlier disease onset, faster radiographic progression, and reduced likelihood of achieving drug-free remission, positioning PRS as potential tools for risk stratification at the preclinical and early-disease stages [39]. Importantly, PRSs have also been associated with subclinical autoimmunity, where elevated scores predict future seroconversion of ACPA and RF in at-risk cohorts [39,40], thus offering a genetic "early warning" system for preventive interventions.

While these studies demonstrate consistent directionality and moderate effect sizes (typically explaining 8–15% of variance), most existing RA PRSs have been developed and optimized using predominantly European-ancestry GWAS datasets, with emerging validation in East Asian cohorts beginning to improve cross-population performance [41]. Predictive power and calibration, however, remain variable across ancestries, underscoring the need for broader representation in training datasets. Limited sample diversity and modest absolute risk separation currently constrain clinical translation, emphasizing that PRSs remain adjunctive rather than stand-alone tools.

Collectively, these findings highlight that although PRSs are promising, their clinical translation will require ancestry-specific calibration and integration with clinical, serologic, and environmental datasets to achieve real-world utility. Long-term implementation will also demand harmonized reporting standards, transparent benchmarking across ancestries, and embedding PRSs within federated, multi-site infrastructures to ensure both generalizability and equity.

2.3. Transcriptomic & Proteomic Signatures

Transcriptomic and proteomic profiling have generated powerful insights into disease heterogeneity and flare risk prediction, particularly in SLE. A longitudinal study in Asian SLE patients used phenome-wide causal proteomics with Mendelian randomization and machine learning to identify five key proteins—SAA1, B4GALT5, GIT2, NAA15, and RPIA—whose expression correlated strongly with one-year flare risk; a composite model integrating these proteins with clinical features achieved an AUC of 0.7 [42]. In parallel, a multi-omics screen of 121 SLE patients compared with healthy controls identified more than 90 differentially expressed proteins and 76 metabolites, including apolipoproteins and arachidonic acid derivatives, with strong correlations to disease activity and renal function; a subset of markers selected via random forest models yielded diagnostic AUCs of 0.86–0.90 [34].

Transcriptomic profiling has further delineated immune-cell-specific activation states that underpin flare dynamics. Single-cell RNA sequencing (scRNA-seq) studies have revealed aberrant type I interferon signatures in plasmacytoid dendritic cells and monocytes, alongside persistent activation of cytotoxic CD8+ T cells in patients with active SLE [43,44]. Longitudinal scRNA-seq has shown that these interferon-driven modules can precede clinical flares by weeks, highlighting their potential as predictive biomarkers. Furthermore, bulk RNA-seq studies consistently identify upregulated interferon-stimulated gene (ISG)

clusters, which not only stratify patients by disease activity but also predict responsiveness to targeted IFN-blocking therapies such as anifrolumab [45].

Proteomic investigations are increasingly complemented by high-throughput platforms such as Olink® proximity extension assays and SomaScan® aptamer-based profiling, which allow simultaneous quantification of thousands of circulating proteins at picogram sensitivity (Figure 1) [46]. Notably, proteomic signatures are proving useful in distinguishing lupus nephritis subtypes, where urinary proteomes (e.g., VCAM-1, NGAL, CD163) track intrarenal inflammation and may reduce reliance on repeat biopsies [47].

Integration of transcriptomic and proteomic layers has demonstrated synergistic value. Network-based models show that proteomic changes in the complement and coagulation cascades are tightly coupled with transcriptomic interferon signatures, pointing to shared upstream drivers of disease amplification [48]. Such integrative approaches also enable "endotype" discovery—subgrouping patients by molecular mechanism rather than clinical phenotype—which is now guiding early adaptive trial designs for precision therapeutics.

Despite high discriminatory performance, many studies remain limited by small cohort sizes (often <150 patients), lack of external validation, and variable assay reproducibility. Reported AUCs > 0.85 are encouraging but may overestimate real-world accuracy due to cross-validation bias [49]. Incorporating standardized pipelines and independent replication cohorts will be essential before these signatures can inform regulatory-grade diagnostics.

These findings demonstrate that integrative multi-omics approaches significantly outperform mono-omics strategies in classification, monitoring, and risk stratification in SLE. Going forward, embedding transcriptomic and proteomic biomarkers into federated learning frameworks, with continuous recalibration across diverse ancestries and treatment contexts, will be essential to move from discovery into clinically deployable decision-support tools.

2.4. Epigenomic Alterations and Cell-Free DNA/Fragmentomics as Emerging Biomarkers

Epigenetic dysregulation and cell-free DNA (cfDNA) signatures are rapidly emerging as minimally invasive, dynamic biomarkers in AIRDs [50]. cfDNA, derived from both nuclear and mitochondrial sources, reflects tissue injury, neutrophil extracellular trap (NET) activity, and systemic inflammation [51,52]. Studies report that elevated plasma cfDNA levels, particularly mitochondrial cfDNA, correlate with disease activity in RA and SLE and could serve as a complement to traditional inflammatory markers such as CRP and ESR for real-time disease monitoring [53,54]. Earlier foundational studies highlighted that cfDNA quantification is highly sensitive to pre-analytical variables, including sample handling, fragmentation bias, and contamination, necessitating rigorous standardization before clinical application [55–57].

Recent advances extend beyond absolute cfDNA concentration toward "fragm entomics"—the analysis of cfDNA fragment size distribution, genomic positioning, and nucleosomal occupancy patterns [58]. These signatures provide clues about tissue-of-origin, cell death pathways, and immune activation states. For example, RA patients demonstrate enrichment of neutrophil-derived cfDNA fragments, consistent with aberrant NETosis, while SLE cohorts exhibit cfDNA fragmentation patterns linked to lymphocyte and endothelial cell injury [59]. Emerging algorithms now integrate cfDNA methylation landscapes with fragmentomic features to improve sensitivity for detecting low-grade inflammation and organ-specific damage.

Epigenetically, multiple studies have reported consistent DNA methylation alterations and N6-methyladenosine (m6A) modifications in RA and SLE, both of which show potential as diagnostic classifiers and mechanistic readouts of immune dysregulation [60–63]. DNA methylation changes at immune regulatory loci (e.g., TNFAIP3, STAT4, IRF5) have

been linked to aberrant cytokine production and treatment non-response, while altered m6A RNA methylation patterns are increasingly implicated in dysregulated T- and B-cell differentiation [64,65]. These findings suggest that epigenomic markers are not only passive correlations of disease but may also represent causal drivers of autoimmune pathogenesis.

While fragmentomic and epigenomic approaches hold considerable promise for early disease detection and molecular stratification, their current translational readiness remains limited (technology readiness level ~3–4). Most published studies involve fewer than 100 participants and are based on retrospective or convenience sampling, which constrains reproducibility and effect-size precision [56,66]. Establishing standardized pre-analytical workflows, cross-platform benchmarking, and multi-center prospective validation will be essential prerequisites for clinical deployment.

Importantly, interpretable machine learning models are beginning to integrate such data. A multi-task deep learning system demonstrated efficacy in learning cross-disease methylation signatures that retain both predictive accuracy and biological interpretability across autoimmune phenotypes (Figure 1) [67]. Other computational pipelines now fuse cfDNA fragmentomics with methylome-wide profiles, enabling the detection of early disease transitions and subclinical flare states [68]. These approaches highlight the promise of dynamic, mechanism-aware biomarkers that could guide preemptive therapy escalation or tapering.

These advances position cfDNA fragmentomics and epigenomic readouts as central candidates for non-invasive, mechanism-aware disease stratification in AIRDs. Looking forward, standardizing analytic pipelines, embedding ancestry-aware epigenomic references, and validating models prospectively in large multi-center cohorts will be critical steps for clinical translation.

2.5. Imaging Biomarkers

Anatomical and molecular imaging have transformed the capacity to detect subclinical inflammation and structural progression in AIRDs. Ultrasound (US) and magnetic resonance imaging (MRI) remain indispensable tools for the detection and monitoring of inflammatory and structural changes in AIRD. US offers high sensitivity for synovitis, tenosynovitis, and superficial erosions, enabling dynamic bedside assessment of disease activity. However, it cannot directly visualize bone marrow edema, which is best characterized by MRI sequences sensitive to water content, such as STIR or T2-weighted fat-suppressed imaging. Together, these modalities provide complementary information that underpins diagnosis, treat-to-target monitoring, and early evaluation of therapeutic response [69,70]. Musculoskeletal imaging has reinforced that US and MRI consistently outperform clinical examination in quantifying inflammatory burden and predicting structural outcomes in RA and psoriatic arthritis [71,72].

Simplified scoring approaches such as RAMRIS-5 have been validated for use in both early and established RA, enabling semi-quantitative assessment of inflammation and joint damage with reduced burden compared to full OMERACT RAMRIS scoring [73,74]. In clinical trials, such simplified scoring systems have accelerated feasibility while retaining high sensitivity to change, thereby supporting their integration into adaptive trial designs and pragmatic real-world monitoring frameworks [75,76]. Importantly, these simplified indices are increasingly paired with AI-assisted image interpretation, which reduces inter-reader variability and enhances reproducibility across centers.

Beyond structural imaging, molecular imaging has advanced rapidly: novel radiotracers targeting activated macrophages or fibroblast-like synoviocytes in inflamed synovium now enable non-invasive visualization of disease-specific biology [69,77]. PET/MRI with tracers such as ˆ18F-FDG or macrophage-specific ligands can complement US/MRI by

providing metabolic signatures of synovitis, creating opportunities for multi-scale phenotyping and drug-response monitoring [78,79]. For instance, PET tracers binding to the folate receptor β on synovial macrophages allow distinction between inflamed and quiescent tissue, while novel fibroblast-activation protein (FAP) ligands provide unique readouts of stromal pathogenicity [80,81]. These molecular approaches are beginning to bridge the gap between static anatomic imaging and dynamic immune-pathobiology, enabling visualization of pathways directly targeted by emerging therapies.

Technological convergence is also driving next-generation imaging biomarkers. Hybrid modalities such as PET/MRI integrate high-resolution anatomic detail with metabolic and immunologic readouts in a single acquisition, while AI-enhanced US leverages automated Doppler signal quantification for real-time flare detection (Figure 1) [82]. Furthermore, machine learning models trained on imaging features—including gray-scale US, power Doppler, and MRI-derived quantitative maps—are being developed for automated disease activity scoring and longitudinal progression prediction [3].

Despite impressive technical accuracy, most imaging-AI studies remain single-center and retrospective, with heterogeneous scanner protocols and modest sample sizes (<300 cases). Reported diagnostic AUCs of 0.85–0.95 often lack external validation, and cost or infrastructure barriers may hinder scalability [83]. Prospective multicenter benchmarking and open-source reference datasets are urgently needed to confirm clinical benefit and cost-effectiveness.

Imaging biomarkers are poised to serve not only as diagnostic adjuncts but as surrogate endpoints in clinical trials, particularly as regulatory agencies begin to recognize imaging-derived metrics of synovitis or erosive change as validated outcome measures. Integration with digital biomarkers (e.g., wearables, motion capture) may further enable remote, multimodal disease activity tracking, ushering in an era of precision monitoring in AIRDs.

Across these domains, imaging biomarkers demonstrate the greatest real-world validation and regulatory readiness, followed by transcriptomic and proteomic signatures that show emerging predictive utility in SLE and RA. Epigenomic and fragmentomic markers remain largely exploratory, whereas polygenic-risk tools provide moderate yet ancestry-limited predictive power. Collectively, this hierarchy highlights that the integration of precision-medicine approaches in AIRD is progressing unevenly—imaging is nearing clinical implementation, multi-omics are in translational consolidation, and digital biomarkers represent the next frontier once standardization and equity challenges are resolved.

2.6. Digital Biomarkers (Wearables/Smartphones)

Digital phenotyping using smartphones and wearable sensors is increasingly recognized as a transformative approach to capture continuous, ecologically valid data in AIRDs (Figure 1) [4]. Digital phenotypes refer to quantifiable behavioral and physiological signatures—such as movement patterns, heart-rate variability, speech prosody, and sleep rhythms—captured through connected sensors and mobile devices [84]. These dynamic data streams mirror disease activity in real time, complementing conventional clinical and laboratory markers. Integrating Apple Watch-derived mobility, fatigue, and heart-rate metrics with smartphone-guided dexterity tasks enabled machine learning models to infer RA disease activity and severity with accuracy exceeding that of intermittent clinical assessments [13,85]. These digital endpoints provide high-frequency functional data that extends beyond clinic visits, offering unprecedented resolution for disease course monitoring. Unlike traditional biomarkers that capture "snapshots" of disease during scheduled visits, digital biomarkers generate dense time-series data that reflect real-world fluctuations in mobility, pain, and fatigue, thus uncovering patterns that may be invisible to episodic clinical assessment.

To ensure clinical reliability, wearable and smartphone measures undergo standardization through multi-device calibration, test–retest reproducibility studies, and validation against established clinical indices such as DAS28, SLEDAI, or CDAI [86]. Current frameworks led by the Digital Rheumatology Network (DRN) and OMERACT recommend benchmarking algorithms for accuracy, precision, and interpretability across vendors, while cross-platform normalization corrects for differences in sensor sampling frequency, placement, and firmware [13]. Such validation enables digital phenotypes to evolve from exploratory metrics into qualified clinical biomarkers capable of regulatory acceptance.

Recent validation studies of single-camera smartphone video analysis have reinforced its feasibility in rheumatology [87]. One investigation demonstrated that finger-joint mobility captured via smartphone correlated strongly with DAS28 scores and physician-assessed disease activity, underscoring its potential as a scalable, low-cost digital biomarker for tele-rheumatology [88]. Other studies have expanded this paradigm to include gait analysis, grip strength estimation, and facial expression monitoring for fatigue and pain detection, showing that even consumer-grade cameras can yield clinically actionable signals [89]. Importantly, these methods democratize access by reducing reliance on specialized imaging infrastructure, particularly in under-resourced settings.

At the systems level, the Digital Rheumatology Network (DRN) is driving consensus guidelines around the validation and integration of digital tools, emphasizing explainable AI, interoperability, and patient-centric design to ensure uptake in PsA, RA, and beyond [90,91]. The DRN and related initiatives are also addressing critical challenges of data privacy, regulatory approval, and clinical workflow integration—highlighting that technical feasibility must be matched with governance and ethical oversight. Moreover, hybrid models that fuse digital biomarkers with molecular and imaging data are emerging, enabling multi-layered disease signatures that better reflect the complexity of AIRDs.

Collectively, these initiatives position digital biomarkers as key elements of next-generation precision rheumatology, complementing molecular and imaging modalities. The integration of smartphone- and wearable-derived biomarkers into adaptive trials, treat-to-target protocols, and remote patient monitoring platforms may shorten feedback loops in clinical care—moving from reactive management of flares toward proactive, personalized disease interception.

3. Harnessing AI and Machine Learning for Autoimmune Rheumatic Diseases

Artificial intelligence (AI) and machine learning (ML) are increasingly applied to AIRD, offering opportunities for earlier diagnosis, continuous monitoring, and personalized treatment optimization [92,93]. Unlike other specialties with more homogeneous data, AIRD research faces the complexity of heterogeneous clinical presentations, multimodal diagnostic pipelines (imaging, serology, -omics), and variable disease trajectories, all of which demand sophisticated modeling approaches. Methods range from traditional regularized regression and gradient boosting to advanced deep learning and multimodal fusion architecture. Importantly, new trends highlight causal inference frameworks and clinician-in-the-loop deployment, reflecting a shift from mere predictive accuracy toward actionable, safe, and interpretable tools [94]. This shift underscores the transition from "black-box prediction" to mechanism-aware, clinically integrated AI that aligns with regulatory expectations for transparency and accountability.

Recent translational advances demonstrate that AI tools are no longer confined to experimental settings but are beginning to enhance clinical decision-making across major AIRDs. In rheumatoid arthritis, transformer-based ultrasound models have achieved near-expert accuracy for detecting synovitis and quantifying power Doppler signal, effectively

standardizing disease-activity scoring and reducing inter-reader variability [95]. In systemic lupus erythematosus, multimodal AI frameworks combining interferon gene signatures with clinical features have successfully predicted response to anifrolumab and other IFN-blocking biologics, guiding precision therapy selection in real-world cohorts [96]. Similarly, in systemic sclerosis, convolutional neural network-based nailfold-capillaroscopy systems (e.g., CAPI-Detect, EfficientNet-B0) have improved early diagnostic accuracy to over 90%, allowing detection of microvascular pathology before overt clinical manifestations [97,98]. These exemplars illustrate how AI is directly enhancing diagnostic reproducibility, flare prediction, and therapeutic response modeling, narrowing the gap between algorithmic innovation and bedside application.

3.1. Phenotyping and EHR Curation

A fundamental challenge in AIRD is the reliable construction of phenotypes from complex EHR and registry data [99,100]. National efforts such as the American College of Rheumatology's RISE registry now aggregate data across >1000 U.S. practices, enabling longitudinal monitoring and real-world model development [101–103]. However, registry-based AI pipelines require harmonization of coding systems (ICD, CPT, SNOMED), management of missingness, and governance against "phenotype drift" as disease definitions evolve (Figure 2) [104,105].

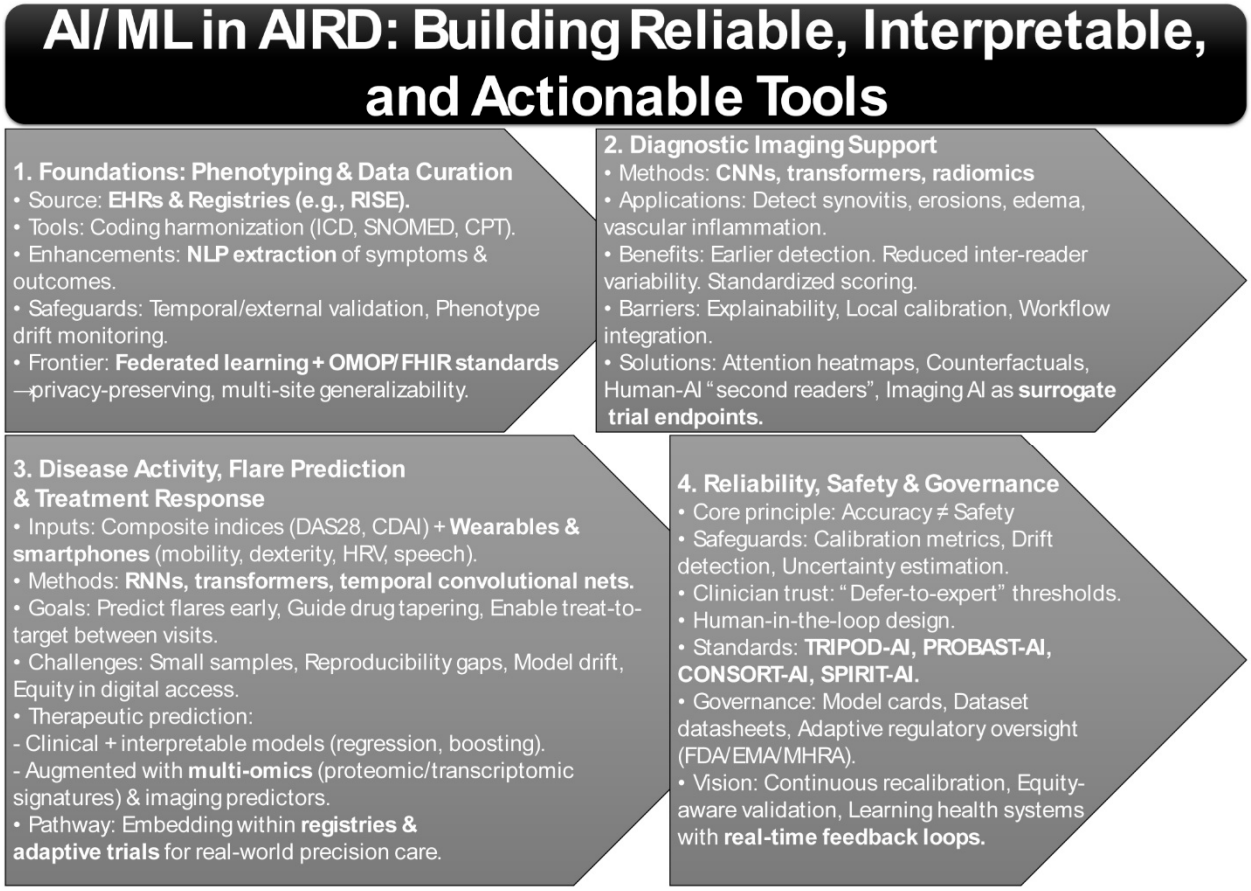


Figure 2. AI/ML Methods That Matter in Autoimmune Rheumatic Diseases: From Phenotyping to Precision Care.

Recent work demonstrates that natural language processing (NLP) can significantly enhance data capture. For example, studies showed that NLP applied to free-text rheumatology notes extracted functional status and pulmonary outcomes with higher sensitivity

than structured coding alone, reducing misclassification in registries [106–108]. Yet, these pipelines face generalizability challenges—site-specific documentation habits and EHR vendor differences often degrade performance when applied across institutions [108,109]. This underscores the need for temporal and external validation, as well as model governance structures to ensure reliability over time and across populations.

Moreover, federated learning approaches are now being piloted to enable cross-institutional model training without centralizing sensitive patient data, thereby improving generalizability while maintaining privacy. Interoperability frameworks such as OMOP-CDM and FHIR are increasingly being paired with AI pipelines to standardize data representation, further reducing barriers to multicenter validation [110,111]. Together, these developments illustrate how robust phenotype construction is evolving into the foundation for downstream predictive and prognostic modeling in AIRD.

3.2. Diagnostic Imaging Support

Imaging is perhaps the most mature application of AI in AIRD. Deep learning methods, particularly convolutional neural networks (CNNs), transformers, and radiomics pipelines, have achieved success in identifying synovitis, erosions, bone marrow edema, and vascular inflammation across conditions such as RA, axial spondyloarthritis (axSpA), systemic sclerosis (SSc), and giant cell arteritis (GCA) (Figure 2) [83,112].

A study demonstrated that AI-enhanced radiographic scoring systems could identify subtle erosive changes in RA earlier than human assessors, with reduced inter-reader variability [113]. Similarly, studies showed that transformer-based ultrasound models accurately quantified synovial proliferation and power Doppler signal, suggesting a role in standardizing clinical scoring across sites [3,114,115]. A meta-analysis of deep learning in musculoskeletal imaging confirmed robust performance but also highlighted heterogeneity in validation and lack of calibration reporting [116].

The translational challenge lies in prospective evaluation and workflow integration. While AI can accelerate reads and standardize interpretation, radiologists and rheumatologists demand explainability, uncertainty estimation, and local calibration before adopting models into routine practice. Without these, there is a risk of automation bias, especially in borderline cases where expert oversight remains indispensable.

Emerging solutions include attention heatmaps, counterfactual visualizations, and uncertainty quantification methods that help clinicians understand why a model made a given prediction [117]. Hybrid human-AI systems are also being tested, where models act as triage or "second readers," flagging high-risk scans while leaving ultimate decision-making with experts [118]. In parallel, early-phase trials are beginning to explore imaging AI as a surrogate endpoint for drug efficacy, raising the possibility that automated quantification of synovitis or vascular inflammation could accelerate therapeutic evaluation [119].

3.3. Disease Activity, Flare Prediction, and Treatment Response

Disease activity monitoring and treatment optimization constitute two of the most critical unmet needs in the application of AI to AIRD. Time-series ML frameworks have been increasingly applied to forecast fluctuations in disease activity by integrating longitudinal measurements of validated composite indices, including the Disease Activity Score in 28 joints (DAS28) and the Clinical Disease Activity Index (CDAI), with continuous streams of patient-generated health data from wearables and smartphones (Figure 2) [120,121]. Recent evidence indicates that accelerometer-derived mobility profiles and touchscreen-based dexterity metrics, when combined with patient-reported outcomes, can anticipate flare events several days prior to their clinical manifestation [122,123]. Such approaches

highlight the potential of AI to extend treat-to-target strategies into inter-visit periods and to enable proactive adjustments in disease management [124].

Importantly, flare prediction extends beyond symptom anticipation: early detection of subclinical activity may prevent irreversible joint damage, reduce corticosteroid dependence, and optimize drug tapering strategies [125]. Novel architectures such as recurrent neural networks (RNNs), transformers, and temporal convolutional networks are particularly well suited to capturing nonlinear disease trajectories and lagged effects of therapy, offering richer predictive insights than static regression-based models [126]. In addition, integration of multi-modal features—such as wearable-derived sleep disruption, HRV fluctuations, and smartphone-based speech prosody—has begun to uncover latent signatures of systemic inflammation, pointing to a broader phenome-wide approach to flare detection.

Nonetheless, current studies remain constrained by small sample sizes, limited reproducibility, and the lack of standardized digital biomarkers, thereby restricting the generalizability of predictive models. Moreover, harmonized data standards and explicit attention to equity in digital monitoring have been emphasized as essential safeguards to ensure that the deployment of flare-prediction algorithms does not exacerbate existing disparities in access to care [127]. Equally important is the issue of model drift: flare-prediction algorithms must undergo continuous recalibration as treatment paradigms, patient behaviors, and sensor technologies evolve, necessitating governance frameworks for lifecycle monitoring [128].

Parallel efforts have focused on the prediction of therapeutic response to biologic disease-modifying antirheumatic drugs (bDMARDs) and JAK inhibitors, representing a major translational frontier in precision rheumatology [129,130]. Systematic reviews report wide variability in predictive performance, with areas under the curve (AUC) ranging substantially, and highlight methodological heterogeneity across studies. Models constructed from routinely available baseline clinical variables and employing interpretable algorithms, such as penalized regression and gradient boosting, demonstrated the most consistent external validation [131,132].

By contrast, models augmented with multi-omics and imaging data often achieved higher discriminatory performance but faced increased risks of overfitting, scalability challenges, and uncertain cost-effectiveness [133]. Proteomic and transcriptomic signatures—such as IFN-stimulated gene modules in SLE or baseline TNF/IL-6 pathway activity in RA—are showing promise for predicting biologic response [134], yet their translation requires harmonized assays, prospective validation, and reimbursement strategies. Imaging-based predictors, including MRI synovitis scores and Doppler ultrasound vascularity, have also correlated with treatment response but are limited by cost and accessibility [135], raising questions about their role in routine practice.

Embedding such predictive frameworks within large-scale, registry-based adaptive infrastructures, such as the Rheumatology Informatics System for Effectiveness (RISE), offers a promising pathway for iterative refinement, prospective validation, and integration into real-world care [136]. Furthermore, adaptive trial designs are beginning to leverage prediction models for dynamic treatment allocation, accelerating drug evaluation while simultaneously generating validation data for the models themselves [76]. This bidirectional integration between AI tools and trial infrastructures represents a crucial step toward mechanism-aware precision therapeutics.

Nevertheless, unresolved challenges remain regarding regulatory approval, calibration across diverse populations, and robust health-economic evaluation, all of which will ultimately determine the feasibility and sustainability of widespread clinical adoption [137]. The long-term vision is the embedding of AI-driven flare prediction and treatment-response

tools within learning health systems, where continuous feedback loops between clinic, registry, and patient-generated data enable real-time precision care.

3.4. Reliability, Safety, and Governance

Reliability and safety remain the defining cornerstones of clinical AI deployment. High discriminatory performance alone offers little value if models are poorly calibrated or become unstable under conditions of data drift [138]. Recent studies have underscored the need for routine evaluation of calibration across multiple dimensions—including calibration-in-the-large, calibration slope, and subgroup-specific calibration—to mitigate the risk of systematic overtreatment or undertreatment, risks that disproportionately affect minority populations [139]. Evidence from medical imaging demonstrates that passive performance monitoring is insufficient to detect covariate or prevalence shifts, reinforcing the necessity of active drift detection strategies and periodic temporal validation [140]. This lesson is particularly salient in rheumatology, where treatment guidelines evolve and registries continue to expand, demanding that data quality audits and recalibration triggers be embedded within deployment pipelines [141].

Equally important is the integration of mechanisms that preserve clinical trust. Research has shown that referral triage models which incorporate uncertainty estimates and implement defer-to-expert thresholds not only enhance clinician confidence but also reduce automation bias [142]. Uncertainty quantification—through Bayesian modeling, conformal prediction, or ensemble variance estimation—is now considered essential for mitigating false reassurance and guiding safe escalation pathways [143]. By explicitly flagging ambiguous cases, AI systems can encourage collaborative decision-making rather than unilateral algorithmic recommendations. This highlights a broader principle: AI in AIRD should be designed to augment rather than replace expert judgment.

Updated reporting frameworks are beginning to institutionalize this ethos. TRIPOD-AI now mandates transparency in the handling of preprocessing steps, missing data, validation strategies, and clinical utility analyses [22]. Complementing this, PROBAST-AI provides a structured framework for assessing risk of bias and applicability in machine-learning prediction models, addressing critical gaps in peer review and regulatory oversight [21]. The CONSORT-AI and SPIRIT-AI extensions further extend this framework to the design and reporting of AI-enabled clinical trials, ensuring reproducibility and regulatory compliance in prospective studies [144]. Together, these frameworks are shaping a standards-based ecosystem that prioritizes interpretability, fairness, and clinical relevance alongside raw accuracy (Figure 2).

Governance infrastructures are also emerging as critical components of reliable AI deployment. Model cards and datasheets for datasets are increasingly required to document training cohorts, limitations, and intended use cases, while regulatory bodies such as the FDA, EMA, and MHRA are developing adaptive oversight frameworks for continuously learning algorithms (Figure 2) [145,146]. In rheumatology, this governance must also account for ancestry-aware PRS, equity in digital biomarker access, and evolving therapeutic landscape issues that amplify the risk of model obsolescence if oversight is not dynamic [147].

Ensuring equity in AI-enabled rheumatology requires deliberate sampling across ancestry, sex, and socioeconomic strata to capture population diversity and avoid algorithmic bias [148,149]. Fairness-auditing pipelines, subgroup calibration metrics, and transparent documentation of data composition are now considered essential safeguards to prevent systemic bias and ensure model generalizability. Moreover, socioeconomic disparities in access to smartphones and wearables introduce a "digital divide" that can distort data inputs and downstream predictions. Addressing these inequities demands inclusive dataset design,

equity-weighted model evaluation, and regulatory alignment with initiatives such as the FDA's Algorithmic Bias Guidance, which encourages demographic reporting and continuous fairness monitoring across product lifecycles [139]. Embedding these safeguards within learning-health-system infrastructures will be vital to guarantee that AI advances translate equitably across all patient populations.

Taken together, these methodological advances outline a pragmatic blueprint for implementation. The process begins with rigorous phenotyping of electronic health records and careful curation of registry data, which establish the foundation for reliable model development. From this base, interpretable baseline models should be benchmarked before advancing to more complex multimodal deep learning frameworks, ensuring that predictive performance is not gained at the expense of transparency. At the deployment stage, systematic incorporation of calibration metrics, decision-curve analysis, and mechanisms for drift monitoring became essential to maintain reliability over time and across settings. Equally important is the design of clinician-in-the-loop interfaces that not only safeguard trust but also promote safety, transparency, and equity—ensuring that AI systems enhance rather than disrupt existing care pathways.

The trajectory of the field has already moved beyond proof-of-concept demonstrations toward registry-scale feasibility. Yet the next decisive frontier lies in prospective validation, the adoption of causal framing to strengthen inference, and the seamless integration of these models into decision support systems capable of delivering demonstrable improvements in patient outcomes. Ultimately, governance in AIRD AI must balance innovation with accountability, ensuring that models are safe, equitable, and continuously aligned with evolving standards of care.

4. Redefining Autoimmune Rheumatic Disease Pathways: From Immune Signatures to AI-Enhanced Precision Medicine

4.1. Rheumatoid Arthritis (RA)

Recent randomized controlled trial demonstrated that abatacept, a CTLA-4–Ig co-stimulation blocking biologic, can delay the transition from autoantibody positivity with arthralgia (pre-RA) to clinically classifiable RA. The APIPPRA trial provided early evidence of preventive efficacy in seropositive individuals without overt synovitis (Table 1) [150], while the ARIAA trial extended these findings to patients with subclinical joint inflammation detected by MRI [151]. In both studies, abatacept was associated with reduced progression to overt RA and diminished inflammatory activity, with benefits persisting beyond the treatment period. Collectively, these trials support the concept that targeted immunomodulation during the pre-clinical phase can alter the natural course of disease development.

Together, these trials underscore a paradigm shift toward early "disease interception" in RA, whereby immune modulation in at-risk individuals may prevent or substantially delay disease onset. This represents a new frontier in rheumatology, where prevention-oriented strategies could reshape the natural history of disease. Yet critical questions remain regarding the long-term durability of benefit, potential rebound activity following treatment cessation, and the cost-effectiveness of extending biologic therapies into pre-clinical or at-risk populations. Moreover, the risk/benefit calculus of exposing asymptomatic individuals to immunosuppressive agents requires careful evaluation through adaptive, stratified trial designs [152].

Table 1. Disease-Focused Advances in AI, Biomarkers, and Digital Health Across Autoimmune Rheumatic Diseases.

Disease	Key Biomarkers/Targets	AI/Digital Innovations	Clinical Impact	Key Limitations/Gaps
Rheumatoid Arthritis (RA)	Pre-RA prevention with abatacept (APIPPRA, ARIAA); autoantibodies (ACPA, RF); MRI-detected subclinical inflammation	Deep learning for US/MRI synovitis segmentation; sub-pixel JSN quantification; smartphone-based fist closure (MeFISTO) as a functional biomarker; ML models combining multi-omics + imaging	Demonstrates feasibility of disease interception; scalable imaging and digital biomarkers; early steps toward individualized drug response prediction	Long-term durability of prevention unknown; small imaging datasets; lack of external validation; heterogeneity in ML pipelines
Systemic Lupus Erythematosus (SLE)	Type I IFN gene signature; SIGLEC-1 expression; proteomic biomarkers (SAA1, B4GALT5, etc.)	IFN-signature guided therapy with anifrolumab; wearables + PROs (OASIS study); EHR-based flare prediction (FLAME); deep learning for lupus nephritis flares; proteomic + ML flare models	Establishes IFN signature as both predictive and monitoring biomarker; digital phenotyping enables early flare detection	Variable organ-specific response; inconsistent LN outcomes; digital tools often under-validated; flare definitions heterogeneous
Systemic Sclerosis (SSc)	Microvascular patterns (giant capillaries, hemorrhages, density loss) on nailfold capillaroscopy	AI-assisted NFC classification: ResNet, EfficientNet, CAPI-Detect; large, annotated NFC datasets; pattern staging (early/active/late)	Enhances reproducibility and early diagnosis; potential for risk stratification (e.g., pulmonary hypertension, ulcers)	Few longitudinal outcome studies; lack of standardized acquisition protocols; external validation limited
Spondyloarthritis (SpA)	HLA-B27, MRI sacroiliac inflammation, PROs, PK parameters	Registry-based ML models (EuroSpA secukinumab cohort); ROC-SpA trial testing PK-guided prediction	Supports treatment persistence and real-world prediction; PK may inform therapeutic drug monitoring	Heterogeneous endpoints; small sample sizes; lack of standardized composite outcomes
Psoriatic Arthritis (PsA)	Disease activity, comorbidities, sonographic inflammation	US-based short-interval predictors (MIJET/2MIJET); early discrimination of JAKi vs. TNFi/ILi responses	Demonstrates feasibility of early imaging response markers; pragmatic outcome (drug retention)	Small pilot cohorts; scarce validated molecular predictors; multi-domain disease complicates modeling
Sjögren's Disease (SjD)	SGUS scores (OMERACT, Hočevar); salivary/tear proteomics; expanded autoantibodies	Standardized SGUS linked to lymphoma risk; proteomic pipelines integrating saliva, plasma, tissue	Non-invasive early diagnosis and risk stratification; complements biopsy	Need for longitudinal validation; risk of over-screening; proteomic candidates require replication
Idiopathic Inflammatory Myopathies (IIM)	Myositis-specific autoantibodies (MSAs); MRI muscle edema; multi-omics panels	ML clustering integrating MSAs + MRI + omics; radiomics-based antibody group prediction	Improved subtype stratification; potential guidance for ILD or therapy selection	Mostly retrospective, single-center; translation to outcomes (e.g., steroid-sparing) unproven
Vasculitides	CRP, ANCA patterns, type I IFN signatures; renal 12-gene transcriptomic panel	PET-CT radiomics/ML distinguishing GCA vs. atherosclerosis; transcriptomics predicting kidney failure in AAV	Enables precision risk stratification (renal outcomes, vascular inflammation)	Early-stage; need for harmonized endpoints; prospective trials embedding predictors into care are lacking

Artificial intelligence (AI) and deep learning methods are increasingly applied to US and MRI for detecting and quantifying synovitis, joint erosion, and joint space narrowing. Recent systematic reviews highlight frameworks aligned with RAMRIS and OMERACT standards, employing architectures such as U-Net, convolutional neural networks, and transformer variants to improve reproducibility compared with human scoring, while also enhancing sensitivity to change [153–156]. Automated segmentation algorithms now achieve near-human accuracy in delineating synovial hypertrophy and erosions [157], while deep radiomics pipelines are beginning to uncover latent imaging features predictive of future structural progression, even before they are visually appreciable [158].

Innovative sub-pixel quantification methods have also been proposed for detecting minute changes in joint space narrowing (JSN) on radiographs, increasing sensitivity in

early disease where structural progression may be subtle [159,160]. Despite progress, many imaging AI studies remain constrained by small, homogeneous datasets, lack of external validation, and inconsistent image acquisition protocols, which collectively limit clinical deployment [161]. Future efforts will require federated learning across multi-center cohorts, harmonization of imaging protocols, and incorporation of calibration metrics to ensure robustness across devices, vendors, and populations.

Digital health approaches are increasingly explored as scalable, objective tools for functional monitoring in RA. A study applied single-camera smartphone motion capture to assess repeated fist closures [88]. Extracted kinematic features—including range of motion, time to maximal flexion, and velocity—correlated strongly with disease activity measured by DAS28. Such approaches illustrate the feasibility of remote functional biomarkers for RA, aligning with treat-to-target strategies and expanding the potential for continuous, home-based disease monitoring. In addition, wearable-derived accelerometry, grip strength sensors, and smartphone-based joint stiffness trackers are being integrated into multimodal pipelines, opening opportunities for near real-time flare detection and longitudinal disease activity profiling [162]. Nevertheless, larger multi-center validation and clear regulatory pathways for digital biomarker adoption remain prerequisites for clinical translation.

Advances in multi-omics integration and machine learning (ML) have accelerated efforts to predict therapeutic response across biologic DMARDs (bDMARDs) and Janus kinase inhibitors (JAKi) [163,164]. A recent scoping review synthesized nearly ninety studies, the majority in RA, with smaller but growing efforts in spondyloarthritis (SpA) and psoriatic arthritis (PsA). These studies leveraged diverse inputs, including clinical biomarkers, genomic variants, proteomic patterns, and imaging-derived features. Reported performance was heterogeneous, with modest to strong discriminatory ability depending on data type and modeling approach [165].

Models that combined multi-omics data with imaging signatures generally outperformed those based on clinical or single-modality inputs. However, most remain exploration, constrained by limited external validation and insufficient reproducibility across cohorts [165]. Emerging candidate predictors include interferon- and B-cell-related gene expression modules, autoantibody glycosylation patterns, and proteomic correlates of TNF/IL-6 pathway activity [12]. While these molecular features highlight promising biological axes for precision therapy, translation into practice is hampered by unresolved issues: the absence of harmonized assay platforms, challenges in standardizing bioinformatic pipelines, and the need for real-world cost-effectiveness evaluations to justify implementation at scale.

In parallel, a study in large European RA cohorts identified clinical and serologic predictors of response to b/tsDMARDs. Factors such as baseline disease activity, age, prior biologic exposure, inflammatory markers (CRP, ESR), and comorbidities influenced therapeutic outcomes, offering pragmatic tools for patient stratification [166]. These pragmatic predictors, while less mechanistically granular than omics-driven models, currently represent the most immediately translatable approach, particularly in health systems where resource constraints limit access to advanced biomarker profiling.

4.2. Systemic Lupus Erythematosus (SLE)

4.2.1. IFN Signature & Targeted Therapy

Type I interferon (IFN) signaling is central to SLE pathogenesis and contributes to disease activity, organ involvement, and long-term prognosis. Disease heterogeneity is reflected in variable IFN gene signature (IFNGS) levels, autoantibody profiles, and downstream pathways such as neutrophil extracellular trap (NET) formation (Table 1) [167,168]. For example, a study demonstrated associations between autoantibodies, elevated IFN signatures, NET

release, and clinical phenotypes, underscoring IFN signaling as a mechanistic driver of disease expression [169].

Anifrolumab, a monoclonal antibody targeting IFNAR1, was approved for moderate-to-severe SLE following pivotal phase III trials. Recent studies confirm that anifrolumab suppresses the IFN signature, reduces cutaneous activity, and lowers flare frequency [170,171]. A study reported durable IFN signature suppression accompanied by clinical improvement [172]. Post-marketing evidence has further validated these findings, showing real-world effectiveness in reducing corticosteroid dependence and improving patient-reported fatigue scores—an especially relevant outcome given the high burden of fatigue in SLE [173].

Studies emphasized that baseline IFN signature magnitude stratifies response likelihood across IFN-targeting therapies, including anifrolumab and anti-IFNα antibodies [170,174]. Patients with a "high IFN" molecular endotype consistently demonstrate greater probability of response, suggesting that the IFNGS may serve as both a predictive and pharmacodynamic biomarker. Conversely, "low IFN" patients often fail to derive meaningful benefit [175], highlighting the necessity of molecular stratification prior to initiating IFN-targeted therapies.

Emerging translational data suggest IFN signatures can serve as monitoring biomarkers. It has been shown that longitudinal changes in SIGLEC-1 expression on monocytes correlated with systemic and cutaneous responses; increased SIGLEC-1 was associated with relapse [176]. Other interferon-inducible proteins, including CXCL10 and ISG15, are also being evaluated as dynamic readouts of pathway activity, with potential to inform early therapeutic switching before clinical relapses are apparent [177].

Not all patients with high IFN signatures respond, and responses vary by organ system. In lupus nephritis, benefits on proteinuria reduction are inconsistent [178]. A study demonstrated that subsets of SLE patients may exhibit uncoupled IFN pathway activation, providing one explanation for variable treatment responses [179]. This heterogeneity may reflect differential activation of type I versus type II interferon pathways, crosstalk with BAFF/TNF signaling, or organ-specific immune microenvironments that alter drug penetrance and pathway dependence. Such findings emphasize the need for combination therapies—pairing IFN blockade with agents targeting B-cell activation, complement, or JAK/STAT signaling—to achieve durable remission across diverse SLE manifestations.

Long-term outcome studies are needed to determine whether sustained IFN signature suppression reduces irreversible organ damage. Key unanswered questions include whether IFN suppression modifies cardiovascular risk, mitigates neuropsychiatric SLE progression, or prevents accrual of organ damage over decades of disease. Prospective registries and adaptive platform trials will be essential to establish the durability, safety, and health-economic impact of IFN-targeted therapies in SLE.

4.2.2. Digital Measures & Flare Prediction

The OASIS study applied biosensors and patient-reported outcomes (PROs) to longitudinally track ~550 SLE patients, including 144 smartwatch users [180]. The study showed that integrating biometric data, quality-of-life metrics, and PROs into ML classifiers achieved strong flare discrimination [180]. These results demonstrate the feasibility of integrating passive physiological data streams—such as heart rate variability, step counts, and sleep duration—with subjective symptom reports, thereby creating multimodal signatures of flare risk that extend beyond clinic-based assessments.

In another investigation, the FLAME pipeline (FLAre Machine learning prediction of SLE) was developed, demonstrating that multivariable EHR data could predict flares with fair accuracy, providing proof-of-concept for real-world data integration [181]. Importantly, FLAME leveraged routinely available structured fields (labs, medications, visit patterns) to

build scalable models, suggesting that pragmatic flare prediction tools can be embedded within existing health record infrastructures without requiring additional patient burden.

In lupus nephritis, a deep learning model incorporating 59 demographic, clinical, and pathological features achieved strong discrimination for predicting renal flares in time-series datasets [182]. This approach highlights the potential of multimodal fusion—linking histopathology with longitudinal clinical variables—to anticipate nephritic flares before overt proteinuria or renal dysfunction emerges, a capability that could transform monitoring strategies and guide earlier therapeutic escalation.

Another study identified patient-prioritized digital concepts of interest (COIs)—walking ability, sustained activity, rest, and sleep—and mapped them to measurable digital clinical measures (DCMs) such as daily steps and sleep efficiency [183]. Patients favored wrist-worn devices, highlighting acceptability. The prioritization of mobility and fatigue-related metrics underscores that patient-valued outcomes may differ from physician-centric disease activity indices, emphasizing the importance of patient-centered design in digital biomarker development.

A separate study used baseline plasma proteomics (data-independent acquisition) plus clinical data to predict flares over one year in an Asian cohort (AUC $\approx$ 0.77). Candidate biomarkers included SAA1, B4GALT5, GIT2, NAA15, and RPIA [42]. These findings illustrate the promise of integrating circulating proteomic signatures with digital and clinical data streams, potentially enabling hybrid molecular-digital biomarkers for robust flare prediction.

Current digital models are often limited by small sample sizes, reliance on internal cross-validation, and inconsistent flare definitions. External validation and regulatory qualification of DCMs remain pressing needs. Standardizing flare definitions across studies, establishing interoperability between devices, and embedding validation cohorts across ancestries and healthcare systems will be essential to achieve regulatory recognition. Ultimately, the goal is to transition digital measures from exploratory research endpoints into qualified biomarkers that can support trial enrichment, adaptive dosing, and real-world disease monitoring in SLE.

4.3. Systemic Sclerosis (SSc)

A pilot study using ResNet-34 + YOLOv3 achieved sensitivity and specificity of ~89% for distinguishing pathological vs. normal nailfold images, with ~96.5% precision for capillary density counts in systemic sclerosis (SSc) cohorts [184]. AI applications in nailfold capillaroscopy (NFC) have highlighted the utility of supervised deep learning for detecting SSc-specific abnormalities such as giant capillaries, hemorrhages, and capillary dropout. However, longitudinal validation linking microvascular metrics to clinical outcomes (e.g., digital ulcers, pulmonary hypertension) remains scarce (Table 1) [97,185].

In one study, CAPI-Detect was developed as a machine learning model trained on >1500 NFC images using 24 quantitative features. It achieved ~91% accuracy in distinguishing scleroderma-specific from nonspecific patterns and improved classification across early, active, and late SSc microvascular stages [98]. Another investigation reported that an EfficientNet-B0 cascade model, applied to NFC images from 225 patients, achieved ROC-AUC values near 1.0, with substantial gains over conventional single-transfer learning methods [97]. A separate group released a large, annotated NFC dataset (321 images, 219 videos), enabling training for both morphological and dynamic feature extraction. Their pipeline reached ~89.9% accuracy for abnormal morphology detection with sub-pixel measurement precision [185].

These advances position AI-assisted NFC as one of the most mature digital applications in SSc, offering potential for automated early diagnosis, microvascular staging, and longitudinal monitoring. Importantly, AI-driven quantification reduces inter-observer

variability—a longstanding challenge in manual NFC interpretation—and provides scalable, reproducible readouts that may standardize clinical practice. Moreover, the ability to detect subtle microvascular alterations could enable earlier identification of patients at risk for vasculopathic complications, including digital ulcers and pulmonary arterial hypertension, both of which are major drivers of morbidity and mortality in SSc.

Despite progress, challenges remain. AI-assisted NFC is approaching clinical readiness, yet lack of multicenter datasets, absence of prospective prognostic validation, and variability in acquisition protocols limit immediate deployment. Most current studies are retrospective and rely on relatively small, single-center image repositories, raising concerns about generalizability across devices, geographic populations, and disease subtypes. Furthermore, few investigations have explicitly linked AI-derived NFC features to downstream organ involvement, therapeutic response, or long-term survival, which are essential for establishing clinical and regulatory value.

4.4. Spondyloarthritis/Psoriatic Arthritis (SpA/PsA)

Predictive modeling in spondyloarthritis (SpA) remains heterogeneous but is steadily advancing [186]. A scoping review encompassing 89 AI studies across inflammatory arthritis identified only 11 studies in SpA/PsA, with reported performance ranging from accuracy ~60–70% and AUC 0.63–0.92. Multi-omics and imaging-augmented models generally outperformed clinical-only baselines, but methodological heterogeneity and limited external validation constrained generalizability [165]. This reflects a broader challenge in SpA research: unlike rheumatoid arthritis, where serologic biomarkers (e.g., ACPA, RF) provide mechanistic anchors, SpA and PsA lack universally validated molecular markers, necessitating greater reliance on composite clinical, imaging, and lifestyle predictors (Table 1).

Within axial SpA, a multiregistry EuroSpA cohort study analyzing secukinumab-treated patients found that clinical characteristics, patient-reported outcomes (PROs), and lifestyle factors predicted achievement of low disease activity (ASDAS-CRP, BASDAI) at 6 months and treatment persistence at 12 months [187]. This underscores the value of real-world registry data for prognostication. Interestingly, registry-based models highlight that baseline PROs—such as patient global assessment and fatigue—can be as predictive of drug persistence as traditional biomarkers, suggesting that patient-centered data streams may be critical for individualized treatment planning.

In parallel, the ROC-SpA randomized protocol will test whether pharmacokinetic (PK) parameters (drug levels, exposure) predict clinical response at 24 weeks following anti-TNF failure, representing a shift toward therapeutic drug monitoring in SpA [188]. If validated, PK-informed personalization could establish a new treatment paradigm, aligning drug exposure with disease endotypes rather than applying uniform dosing strategies. Such approaches may also help rationalize costs by avoiding unnecessary biological cycling in non-responders.

For psoriatic arthritis (PsA), imaging biomarkers are emerging as key translational tools. A recent prospective pilot study showed that short-interval ultrasound changes in inflammation (MIJET/2MIJET/GUIS scores) at 1–3 months predicted 6-month drug retention, with faster responses observed in JAK inhibitor-treated patients compared with TNF, IL-17, or IL-12/23 inhibitor therapy [189]. This suggests that dynamic imaging readouts may serve as early surrogate markers of therapeutic persistence, accelerating adaptive decision-making in PsA. Larger PsA cohorts confirm that baseline disease activity, prior biologic exposure, and comorbidity burden influence therapeutic response, but validated molecular predictors remain scarce [190,191]. Emerging multi-omics studies have identified candidate pathways—including IL-23/Th17 signaling, keratinocyte-derived

cytokines, and metabolic dysregulation—that may differentiate responders from non-responders [192], but these remain exploratory pending replication in diverse cohorts.

Key gaps remain, as modest sample sizes, variable imaging protocols, and inconsistent composite outcomes across studies limit generalizability. Moreover, heterogeneity in disease domains—peripheral arthritis, axial involvement, enthesitis, dactylitis, and skin disease—complicates biomarker validation, since predictors may differ by dominant phenotype. This necessitates domain-specific models or modular prediction frameworks that can adapt to different PsA presentations. Large, multicenter studies integrating PK data with multi-modal predictors (clinical, imaging, genomic, and lifestyle) are required to establish clinically deployable prediction tools. Ultimately, the integration of SpA/PsA predictive models into learning health systems and adaptive trial designs will be essential to move from exploratory research toward precision, mechanism-guided care.

4.5. Other Conditions

4.5.1. Sjögren's Disease (SjD)

Salivary gland ultrasound (SGUS) is increasingly validated for diagnosis and risk stratification. A study comparing OMERACT vs. Hočevar scoring demonstrated that parotid ultrasound features correlated with lymphoma risk in SjD [193]. A new multicenter study confirmed that SGUS correlates with secretory function, systemic disease activity, and lymphoma risk factors, supporting its broader clinical use [194]. Importantly, updated guidelines caution against routine repeat SGUS in asymptomatic patients, highlighting the need to avoid over-screening [195].

On the biomarker front, salivary proteomics are rapidly evolving. A study review emphasized proteomic pipelines for identifying candidate diagnostic and prognostic markers [196], while an integrative study combining saliva, plasma, and gland tissue proteomics identified novel biomarker candidates for SjD classification [197]. Emerging evidence also suggests that salivary exosomal microRNAs and proteoforms may offer superior sensitivity for early-stage disease compared with conventional serologic markers such as anti-Ro/SSA and anti-La/SSB (Table 1) [198]. Integration of SGUS and proteomics into multimodal diagnostic algorithms could therefore accelerate detection, stratify lymphoma risk, and refine patient selection for clinical trials.

Standardized SGUS and proteomic panels represent complementary tools for early diagnosis and lymphoma risk stratification, but longitudinal validation linking imaging/omics outputs to clinical outcomes remains essential. Future directions include embedding SGUS-proteomic fusion models within registry-based cohorts and testing their capacity to predict long-term systemic involvement, malignancy risk, and response to B-cell-targeted therapies.

4.5.2. Idiopathic Inflammatory Myopathies (IIM)

Stratification of IIM is increasingly driven by myositis-specific autoantibodies (MSAs), imaging, and ML approaches. Reviews emphasize how MSAs have redefined subtype classification and prognosis [199], while methodological papers detail ML opportunities for biomarker discovery and patient clustering [200]. For instance, anti-MDA5 positivity is strongly associated with rapidly progressive interstitial lung disease (ILD), while anti-TIF1-γ predicts malignancy risk, underscoring the prognostic utility of serologic stratification [201].

Applications of ML to muscle MRI have demonstrated feasibility in predicting antibody-defined subgroups and disease clusters via radiomics and texture features [202]. Similarly, multi-omics pipelines are being tested to improve stratification, though most remain single-center, retrospective, and exploratory. Deep learning applied to T2-weight and STIR MRI sequences has revealed latent imaging phenotypes that correspond to distinct

histopathological patterns, suggesting potential for early detection of subclinical muscle inflammation [203]. Moreover, integrative ML models combining autoantibody profiles, transcriptomics, and MRI data are beginning to uncover mechanistic endotypes that may guide immunosuppressive therapy selection (Table 1) [204,205].

Despite proof-of-concept success, prospective multicenter validation linking ML-based stratification to treatment response and clinical outcomes (e.g., ILD progression, steroid-sparing) is urgently needed [206]. Next steps will require harmonization of MRI acquisition protocols, integration of patient-reported outcomes, and trial-based testing of whether biomarker-guided stratification can optimize therapeutic decisions in IIM.

4.5.3. Vasculitides

In large-vessel vasculitis (LVV), [ˆ18F] FDG-PET/CT is increasingly central to diagnosis and monitoring. Studies show PET/CT can confirm vascular involvement when biopsies are negative, and emerging radiomics/ML models distinguish active giant cell arteritis (GCA) from atherosclerosis, potentially reducing diagnostic uncertainty [207]. Hybrid imaging approaches, including PET/MRI, are further expanding the toolkit, enabling simultaneous metabolic and anatomical assessment of vessel inflammation [208]. These modalities may serve not only as diagnostic adjuncts but also as surrogate endpoints for treatment response in clinical trials.

In ANCA-associated vasculitis (AAV), renal transcriptomic signatures are advancing risk prediction. A study developed a 12-gene renal signature that outperformed clinico-pathologic scores in predicting kidney failure [209]. More recent data suggest that stronger type I IFN signatures predict worse renal outcomes and distinct clinical phenotypes, underscoring immune-pathway-guided precision medicine [210]. Integration of kidney biopsy transcriptomics with digital pathology and single-cell sequencing is beginning to delineate cellular drivers of renal injury, potentially guiding therapeutic targeting of pathogenic myeloid and interferon-driven networks (Table 1).

A recent review synthesized AI applications across vasculitides, highlighting progress in diagnostic imaging, biomarker discovery, and outcome prediction, while stressing the need for larger, prospective harmonized datasets [211]. Moving forward, federated AI pipelines across vasculitis consortia and international biobanks will be essential to achieve sufficient statistical power and ensure equitable performance across ancestries.

PET-based radiomics and renal transcriptomics exemplify organ-specific precision tools; the next step is embedding these predictors into decision-impact trials to guide therapy. Ultimately, the goal is to transform these tools from retrospective predictors into real-time decision-support instruments capable of improving patient survival, organ preservation, and quality of life.

5. Artificial Intelligence in Rheumatology: From Triage to Therapy Selection

5.1. AI-Enhanced Triage and Access

Among rheumatology applications, text-based triage is the furthest along the translation curve because it addresses a clear bottleneck—waiting times—without displacing diagnostic authority. A recent multicenter study processed 8044 GP referral letters (5728 patients) from 12 clinics, training models in two centers and testing in the remaining ten. This external-site validation design reduces the risk of "center overfitting" and provides evidence of genuine generalizability across healthcare systems [212]. The system prioritized likely RA, OA, fibromyalgia, and long-term care needs, showing that machine learning can augment queue management and equity of access.

The translational value here lies in optimizing time-to-assessment rather than automating diagnosis. Deployment, however, depends on safeguards such as calibration monitoring (slopes and intercepts), deferral rules for low-confidence predictions, and post-deployment drift audits to detect changes in case-mix or letter style. Comparable patient-facing systems, such as RhePort (digital rheumatology patient intake and referral platform), demonstrate that combining structured digital intake with NLP triage could further streamline referral pathways [212,213]. Such hybrid platforms not only shorten diagnostic delays but may also reduce inequities in access, particularly for patients in underserved regions, by providing consistent triage independent of referral letter quality (Table 2).

Table 2. AI in Rheumatology—From Diagnostics to Decisions.

Domain	Data/ Input	Model Types	Validation Status	Clinical Maturity	Key Challenges	Next Translational Step
Triage & Access	Referral letters (NLP), structured intake	NLP (transformers, boosting)	External-site validation [212]	High—first in line for deployment	Calibration drift, subgroup fairness, governance rules	Registry/workflow embedding, prospective drift monitoring
Imaging Decision Support	MRI, US, NVC images	CNNs, end-to-end pipelines	Multicenter feasibility (MRI/US/NVC); reproducibility demonstrated [83,98,112]	Moderate—reader-assist tools maturing	Scanner/vendor heterogeneity, lack of prospective trials	Workflow-embedded prospective evaluation; standardized reporting
Therapy Selection	Clinical data, serology, omics, imaging	Gradient boosting, multimodal ML, RNA-seq signatures	Mostly retrospective; limited external/temporal validation [132,165]	Early—promising but not trial-ready	Heterogeneity, lack of harmonized endpoints, no impact trials	Registry pilots; decision-curve analysis; prospective impact studies
Risk Stratification (Adjacent: RA-ILD)	Biomarkers (KL-6), imaging, clinical data	XGBoost, ensemble ML	Cohort-level validation [214]	Moderate—translational potential for early screening	Generalizability, integration into care pathway	

These findings suggest that triage augmentation—supported by transparent reporting frameworks such as TRIPOD-AI and PROBAST-AI—is the most immediate clinical application of AI in rheumatology [22].

5.2. Imaging Decision Support

Imaging AI in rheumatology is converging on a reader-assist model, standardizing quantification rather than replacing interpretation. In RA, deep learning pipelines now achieve volumetric quantification of synovitis and erosions on contrast-enhanced MRI, correlating strongly with RAMRISs and matching expert reproducibility [83,112]. Similarly, the ARTHUR v2.0 ultrasound platform integrates segmentation and activity grading aligned with OMERACT/EULAR definitions, supporting consistent scoring across sites [215].

Systemic sclerosis provides an instructive parallel. A paper demonstrated that convolutional neural networks could fully automate nailfold capillaroscopy (NVC) interpretation [98]. More recently, the multicenter CAPI-Detect initiative refined ML-based scoring of capillary density, hemorrhages, and the scleroderma pattern, reporting reproducible accuracy across independent datasets [98,216]. Complementary work further validated automated microvascular abnormality detection, underscoring the feasibility of AI-assisted NVC for SSc [185].

The clinical takeaway is that imaging AI is most advanced in standardized scoring, workload reduction, and trial reproducibility, not in stand-alone diagnosis. Prospective workflow-embedded studies remain the key translational step before these systems can enter routine care [217]. In parallel, explainability tools such as saliency maps and uncertainty quantification are increasingly incorporated into imaging AI pipelines to address clinician trust and mitigate automation bias [218,219]. Furthermore, the incorporation of

imaging AI as surrogate endpoints in drug trials may accelerate therapeutic evaluation by providing objective, reproducible readouts of disease activity.

5.3. Predictive Tools for Therapy Selection

The prediction of biologic or targeted synthetic DMARD response represents a more ambitious but less mature AI application. A systematic review of 89 AI studies across inflammatory arthritis reported AUCs ranging from 0.63 to 0.92, with multi-omics and imaging features consistently improving discrimination over clinical baselines. Yet, methodological heterogeneity and limited external validation restrict current clinical use [132,165].

Encouragingly, registry-based models using only baseline clinical features and gradient boosting have achieved clinically plausible predictive accuracy for 6- and 12-month bDMARD outcomes, illustrating the value of transparent, implementable baselines [220]. At the mechanistic layer, whole-blood RNA-seq studies are beginning to identify molecular predictors of JAK inhibitor response, though these remain at the discovery stage [221]. Early findings suggest that interferon-driven transcriptional modules and metabolic pathway activity may stratify JAK inhibitor responders, but harmonization of RNA-seq assays and prospective biomarker-guided trials will be needed before translation [222].

Adjacent work in RA complications underscores translational potential: ML models predicting RA-associated interstitial lung disease (RA-ILD) using clinical and biomarker features such as KL-6 have shown robust performance, supporting early screening applications [214]. This illustrates that prediction models can also be extended beyond drug response toward complication forecasting, potentially enabling proactive surveillance strategies that pre-empt irreversible organ damage (Table 2).

The critical point is that predictive models must move beyond retrospective accuracy toward impact-on-care trials. Embedding calibrated decision aids into registries, running decision-curve analyses, and tracking equity across subgroups will be prerequisites before routine deployment. Reporting should adhere to TRIPOD-AI and PROBAST-AI, while early evaluations align with DECIDE-AI guidance before progression to randomized impact studies [22,223]. Ultimately, the clinical value of predictive tools will be judged not by ROC curves but by their ability to change physician behavior, improve patient outcomes, and demonstrate cost-effectiveness in real-world healthcare systems.

6. Data Infrastructures for AI in Rheumatology: Registries, Interoperability, and Federated Collaboration

6.1. Registries and EHR as Foundational Substrates

The American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) registry exemplifies how electronic health record (EHR)-enabled infrastructures can serve simultaneously as clinical quality improvement (QI) engines and research substrates [18]. Recognized by the Centers for Medicare & Medicaid Services (CMS) as a Qualified Clinical Data Registry (QCDR), RISE aggregates encounter-level data from participating practices and returns interactive dashboards that benchmark performance, track patient-level quality measures, and support reporting for the Merit-based Incentive Payment System (MIPS) and MIPS Value Pathways (MVPs) [224].

An interrupted time-series analysis showed that enrollment in RISE was associated with sustained improvements across rheumatology quality measures, with the strongest gains observed in RA disease activity documentation and functional status assessment [103,225]. These findings provide robust real-world evidence that registry-embedded feedback loops can alter practice behavior and advance care quality, particularly among lower-performing practices. Importantly, RISE demonstrates that QCDRs can function as "learning health

systems," where continuous data capture and real-time analytics translate directly into measurable practice improvement (Table 3).

Table 3. Data Infrastructure for Precision Rheumatology.

Infrastructure Pillar	Strengths	Limitations	Clinical Applications
Registries & EHR (e.g., RISE)	National-scale QCDR; supports quality improvement and reimbursement; real-world evidence of improved care quality.	Dependent on practice adoption and data quality; disease coverage is still limited (e.g., lupus measures emerging).	Quality benchmarking, CMS Quality Payment Program reporting, registry-based research.
Interoperability (OMOP/FHIR)	FHIR enables clinical data exchange; OMOP supports multi-site analytics; hybrid architectures proven feasible.	Standards alone are insufficient; require metadata, governance, and ontology alignment (SNOMED CT, LOINC).	Multi-site analytics, phenotyping, clinical trial recruitment, harmonized real-world evidence generation.
Privacy-Preserving Collaboration (Federated Learning)	Allows multi-site model training without centralizing patient data; governance frameworks emerging; feasible in diverse clinical tasks.	Technically complex; potential for bias and fairness issues; resource-intensive implementation.	Comparative effectiveness research (e.g., RA biologics), collaborative risk prediction across sites.
Multisite Modeling Pitfalls & Mitigations	Recognition of covariate shift, site bias, and acquisition drift; new methods (FedWeight, COLA-GLMM) improve calibration and cross-site validity.	Residual generalization challenges: continuous monitoring and retraining required.	Development of fairer, more robust models; deployment with embedded recalibration triggers.
Case Example: RA Disease Activity Prediction	Demonstrated feasibility using EHR + PRO features; established templates for computable disease-activity endpoints.	Early studies lacked robust external validation; not yet embedded in clinical dashboards.	Risk-stratified dashboards for RA management; integration of prediction models into quality improvement cycles.

The sustainability of such systems is reinforced by policy integration. RISE participation aligns directly with CMS QPP pathways, ensuring financial and regulatory incentives for ongoing data submission and quality reporting. Governance structures continue to expand measure sets (e.g., for lupus) and mandate rigorous validation prior to national roll-out—practices that are essential to maintain scientific integrity in large-scale QCDRs. Future expansions are expected to include digital biomarkers, patient-reported outcomes, and imaging data streams, enabling multimodal precision analytics within a single national registry.

6.2. Interoperability and Common Data Models

The utility of registries such as RISE is contingent on their ability to interoperate across sites and vendors. Two standards have become foundational: Fast Healthcare Interoperability Resources (FHIR), designed for transactional data exchange and app integration, and the Observational Medical Outcomes Partnership (OMOP) common data model (CDM), developed to harmonize multi-site analytics [226,227]. Increasingly, hybrid approaches combine FHIR for resource-level exchange with OMOP for large-scale analytic queries (Table 3).

Recent work illustrates both technical progress and the need for governance. A study described a reproducible FHIR mapping pipeline that ensures consistency when converting heterogeneous clinical data into standard resources [105]. Reviews emphasize that open standards are necessary but insufficient; without metadata, provenance, and ontology alignment (e.g., SNOMED CT, LOINC), true interoperability remains elusive [228,229]. A practical illustration of these principles is the *recruIT* platform, which integrates FHIR for exchange and OMOP for analytic cohorting, providing a dual-model design that could inform registry-embedded clinical trials [230,231]. Such dual-model architectures exemplify how interoperability standards can transition from theoretical frameworks to operational platforms, supporting both point-of-care decision support and large-scale causal inference.

6.3. Privacy-Preserving Collaboration Through Federated Learning

Decentralized data environments present a challenge for precision rheumatology, as many practices cannot legally or technically centralize patient-level data. Federated Learning (FL) provides a potential solution by enabling distributed model training where only model parameters are exchanged across sites.

Recent reviews highlight the growing maturity of FL frameworks, describing governance structures for role assignment, auditing, and balancing privacy with fairness [20,232]. Empirical demonstrations have validated FL in clinical prediction tasks ranging from intensive care outcomes to imaging diagnostics, even under conditions of substantial data heterogeneity [233,234]. Of particular relevance is the development of federated target trial emulation (FL-TTE), which allows comparative effectiveness research (e.g., evaluating anti-TNF versus JAK inhibitor strategies in RA) without sharing individual patient records [235,236]. This innovation positions FL as a bridge between traditional observational research and pragmatic randomized trials, enabling registry-linked networks to generate causal evidence while preserving data sovereignty (Table 3).

6.4. Pitfalls of Multisite Modeling and Mitigation Strategies

Despite technical advances, multi-site modeling is vulnerable to well-documented pitfalls. Covariate shift degrades external validity and generalizability [237,238]. Acquisition drift, including changes in imaging protocols or laboratory platforms, similarly threatens longitudinal stability.

Emerging strategies seek to mitigate these risks. FedWeight, a density-based reweighting method for FL, improved cross-site calibration [239]. COLA-GLMM, a one-shot distributed algorithm for generalized linear mixed models, achieved exact multi-site inference with minimal communication overhead [240]. Complementary advances in secure aggregation, fairness-aware updates, and transparent model documentation ("model cards") provide operational scaffolding for responsible deployment (Table 3) [241].

Best practices now emphasize coupling technical safeguards with procedural governance: site-specific temporal validation, routine calibration checks, and embedding retraining triggers within registry-driven QI cycles. Moreover, multidisciplinary governance boards, including clinicians, statisticians, ethicists, and patients—are increasingly recommended to ensure that technical adjustments align with clinical priorities and ethical standards.

6.5. Case Illustration: Predicting RA Disease Activity Using RISE

Early feasibility studies, often reported through ACR Meeting Abstracts, have demonstrated that machine learning models trained on routine EHR and patient-reported outcomes (PROs) can classify and forecast RA disease activity [242]. These prototypes established a methodological template: define computable endpoints (e.g., DAS28), derive features from EHR and PRO data, and validate performance in temporally held-out datasets.

A modernized, RISE-embedded pipeline would extend this framework through rigorous standards. Data harmonization should map incoming records into canonical FHIR and OMOP layers, with provenance and ontology alignment maintained [226,243,244]. Cohort definitions and computable phenotypes should be pre-registered in the RISE Hub to ensure reproducibility [18]. Modeling strategies should benchmark interpretable methods (e.g., regularized GLMs, gradient boosting) alongside multimodal extensions incorporating labs, medications, and potentially imaging or digital biomarkers. Validation protocols should enforce both temporal and cross-site external testing, with calibration, decision-curve analyses, and equity stratification by demographic subgroups (Table 3).

Finally, deployment should surface predictions within the RISE dashboard as risk-stratified panels, with drift detection and recalibration embedded into routine QI cycles. Such an infrastructure would not only accelerate clinical adoption but also serve as a testbed for adaptive trial designs, where predictive models guide enrichment strategies, treatment allocation, and real-world monitoring of therapeutic impact.

This case illustrates a practical pathway for embedding precision analytics into daily rheumatology care, aligning methodological rigor with operational feasibility. Ultimately, the integration of RISE-like infrastructures with interoperable standards, federated learning, and robust governance represents the cornerstone of a scalable precision rheumatology ecosystem.

6.6. Implementation Costs and Regulatory Readiness

The translation of AI-enabled frameworks from research to clinical rheumatology faces substantial operational and regulatory challenges. Implementation costs extend beyond algorithm development to include data-storage infrastructure, cybersecurity safeguards, and continuous validation expenses required to sustain model reliability. Large-scale deployment also necessitates clinician re-training, workflow redesign, and ongoing technical support to ensure that AI recommendations integrate seamlessly with established EHR systems. Interoperability remains a persistent barrier, as variability across hospital information systems often impedes the standardized exchange of multimodal data essential for model recalibration and longitudinal monitoring [245–247].

Economic and policy dimensions further complicate adoption. The absence of clear reimbursement pathways for algorithmic decision-support tools constrains institutional uptake, while cost–benefit analyses remain scarce. Initial evidence suggests that maintenance of federated-learning infrastructures and data-governance frameworks may rival, or even exceed, traditional trial costs—underscoring the need for early economic evaluation alongside technical validation [245,248].

Regulatory oversight is also evolving rapidly. The U.S. Food and Drug Administration's Software as a Medical Device (SaMD) AI/ML Action Plan and the European Medicines Agency's Guideline on Artificial Intelligence in Medicinal Products emphasize adaptive regulatory pathways, lifecycle monitoring, and transparency in model updating [249,250]. These initiatives reflect a shift toward continuous assurance, where safety, equity, and performance are monitored dynamically rather than through static pre-approval processes.

Ultimately, the feasibility of clinical AI implementation depends as much on system-level readiness as on algorithmic sophistication. Sustainable adoption will require parallel investment in digital infrastructure, regulatory harmonization, and reimbursement models that recognize the value of predictive analytics within precision-medicine care pathways.

7. Standards and Study Designs for AI Prediction Models in Clinical Research

7.1. Core Reporting Standards for Prediction Models

Transparent and comprehensive reporting remains a cornerstone of credibility, reproducibility, and clinical translation for artificial intelligence (AI)-based prediction models (Figure 3). The TRIPOD+AI extension has established itself as the reference standard for documenting prediction model development and validation in clinical research. This guideline builds on the original TRIPOD framework by requiring detailed disclosure of data provenance and linkage processes, strategies for addressing missing data, and explicit specifications of feature engineering, model hyperparameters, and final architecture. Furthermore, TRIPOD+AI mandates robust internal validation procedures—such as bootstrap resampling or nested cross-validation—alongside external and temporal validation to ensure model transportability. Emphasis is placed on calibration analyses and the integration of clinical utility assessments (e.g., decision-analytic frameworks), which enable reviewers and clinicians to evaluate both methodological rigor and real-world applicability [22].

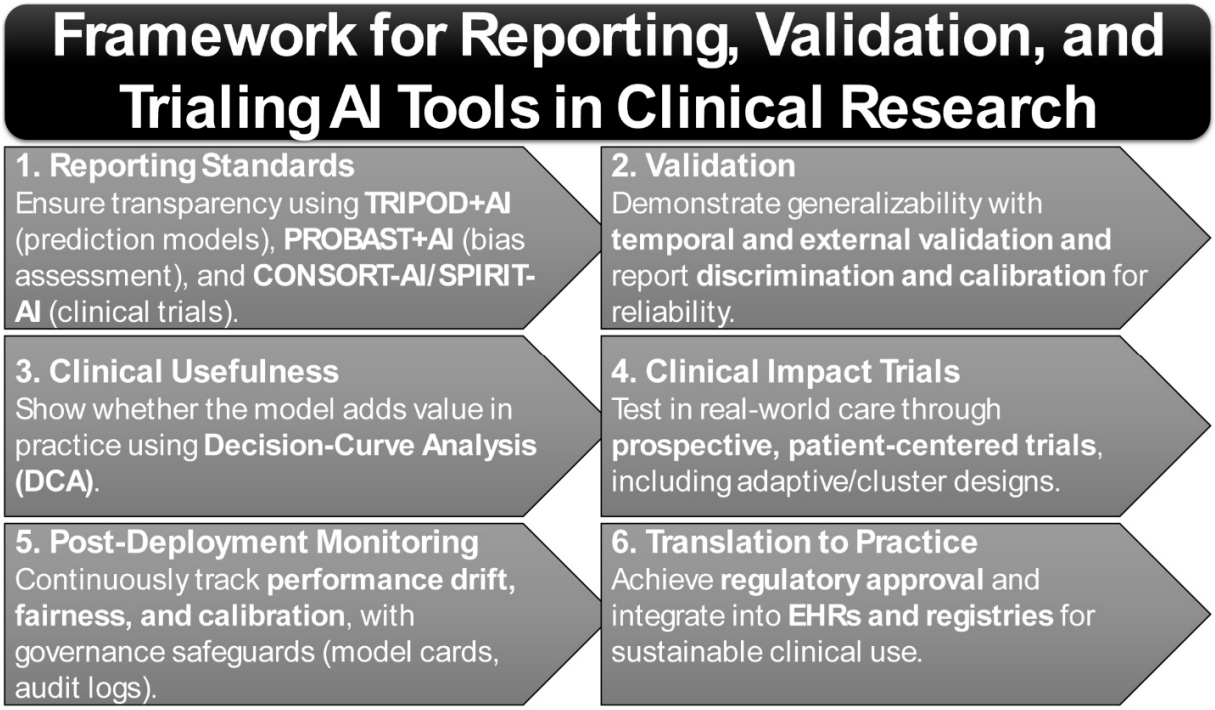


Figure 3. Framework for Reporting, Validation, and Trialing AI Tools in Clinical Research.

Complementing reporting guidelines, the PROBAST+AI (2025) tool provides an updated framework for systematically evaluating risk of bias and applicability in AI prediction models. By introducing specific signaling questions related to dataset shift, preprocessing leakage, subgroup fairness, and model selection, PROBAST+AI directly addresses methodological vulnerabilities unique to machine learning-based approaches. This tool is increasingly recommended for peer reviewers, systematic reviewers, and clinical methodologists to ensure that AI models are critically appraised with rigor comparable to traditional epidemiological studies [21]. Crucially, PROBAST+AI has also emphasized fairness auditing across demographic strata, ensuring that model performance disparities—often hidden by aggregate metrics—are explicitly reported and mitigated.

When AI constitutes the intervention itself rather than an auxiliary decision-support tool, the CONSORT-AI and SPIRIT-AI extensions define the reporting standards for randomized controlled trial reports and protocols. These frameworks require explicit specification

of the intended clinical role of the AI system, documentation of human–AI interaction within the trial, prespecified monitoring strategies for adaptive or learning systems, and transparent disclosure of algorithm updates. Adherence to these guidelines not only strengthens methodological integrity but also facilitates regulatory evaluation and clinical acceptance [23,251]. Recent trial protocols have also begun incorporating "algorithmic accountability statements," disclosing model update frequency, governance structures, and pathways for patient feedback—features increasingly demanded by regulators and ethics boards.

7.2. Study Design Foundations: Reviewer Expectations and Best Practices

From the perspective of peer reviewers and regulators, external and temporal validation is now regarded as minimum methodological requirements (Figure 3). Unlike traditional random-split validation, contemporary standards demand assessment on temporally distinct cohorts (reflecting later patient populations) and geographically external datasets (reflecting different clinical environments). Such practices provide robust evidence for generalizability and resilience against dataset shift. Alongside discrimination indices such as the area under the receiver operating characteristic curve (AUC) and precision-recall AUC, rigorous calibration reporting—including slope, intercept, and visual calibration plots across clinically relevant risk ranges—is essential. The BMJ's instructional series on prediction model development and validation remains a key reference point for best practices in this area [252,253].

Beyond validation, decision-curve analysis (DCA) has emerged as a central tool for quantifying the clinical utility of AI models. By assessing net benefit across plausible decision thresholds, DCA addresses a critical limitation of discrimination metrics, which do not reflect the consequences of clinical decision-making. Increasingly, reviewers expect inclusion of DCA plots with correct interpretation, while best-practice tutorials highlight frequent errors such as threshold misspecification or misinterpretation of net benefit curves [254–256]. In addition, newer utility frameworks such as cost–benefit analysis and value-of-information modeling are being paired with DCA to provide health-economic perspectives on whether AI adoption meaningfully improves care efficiency (Figure 3) [245].

The next methodological frontier involves clinical impact trials, designed to evaluate whether AI-driven tools produce measurable improvements in patient outcomes. When systems are intended to alter care delivery—such as accelerating rheumatology triage or optimizing treatment selection—prospective designs such as stepped-wedge or cluster-randomized trials are recommended. Importantly, these studies should prioritize patient-centered outcomes, including time-to-specialist assessment, flare prevention, or remission rates, rather than algorithmic performance metrics alone. Emerging adaptive trial designs also allow for real-time updating of AI tools under prespecified governance frameworks, ensuring that continuously learning systems can be evaluated safely without compromising trial integrity. For continuously adaptive systems (e.g., large language model-based triage tools), dynamic monitoring frameworks that integrate model updates under prespecified safety and governance conditions are becoming essential (Figure 3) [257–259].

Finally, post-deployment monitoring has shifted from being an optional safeguard to a fundamental requirement for AI implementation in clinical practice. Deployed models must be monitored for performance drift, calibration decay, subgroup fairness erosion, and degradation in label quality (Figure 3). Recent position papers emphasize the importance of statistically efficient, label-sparing surveillance strategies with clearly defined triggers for recalibration, retraining, or system rollback. These mechanisms should be integrated into formal governance frameworks, incorporating tools such as model cards, audit logs, and institutional oversight committees [128,140,260]. Regulatory momentum is also mov-

ing toward "continuous assurance," where monitoring data are periodically submitted to oversight agencies, ensuring that AI systems remain safe and equitable throughout their lifecycle.

7.3. Methodological Appraisal and Evidence Grading

In line with emerging best practices, the studies discussed were appraised with reference to the TRIPOD-AI and PROBAST-AI frameworks to evaluate reporting quality, calibration transparency, and potential sources of bias. These standards emphasize reproducibility, calibration fidelity, and bias-mitigation practices that remain inconsistently implemented across current AI studies in rheumatology. Most published models align with Technology Readiness Levels (TRL) 4–6, reflecting pre-clinical or validation phases rather than tools ready for clinical deployment [16,22,252]. Despite promising discrimination metrics, many studies are limited by modest cohort sizes, single-center designs, and incomplete external or temporal validation, contributing to uncertainty in real-world generalizability. Common sources of bias include overfitting in high-dimensional data, insufficient stratified calibration across demographic subgroups, and selective reporting of best-performing metrics.

Imaging-based AI systems generally demonstrate greater reproducibility owing to standardized acquisition protocols, whereas multi-omics and digital-biomarker models exhibit higher variability in feature selection and preprocessing pipelines. These contrasts underscore the need for harmonized validation frameworks and transparent reporting to enhance comparability across studies [261]. Future AI research in AIRD should routinely disclose TRL stage, external-validation approach, and fairness-audit outcomes, enabling journals and regulators to gauge the maturity of evidence relative to deployment readiness. Embedding such structured methodological critique within AI research will accelerate the transition from experimental modeling to clinically trusted decision-support systems.

8. Equity and Portability in Polygenic Risk and AI Models: Addressing Ancestry Gaps and Bias in Precision Medicine

8.1. PRS Portability and Ancestry Gaps

Polygenic risk scores (PRSs) frequently exhibit attenuated predictive accuracy when applied to individuals from ancestries other than those in which the underlying GWAS discovery was conducted. This deterioration arises due to differences in allele frequencies, linkage disequilibrium (LD) architecture, and cohort-specific artifacts, which distort effect sizes and impair both discrimination and calibration. For example, a recent large-scale evaluation of PRS performance for 14 traits across four ancestry groups (Africans, Europeans, East Asians, South Asians) found that scores trained on European datasets lost ~50% or more of their predictive power when applied to African or East Asian populations. However, when ancestry-specific submodels or ancestry-aware training was used, performance improved substantially though still lagged European benchmarks (Table 4) [262].

Another methodological advance is represented by the "X-Wing" framework, which quantifies local genetic correlation across populations and weights contributions to combine population-specific PRSs with summary statistics alone. This was shown to yield relative gains in R^2 of 14–119% in non-European populations compared to conventional single-ancestry PRS approaches [263]. These innovations highlight a shift from "one-size-fits-all" genetics toward ancestry-aware precision genomics that can equitably inform risk prediction across global populations.

Table 4. Equity, Generalizability, and Safety Considerations for AI in Autoimmune Rheumatic Diseases.

Domain	Key Challenges	Risks if Unaddressed	Mitigation Strategies	Minimum Reporting Set
PRS Portability & Ancestry Gaps	Loss of accuracy across ancestries due to allele-frequency and LD differences; poor calibration across sex, age, and SES strata.	Worsening health disparities; misleading risk predictions; inequitable clinical recommendations.	Multi-ancestry GWAS; ancestry-aware modeling; ancestry/site-specific recalibration; subgroup reporting of performance.	Ancestry-stratified R^2/AUC; calibration curves; decision-curve analysis by subgroup.
Data Drift & Bias Audits	Model degradation due to covariate, prior-probability, acquisition, and concept drift; hidden biases in datasets.	Silent failure of AI models; unfair treatment allocation; erosion of clinical trust.	Drift detectors with temporal validation; scheduled recalibration; bias audits; integration of bias dashboards in registries.	Performance stratified by sex, age, ancestry, SES proxies, and site; bias dashboard outputs.
Governance & Safety by Design	Ensuring continuous safety across lifecycles: bias detection, calibration monitoring, clinician-defer thresholds, and accountability.	Unsafe deployment; lack of transparency; patient harm; regulatory non-compliance.	Pre-deployment bias assessment; live calibration monitoring; safety circuit-breakers; audit trails; periodic re-validation and changelogs.	Intended-use statement; subgroup performance reports; update logs with versioning; governance protocols.

These findings underscore that multi-ancestry GWAS, ancestry-aware modeling, and local/site-specific recalibration are not optional but essential for responsible translational PRS work. Moreover, subgroup reporting (sex, age, ancestry, SES proxies, recruitment site) is required to detect calibration instability or bias across strata. Without systematic subgroup analyses, there is a substantial risk that PRS-driven clinical tools will exacerbate health disparities, particularly in underrepresented populations.

8.2. Data Drift, Bias Audits, and Transparent Documentation

Even a well-validated PRS or AI model will degrade over time or across settings if it is not safeguarded against various forms of drift: covariate drift, prior-probability drift, concept drift, and changes in data acquisition protocols. Studies in medical imaging show that automatic acquisition drift correction helps maintain performance, but such procedures must be accompanied by robust supervision, retraining triggers, or fallback systems to avoid "silent failures" where performance drops go undetected [140,264,265].

Bias audits form another critical safety mechanism. Recent work proposes frameworks for privacy-preserving subgroup audits, operational bias dashboards, and embedding equity metrics into clinical AI QA cycles. Transparent documentation via model cards or "nutrition labels" is also maturing. The Coalition for Health AI (CHAI) has released an open-source applied model card template (Aidoc example) that includes sections on data provenance, intended-use populations, fairness, known risks and limitations, performance in subgroups, and maintenance/versioning (Table 4) [266–268].

Post-deployment monitoring must include scheduled validations (temporal/site), performance surveillance, and predefined alerting thresholds. Importantly, reviews now argue that AI governance in healthcare should parallel pharmacovigilance systems, with structured monitoring playbooks, mandatory safety reporting, and continuous audit trails

to ensure accountability. This reframing positions clinical AI not as static "devices" but as evolving interventions requiring lifecycle oversight.

8.3. Governance and Safety by Design

Governance must be embedded across all stages of the model life cycle. Before deployment, developers should perform formal bias assessments across major demographic axes, simulate distribution shifts (e.g., by holding out data from future time periods or distinct sites), and define intended-use statements and contraindications in documentation [269]. Best practices now recommend the inclusion of patient and clinician stakeholders in early governance discussions, ensuring that the design of AI systems aligns with real-world priorities and ethical expectations.

During deployment, monitoring of calibration (e.g., comparing predicted vs. observed risk over time), safety "circuit-breakers" (e.g., thresholded uncertainty where the model defers to clinician rather than forcing a possibly erroneous prediction), and comprehensive audit trails of predictions and model updates are essential for accountability [270]. Such mechanisms not only enhance trust but also reduce automation bias, ensuring that clinicians retain the ultimate authority in ambiguous or high-risk scenarios (Table 4).

After deployment, models should undergo scheduled revalidation (site-/time-specific), use label-efficient monitoring strategies (to reduce the annotation burden but still detect drift and bias), and release public changelogs whenever model versions change. For models that adapt in situ, dynamic or adaptive trial designs may provide a way to learn safely while controlling risk [75,271,272]. Regulators are increasingly calling for "continuous assurance frameworks," in which adaptive AI models must submit periodic evidence of safety, calibration, and equity performance as part of their lifecycle management—an approach likely to become central to future AI governance in healthcare.

9. Future Directions

9.1. Multimodal Fusion (Omics, Imaging, and Digital Phenotypes)

The evolution of precision rheumatology is increasingly favoring multimodal fusion frameworks that integrate omics (genome, transcriptome, proteome), quantitative imaging (e.g., ultrasound, MRI, capillaroscopy), and digital phenotypes (wearables, smartphone-based functional tasks). Fusion methodologies—early fusion (feature concatenation), late fusion (decision-level aggregation), and hybrid attention-based architectures to yield superior discrimination and calibration compared to single-modality models, particularly when external validation and clinical workflow integration are prioritized.

Although many of the most advanced examples originate outside rheumatology, their methodological components—feature curation, label hygiene, and cross-site evaluation—are directly applicable to AIRD programs. For AIRD, a recommended strategy is to establish a minimal fusion core: routine electronic health record (EHR) and laboratory data combined with standardized imaging outputs (such as RAMRIS-aligned MRI or ultrasound features) and a concise proteomic panel; onto this core, add digital phenotypes (e.g., smartphone-based range-of-motion tasks; step-to-symptom coupling) to capture longitudinal fluctuations. Crucially, studies should pre-register ablation experiments modality by modality and report the marginal utility of each data modality, so that the additional cost and effort of collecting extra modalities are justified in clinical contexts.

9.2. Mechanism-Aware Machine Learning to Guide Drug Targeting

A well-recognized limitation of purely statistical models is their lack of inherent biological interpretability. Mechanism-aware machine learning seeks to address this limitation by explicitly embedding disease pathophysiology into predictive frameworks. In SLE,

this paradigm has recently shown considerable promise. Emerging work highlights the profound heterogeneity across type I, II, and III interferon (IFN) biology, demonstrating that measured IFN gene-expression signatures often fail to correspond precisely with the underlying functional IFN activity. This discordance offers an important explanation for the variable therapeutic responses observed with IFN-pathway inhibitors, such as anifrolumab, and emphasizes the necessity of integrating mechanistic insights into model design to achieve more reliable patient stratification [179].

Complementary to that, quantitative systems pharmacology (QSP) models are growing more robust. These integrate patient-level and aggregated trial data to simulate IFN-inducible gene dynamics under different therapeutic regimens, providing foundations for patient stratification, dose-optimization, and exploration of drug combinations. Such models, aligned with the model-informed drug development (MIDD) paradigm, promise biologically interpretable decision support systems capable of both forecasting and prescribing.

9.3. Digital Twins, N-of-1 Trials, Adaptive Platforms, and Home Testing

Digital twins, defined as detailed computational replicas of patient physiology that integrate multi-scale biological, clinical, and environmental data, are emerging as transformative tools in AIRD research. Recent studies illustrate their potential, such as the construction of a modular, multicellular virtual twin of the arthritic joint encompassing more than 1000 biomolecules. This model was validated against gene-expression data and subsequently applied to interrogate both existing therapeutic agents and novel candidate targets, underscoring the value of digital twins in accelerating translational discovery and precision medicine [273].

The resurgence of N-of-1 designs offers potential for individualized treatment evaluation. Platforms enabling patient-level comparisons using app-based outcome capture (e.g., Arthritis Power) allow for adaptive dosing decisions and trajectories that respond to individual response dynamics.

Adaptive and platform trial designs are also becoming more prevalent in rheumatology. EULAR has issued guidance supporting such designs; for instance, CONQUEST in systemic sclerosis-associated interstitial lung disease (SSc-ILD) is among the first platform trials in the field, facilitating evaluation of multiple interventions against a shared control and permitting response-adaptive features. Prevention trials (e.g., RA risk-cohort platforms) align with this trend via shared infrastructure and biomarker-guided enrichment.

Home and remote monitoring technologies are advancing at a rapid pace, with studies demonstrating that patients can reliably self-test inflammatory markers such as C-reactive protein (CRP) and white blood cell counts, including through dried blood spot sampling. Smartphone-read lateral flow assays (LFAs), augmented with machine learning-based quantification, now achieve close concordance with conventional laboratory assays, while innovations in electrochemical and distance-based formats are further enhancing sensitivity without compromising portability. Wearable devices add another layer of promise, with preclinical studies in RA models showing on-body inflammatory feedback and closed-loop modulation, thereby foreshadowing the emergence of theranostic systems that integrate diagnostics and interventions. Looking ahead, remote monitoring is expected to evolve toward multimodal integration, in which self-testing platforms, wearable biosensors, and smartphone-based analytics converge into comprehensive disease activity dashboards capable of supporting real-time treat-to-target strategies, early flare detection, and adaptive therapy adjustment in both clinical and home settings. Future priorities will include large-scale validation across diverse populations, the establishment of regulatory and technical frameworks for seamless integration of device-derived data into electronic health records, and equity-oriented implementation strategies to ensure accessibility in resource-limited

contexts. Collectively, the convergence of biosensing, digital health infrastructure, and AI-driven analytics is poised to transform remote monitoring from a supplementary adjunct into a central pillar of precision rheumatology.

10. Conclusions

Precision rheumatology has now advanced beyond theoretical aspiration to demonstrable early-stage implementation. Several milestones illustrate this transition. First, preventive immunomodulation in at-risk RA has moved from proof-of-concept to randomized trial evidence: abatacept delayed or reduced the onset of clinical RA in high-risk cohorts (e.g., APIPPRA and ARIAA). Second, artificial intelligence (AI)-augmented imaging is now capable of reader-assist scoring for synovitis and microvascular changes, improving reproducibility in musculoskeletal ultrasound and MRI interpretation. Third, digital biomarkers derived from smartphones and wearable devices—such as motion signatures, circadian activity, and patient-reported passive sensing—are extending disease monitoring beyond the clinic, offering opportunities for remote and continuous assessment.

However, the deployment of AI for therapeutic decision support remains premature. No system for drug-selection is currently "clinic-ready" without rigorous methodological and regulatory evaluation. This requires adherence to evolving standards: transparent model reporting through TRIPOD+AI, systematic risk-of-bias appraisal with PROBAST+AI, external and temporal validation across diverse populations, calibration testing, and clinical utility evaluation using decision-analytic frameworks. Beyond these prerequisites, true clinical integration necessitates prospective impact trials, alignment with professional guidelines, and post-deployment monitoring for dataset drift, fairness, and equity of outcomes.

The most actionable near-term opportunity lies in embedding multimodal, mechanism-aware models into rheumatology registries and electronic health record (EHR) platforms, where they can be linked to longitudinal outcome data. Coupled with N-of-1 methodologies, adaptive trial designs, and home-based biosample or digital testing, such systems could shorten treat-to-target cycles and personalize therapy adjustment at scale. The trajectory of precision rheumatology thus reflects a broader paradigm shift: from promising pilot studies to the construction of reliable, equitable, and continuously monitored decision-support ecosystems.

Looking ahead, the integration of AI with molecular, imaging, and digital biomarkers holds the potential to transform the diagnostic and therapeutic landscape of immune-mediated inflammatory diseases (IMIDs). Predictive modeling approaches can refine early disease interception, identify preclinical immune activation, and support individualized therapy selection based on mechanistic profiles rather than population averages. The convergence of these technologies within learning health systems, where model outputs are continuously validated against real-world data, will enable a dynamic cycle of feedback, recalibration, and improvement. Such an ecosystem promises not only earlier and more precise diagnosis but also sustained personalization of care, aligning therapeutic intervention with the molecular and behavioral signatures of each patient.

Author Contributions: O.A.A.-E. and M.M.N. equally contributed to the design and writing of the main manuscript text. 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: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Tanaka, H.; Okada, Y.; Nakayamada, S.; Miyazaki, Y.; Sonehara, K.; Namba, S.; Honda, S.; Shirai, Y.; Yamamoto, K.; Kubo, S.; et al. Extracting immunological and clinical heterogeneity across autoimmune rheumatic diseases by cohort-wide immunophenotyping. *Ann. Rheum. Dis.* **2024**, *83*, 242–252. [CrossRef]
2. Al-Ewaidat, O.A.; Naffaa, M.M. Stroke risk in rheumatoid arthritis patients: Exploring connections and implications for patient care. *Clin. Exp. Med.* **2024**, *24*, 30. [CrossRef]
3. Bilgin, E. Current application, possibilities, and challenges of artificial intelligence in the management of rheumatoid arthritis, axial spondyloarthritis, and psoriatic arthritis. *Ther. Adv. Musculoskelet. Dis.* **2025**, *17*, 1759720X251343579. [CrossRef] [PubMed]
4. Stafford, I.S.; Kellermann, M.; Mossotto, E.; Beattie, R.M.; MacArthur, B.D.; Ennis, S. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit. Med.* **2020**, *3*, 30. [CrossRef]
5. Smolen, J.S.; Aletaha, D.; McInnes, I.B. Rheumatoid arthritis. *Lancet* **2016**, *388*, 2023–2038. [CrossRef]
6. Tsokos, G.C.; Lo, M.S.; Reis, P.C.; Sullivan, K.E. New insights into the immunopathogenesis of systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **2016**, *12*, 716–730. [CrossRef] [PubMed]
7. Kumar, A.; Vasdev, V.; Patnaik, S.K.; Bhatt, S.; Singh, R.; Bhayana, A.; Hegde, A.; Kumar, A. The diagnostic utility of rheumatoid factor and anticitrullinated protein antibody for rheumatoid arthritis in the Indian population. *Med. J. Armed Forces India* **2022**, *78*, S69–S74. [CrossRef]
8. Al-Ewaidat, O.A.; Naffaa, M.M. Deciphering Mechanisms, Prevention Strategies, Management Plans, Medications, and Research Techniques for Strokes in Systemic Lupus Erythematosus. *Medicines* **2024**, *11*, 15. [CrossRef] [PubMed]
9. Dai, X.; Fan, Y.; Zhao, X. Systemic lupus erythematosus: Updated insights on the pathogenesis, diagnosis, prevention and therapeutics. *Signal Transduct. Target. Ther.* **2025**, *10*, 102. [CrossRef]
10. Findeisen, K.E.; Sewell, J.; Ostor, A.J.K. Biological Therapies for Rheumatoid Arthritis: An Overview for the Clinician. *Biologics* **2021**, *15*, 343–352. [CrossRef]
11. Salaffi, F.; Carotti, M.; Di Carlo, M.; Ceccarelli, L.; Farah, S.; Poliseno, A.C.; Di Matteo, A.; Bandinelli, F.; Giovagnoni, A. Magnetic Resonance Imaging (MRI)-Based Semi-Quantitative Methods for Rheumatoid Arthritis: From Scoring to Measurement. *J. Clin. Med.* **2024**, *13*, 4137. [CrossRef]
12. Parodis, I.; Lindblom, J.; Toro-Dominguez, D.; Beretta, L.; Borghi, M.O.; Castillo, J.; Carnero-Montoro, E.; Enman, Y.; Mohan, C.; Alarcon-Riquelme, M.E.; et al. Interferon and B-cell Signatures Inform Precision Medicine in Lupus Nephritis. *Kidney Int. Rep.* **2024**, *9*, 1817–1835. [CrossRef] [PubMed]
13. Creagh, A.P.; Hamy, V.; Yuan, H.; Mertes, G.; Tomlinson, R.; Chen, W.H.; Williams, R.; Llop, C.; Yee, C.; Duh, M.S.; et al. Digital health technologies and machine learning augment patient reported outcomes to remotely characterise rheumatoid arthritis. *NPJ Digit. Med.* **2024**, *7*, 33. [CrossRef] [PubMed]
14. Creagh, A.P.; Dondelinger, F.; Lipsmeier, F.; Lindemann, M.; De Vos, M. Longitudinal Trend Monitoring of Multiple Sclerosis Ambulation Using Smartphones. *IEEE Open J. Eng. Med. Biol.* **2022**, *3*, 202–210. [CrossRef] [PubMed]
15. Liu, S.; Liu, Y.; Li, M.; Shang, S.; Cao, Y.; Shen, X.; Huang, C. Artificial intelligence in autoimmune diseases: A bibliometric exploration of the past two decades. *Front. Immunol.* **2025**, *16*, 1525462. [CrossRef]
16. Sequi-Sabater, J.M.; Benavent, D. Artificial intelligence in rheumatology research: What is it good for? *RMD Open* **2025**, *11*, e004309. [CrossRef]
17. Hammam, N.; Izadi, Z.; Li, J.; Evans, M.; Kay, J.; Shiboski, S.; Schmajuk, G.; Yazdany, J. The Relationship Between Electronic Health Record System and Performance on Quality Measures in the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) Registry: Observational Study. *JMIR Med. Inform.* **2021**, *9*, e31186. [CrossRef]
18. Yazdany, J.; Bansback, N.; Clowse, M.; Collier, D.; Law, K.; Liao, K.P.; Michaud, K.; Morgan, E.M.; Oates, J.C.; Orozco, C.; et al. Rheumatology Informatics System for Effectiveness: A National Informatics-Enabled Registry for Quality Improvement. *Arthritis Care Res.* **2016**, *68*, 1866–1873. [CrossRef]
19. Prot, V.; Aguilera, H.M.; Skallerud, B.; Persson, R.; Urheim, S. A method for non-invasive estimation of mitral valve annular regional strains. *Comput. Biol. Med.* **2025**, *187*, 109773. [CrossRef]
20. 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]
21. Moons, K.G.M.; Damen, J.A.A.; Kaul, T.; Hooft, L.; Andaur Navarro, C.; Dhiman, P.; Beam, A.L.; Van Calster, B.; Celi, L.A.; Denaxas, S.; et al. PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ* **2025**, *388*, e082505. [CrossRef] [PubMed]
22. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Ghassemi, M.; Liu, X.; Reitsma, J.B.; van Smeden, M.; et al. TRIPOD + AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **2024**, *385*, e078378. [CrossRef]

23. Ibrahim, H.; Liu, X.; Rivera, S.C.; Moher, D.; Chan, A.W.; Sydes, M.R.; Calvert, M.J.; Denniston, A.K. Reporting guidelines for clinical trials of artificial intelligence interventions: The SPIRIT-AI and CONSORT-AI guidelines. *Trials* **2021**, *22*, 11. [CrossRef]
24. Ribeiro Junior, H.L.; Nepomuceno, F.; Pessoa, C.D.O. AI in clinical trials is missing from CONSORT and SPIRIT 2025 guidelines. *Lancet* **2025**, *406*, 25. [CrossRef]
25. Hammam, N.; Evans, M.; Morgan, E.; Reimold, A.; Anastasiou, C.; Kay, J.L.; Yazdany, J.; Schmajuk, G. Treatment of Sarcoidosis in US Rheumatology Practices: Data from the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) Registry. *Arthritis Care Res.* **2022**, *74*, 371–376. [CrossRef]
26. Guthridge, J.M.; Wagner, C.A.; James, J.A. The promise of precision medicine in rheumatology. *Nat. Med.* **2022**, *28*, 1363–1371. [CrossRef]
27. Birtane, M.; Yavuz, S.; Tastekin, N. Laboratory evaluation in rheumatic diseases. *World J. Methodol.* **2017**, *7*, 1–8. [CrossRef]
28. Martinez-Prat, L.; Nissen, M.J.; Lamacchia, C.; Bentow, C.; Cesana, L.; Roux-Lombard, P.; Gabay, C.; Mahler, M. Comparison of Serological Biomarkers in Rheumatoid Arthritis and Their Combination to Improve Diagnostic Performance. *Front. Immunol.* **2018**, *9*, 1113. [CrossRef]
29. Steiner, G.; Toes, R.E.M. Autoantibodies in rheumatoid arthritis—Rheumatoid factor, anticitrullinated protein antibodies and beyond. *Curr. Opin. Rheumatol.* **2024**, *36*, 217–224. [CrossRef] [PubMed]
30. Brink, M.; Hansson, M.; Mathsson-Alm, L.; Wijayatunga, P.; Verheul, M.K.; Trouw, L.A.; Holmdahl, R.; Ronnelid, J.; Klareskog, L.; Rantapaa-Dahlqvist, S. Rheumatoid factor isotypes in relation to antibodies against citrullinated peptides and carbamylated proteins before the onset of rheumatoid arthritis. *Arthritis Res. Ther.* **2016**, *18*, 43. [CrossRef] [PubMed]
31. Perera, J.; Delrosso, C.A.; Nerviani, A.; Pitzalis, C. Clinical Phenotypes, Serological Biomarkers, and Synovial Features Defining Seropositive and Seronegative Rheumatoid Arthritis: A Literature Review. *Cells* **2024**, *13*, 743. [CrossRef]
32. Avouac, J.; Kay, J.; Choy, E. Personalised treatment of rheumatoid arthritis based on cytokine profiles and synovial tissue signatures: Potentials and challenges. *Semin. Arthritis Rheum.* **2025**, *73*, 152740. [CrossRef]
33. Rayner, F.; Hiu, S.; Melville, A.; Bigirumurame, T.; Anderson, A.; Dyke, B.; Kerrigan, S.; McGucken, A.; Prichard, J.; Shahrokhabadi, M.S.; et al. Clinical predictors of flare and drug-free remission in rheumatoid arthritis: Preliminary results from the prospective BIO-FLARE experimental medicine study. *BMJ Open* **2025**, *15*, e092478. [CrossRef]
34. Huang, X.; Luu, L.D.W.; Jia, N.; Zhu, J.; Fu, J.; Xiao, F.; Liu, C.; Li, S.; Shu, G.; Hou, J.; et al. Multi-Platform Omics Analysis Reveals Molecular Signatures for Pathogenesis and Activity of Systemic Lupus Erythematosus. *Front. Immunol.* **2022**, *13*, 833699. [CrossRef]
35. Huang, H.; Sun, X.; Zhang, Q.; Liu, C.; Cao, X.; Zhang, D.; Wang, G.; Pu, C. Combined serum IFN-gamma and IL-22 levels as predictive biomarkers for hepatocellular carcinoma risk: A clinical investigation. *Biomed. Rep.* **2025**, *23*, 149. [CrossRef]
36. Pisetsky, D.S. Pathogenesis of autoimmune disease. *Nat. Rev. Nephrol.* **2023**, *19*, 509–524. [CrossRef] [PubMed]
37. Kutsuna, Y.J.; Aibara, N.; Hashizume, J.; Kawarabayashi, S.; Tamai, M.; Miyata, J.; Yoshifuji, H.; Miyamoto, H.; Sato, K.; Kodama, Y.; et al. Identification of immune complex antigens that are detected prior to early rheumatoid arthritis symptoms and increase with disease progression: Comprehensive serum immune complexome analysis to identify candidate disease biomarkers in health checkup cohort study. *Clin. Immunol.* **2025**, *281*, 110591. [CrossRef] [PubMed]
38. Lerga-Jaso, J.; Terpolovsky, A.; Novkovic, B.; Osama, A.; Manson, C.; Bohn, S.; De Marino, A.; Kunitomi, M.; Yazdi, P.G. Optimization of multi-ancestry polygenic risk score disease prediction models. *Sci. Rep.* **2025**, *15*, 17495. [CrossRef] [PubMed]
39. Wu, T.S.; Chen, Y.J.; Hsiung, C.N.; Mao, C.L.; Wei, C.Y.; Chen, I.C.; Kao, C.M.; Hsiao, T.H.; Huang, W.N.; Chen, Y.H.; et al. Polygenic risk scores of rheumatoid arthritis associated with seropositivity and bone erosions in a Taiwanese population. *Sci. Rep.* **2025**, *15*, 25700. [CrossRef]
40. Honda, S.; Ikari, K.; Yano, K.; Terao, C.; Tanaka, E.; Harigai, M.; Kochi, Y. Association of Polygenic Risk Scores with Radiographic Progression in Patients with Rheumatoid Arthritis. *Arthritis Rheumatol.* **2022**, *74*, 791–800. [CrossRef]
41. Ishigaki, K.; Sakaue, S.; Terao, C.; Luo, Y.; Sonehara, K.; Yamaguchi, K.; Amariuta, T.; Too, C.L.; Laufer, V.A.; Scott, I.C.; et al. Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nat. Genet.* **2022**, *54*, 1640–1651. [CrossRef] [PubMed]
42. Chen, L.; Deng, O.; Fang, T.; Chen, M.; Zhang, X.; Cong, R.; Lu, D.; Zhang, R.; Jin, Q.; Wang, X. Phenome-wide causal proteomics enhance systemic lupus erythematosus flare prediction: A study in Asian populations. *medRxiv* **2024**. [CrossRef]
43. Paek, S.J.; Lee, H.S.; Lee, Y.J.; Bang, S.Y.; Kim, D.; Kang, B.K.; Park, D.J.; Joo, Y.B.; Kim, M.; Kim, H.; et al. Tracking clonal dynamics of CD8 T cells and immune dysregulation in progression of systemic lupus erythematosus with nephritis. *Exp. Mol. Med.* **2025**, *57*, 1700–1710. [CrossRef]
44. Zhang, J.; Zhuang, W.; Li, Y.; Deng, C.; Xuan, J.; Sun, Y.; He, Y. Bioinformatic analysis and experimental verification reveal expansion of monocyte subsets with an interferon signature in systemic lupus erythematosus patients. *Arthritis Res. Ther.* **2025**, *27*, 96. [CrossRef] [PubMed]
45. Shen, M.; Duan, C.; Xie, C.; Wang, H.; Li, Z.; Li, B.; Wang, T. Identification of key interferon-stimulated genes for indicating the condition of patients with systemic lupus erythematosus. *Front. Immunol.* **2022**, *13*, 962393. [CrossRef]

46. Haslam, D.E.; Li, J.; Dillon, S.T.; Gu, X.; Cao, Y.; Zeleznik, O.A.; Sasamoto, N.; Zhang, X.; Eliassen, A.H.; Liang, L.; et al. Stability and reproducibility of proteomic profiles in epidemiological studies: Comparing the Olink and SOMAscan platforms. *Proteomics* **2022**, *22*, e2100170. [CrossRef]
47. Shang, S.; Xia, J.; He, G.; Zheng, Y.; Zhang, J.; Lu, H.; Wang, H.; Li, W.; Li, Q.; Chen, X. Advances in precision medicine for lupus nephritis: Biomarker- and AI-driven diagnosis and treatment response prediction and targeted therapies. *EBioMedicine* **2025**, *117*, 105785. [CrossRef]
48. Cervia-Hasler, C.; Bruningk, S.C.; Hoch, T.; Fan, B.; Muzio, G.; Thompson, R.C.; Ceglarek, L.; Meledin, R.; Westermann, P.; Emmenegger, M.; et al. Persistent complement dysregulation with signs of thromboinflammation in active Long Covid. *Science* **2024**, *383*, eadg7942. [CrossRef]
49. Galozzi, P.; Basso, D.; Plebani, M.; Padoan, A. Artificial intelligence and laboratory data in rheumatic diseases. *Clin. Chim. Acta* **2023**, *546*, 117388. [CrossRef]
50. Duvvuri, B.; Lood, C. Cell-Free DNA as a Biomarker in Autoimmune Rheumatic Diseases. *Front. Immunol.* **2019**, *10*, 502. [CrossRef]
51. Dihlmann, S.; Kaduk, C.; Passek, K.H.; Spieler, A.; Bockler, D.; Peters, A.S. Exploring circulating cell-free DNA as a biomarker and as an inducer of AIM2-inflammasome-mediated inflammation in patients with abdominal aortic aneurysm. *Sci. Rep.* **2025**, *15*, 20196. [CrossRef] [PubMed]
52. Jackson Chornenki, N.L.; Coke, R.; Kwong, A.C.; Dwivedi, D.J.; Xu, M.K.; McDonald, E.; Marshall, J.C.; Fox-Robichaud, A.E.; Charbonney, E.; Liaw, P.C. Comparison of the source and prognostic utility of cfDNA in trauma and sepsis. *Intensive Care Med. Exp.* **2019**, *7*, 29. [CrossRef]
53. Liu, F.; Su, Y.; Liu, X.; Zhao, L.; Wu, Z.; Liu, Y.; Zhang, L. Cell-free DNA: A metabolic byproduct with diagnostic and prognostic potential in rheumatic disorders. *Front. Pharmacol.* **2025**, *16*, 1537934. [CrossRef]
54. Lehmann, J.; Giaglis, S.; Kyburz, D.; Daoudlarian, D.; Walker, U.A. Plasma mtDNA as a possible contributor to and biomarker of inflammation in rheumatoid arthritis. *Arthritis Res. Ther.* **2024**, *26*, 97. [CrossRef]
55. Kerachian, M.A.; Azghandi, M.; Mozaffari-Jovin, S.; Thierry, A.R. Guidelines for pre-analytical conditions for assessing the methylation of circulating cell-free DNA. *Clin. Epigenetics* **2021**, *13*, 193. [CrossRef]
56. Peng, H.; Pan, M.; Zhou, Z.; Chen, C.; Xing, X.; Cheng, S.; Zhang, S.; Zheng, H.; Qian, K. The impact of preanalytical variables on the analysis of cell-free DNA from blood and urine samples. *Front. Cell Dev. Biol.* **2024**, *12*, 1385041. [CrossRef]
57. Sathyanarayana, S.H.; Spracklin, S.B.; Deharvengt, S.J.; Green, D.C.; Instasi, M.D.; Gallagher, T.L.; Shah, P.S.; Tsongalis, G.J. Standardized Workflow and Analytical Validation of Cell-Free DNA Extraction for Liquid Biopsy Using a Magnetic Bead-Based Cartridge System. *Cells* **2025**, *14*, 1062. [CrossRef] [PubMed]
58. Qi, T.; Pan, M.; Shi, H.; Wang, L.; Bai, Y.; Ge, Q. Cell-Free DNA Fragmentomics: The Novel Promising Biomarker. *Int. J. Mol. Sci.* **2023**, *24*, 1503. [CrossRef] [PubMed]
59. Fresneda Alarcon, M.; McLaren, Z.; Wright, H.L. Neutrophils in the Pathogenesis of Rheumatoid Arthritis and Systemic Lupus Erythematosus: Same Foe Different M.O. *Front. Immunol.* **2021**, *12*, 649693. [CrossRef]
60. Huang, Y.; Xue, Q.; Chang, J.; Wang, Y.; Cheng, C.; Xu, S.; Wang, X.; Miao, C. M6A methylation modification in autoimmune diseases, a promising treatment strategy based on epigenetics. *Arthritis Res. Ther.* **2023**, *25*, 189. [CrossRef]
61. Wu, J.; Deng, L.J.; Xia, Y.R.; Leng, R.X.; Fan, Y.G.; Pan, H.F.; Ye, D.Q. Involvement of N6-methyladenosine modifications of long noncoding RNAs in systemic lupus erythematosus. *Mol. Immunol.* **2022**, *143*, 77–84. [CrossRef]
62. Guo, D.; Liu, J.; Li, S.; Xu, P. Analysis of m6A regulators related immune characteristics in ankylosing spondylitis by integrated bioinformatics and computational strategies. *Sci. Rep.* **2024**, *14*, 2724. [CrossRef]
63. Cheng, L.; Li, H.; Zhan, H.; Liu, Y.; Li, X.; Huang, Y.; Wang, L.; Zhang, F.; Li, Y. Alterations of m6A RNA methylation regulators contribute to autophagy and immune infiltration in primary Sjogren's syndrome. *Front. Immunol.* **2022**, *13*, 949206. [CrossRef]
64. Gao, Y.; Zhang, Y.; Liu, X. Rheumatoid arthritis: Pathogenesis and therapeutic advances. *MedComm* **2024**, *5*, e509. [CrossRef] [PubMed]
65. Wardowska, A. m6A RNA Methylation in Systemic Autoimmune Diseases-A New Target for Epigenetic-Based Therapy? *Pharmaceuticals* **2021**, *14*, 218. [CrossRef] [PubMed]
66. Wang, H.; Mennea, P.D.; Chan, Y.K.E.; Cheng, Z.; Neofytou, M.C.; Surani, A.A.; Vijayaraghavan, A.; Ditter, E.J.; Bowers, R.; Eldridge, M.D.; et al. A standardized framework for robust fragmentomic feature extraction from cell-free DNA sequencing data. *Genome Biol.* **2025**, *26*, 141. [CrossRef]
67. Xu, P.; Cai, J.; Gao, Y.; Rong, Z. MIRACLE: Multi-task Learning based Interpretable Regulation of Autoimmune Diseases through Common Latent Epigenetics. *arXiv* **2023**. [CrossRef]
68. Kim, S.; Zhang, L.; Qin, Y.; Bohn, R.I.C.; Park, H.J. Pathway information on methylation analysis using deep neural network (PROMINENT): An interpretable deep learning method with pathway prior for phenotype prediction using gene-level DNA methylation. *Artif. Intell. Med.* **2025**, *170*, 103236. [CrossRef]

69. Lee, K.; Niku, S.; Koo, S.J.; Belezzuoli, E.; Guma, M. Molecular imaging for evaluation of synovitis associated with osteoarthritis: A narrative review. *Arthritis Res. Ther.* **2024**, *26*, 25. [CrossRef]
70. Boeren, A.M.P.; Oei, E.H.G.; van der Helm-van, A.H.M. The value of MRI for detecting subclinical joint inflammation in clinically suspect arthralgia. *RMD Open* **2022**, *8*, e002128. [CrossRef]
71. So, H.; Cheng, I.; Tam, L.S. The Role of Imaging in Predicting the Development of Rheumatoid Arthritis. *Rheumatol. Immunol. Res.* **2021**, *2*, 27–33. [CrossRef]
72. Ogdie, A.; Coates, L.C.; Mease, P. Measuring Outcomes in Psoriatic Arthritis. *Arthritis Care Res.* **2020**, *72* (Suppl. S10), 82–109. [CrossRef] [PubMed]
73. Frenken, M.; Schleich, C.; Brinks, R.; Abrar, D.B.; Goertz, C.; Schneider, M.; Ostendorf, B.; Sewerin, P. The value of the simplified RAMRIS-5 in early RA patients under methotrexate therapy using high-field MRI. *Arthritis Res. Ther.* **2019**, *21*, 21. [CrossRef]
74. Schleich, C.; Buchbender, C.; Sewerin, P.; Miese, F.; Aissa, J.; Brinks, R.; Schneider, M.; Antoch, G.; Ostendorf, B. Evaluation of a simplified version of the Rheumatoid Arthritis Magnetic Resonance Imaging Score (RAMRIS) comprising 5 joints (RAMRIS5). *Clin. Exp. Rheumatol.* **2015**, *33*, 209–215.
75. Ben-Eltriki, M.; Rafiq, A.; Paul, A.; Prabhu, D.; Afolabi, M.O.S.; Baslhaw, R.; Neilson, C.J.; Driedger, M.; Mahmud, S.M.; Lacaze-Masmonteil, T.; et al. Adaptive designs in clinical trials: A systematic review-part I. *BMC Med. Res. Methodol.* **2024**, *24*, 229. [CrossRef]
76. Kaizer, A.M.; Belli, H.M.; Ma, Z.; Nicklawsky, A.G.; Roberts, S.C.; Wild, J.; Wogu, A.F.; Xiao, M.; Sabo, R.T. Recent innovations in adaptive trial designs: A review of design opportunities in translational research. *J. Clin. Transl. Sci.* **2023**, *7*, e125. [CrossRef] [PubMed]
77. Noversa de Sousa, R.; Tascilar, K.; Corte, G.; Atzinger, A.; Minopoulou, I.; Ohrndorf, S.; Waldner, M.; Schmidkonz, C.; Kuwert, T.; Knieling, F.; et al. Metabolic and molecular imaging in inflammatory arthritis. *RMD Open* **2024**, *10*, e003880. [CrossRef]
78. MacKay, J.W.; Watkins, L.; Gold, G.; Kogan, F. [(18)F]NaF PET-MRI provides direct in-vivo evidence of the association between bone metabolic activity and adjacent synovitis in knee osteoarthritis: A cross-sectional study. *Osteoarthr. Cartil.* **2021**, *29*, 1155–1162. [CrossRef] [PubMed]
79. MacRitchie, N.; Frleta-Gilchrist, M.; Sugiyama, A.; Lawton, T.; McInnes, I.B.; Maffia, P. Molecular imaging of inflammation—Current and emerging technologies for diagnosis and treatment. *Pharmacol. Ther.* **2020**, *211*, 107550. [CrossRef]
80. Chandrupatla, D.; Molthoff, C.F.M.; Lammertsma, A.A.; van der Laken, C.J.; Jansen, G. The folate receptor beta as a macrophage-mediated imaging and therapeutic target in rheumatoid arthritis. *Drug Deliv. Transl. Res.* **2019**, *9*, 366–378. [CrossRef]
81. Mori, Y.; Novruzov, E.; Schmitt, D.; Cardinale, J.; Watabe, T.; Choyke, P.L.; Alavi, A.; Haberkorn, U.; Giesel, F.L. Clinical applications of fibroblast activation protein inhibitor positron emission tomography (FAPI-PET). *NPJ Imaging* **2024**, *2*, 48. [CrossRef]
82. Kastelik-Hryniewiecka, A.; Jewula, P.; Bakalorz, K.; Kramer-Marek, G.; Kuznik, N. Targeted PET/MRI Imaging Super Probes: A Critical Review of Opportunities and Challenges. *Int. J. Nanomed.* **2021**, *16*, 8465–8483. [CrossRef]
83. Xu, L.; Bressem, K.; Adams, L.; Poddubnyy, D.; Proft, F. AI for imaging evaluation in rheumatology: Applications of radiomics and computer vision-current status, future prospects and potential challenges. *Rheumatol. Adv. Pract.* **2025**, *9*, rkae147. [CrossRef]
84. Thesia, J.; Pandya, A. Wearable biosensors for autoimmune disorders. *Prog. Mol. Biol. Transl. Sci.* **2025**, *215*, 405–418. [CrossRef] [PubMed]
85. Hamy, V.; Llop, C.; Yee, C.W.; Garcia-Gancedo, L.; Maxwell, A.; Chen, W.H.; Tomlinson, R.; Bobbili, P.; Bendelac, J.; Landry, J.; et al. Patient-centric assessment of rheumatoid arthritis using a smartwatch and bespoke mobile app in a clinical setting. *Sci. Rep.* **2023**, *13*, 18311. [CrossRef]
86. Wagner, S.R.; Gregersen, R.R.; Henriksen, L.; Hauge, E.M.; Keller, K.K. Smartphone Pedometer Sensor Application for Evaluating Disease Activity and Predicting Comorbidities in Patients with Rheumatoid Arthritis: A Validation Study. *Sensors* **2022**, *22*, 9396. [CrossRef]
87. Reed, M.; Rampono, B.; Turner, W.; Harsanyi, A.; Lim, A.; Paramalingam, S.; Massasso, D.; Thakkar, V.; Mundae, M.; Rampono, E. A multicentre validation study of a smartphone application to screen hand arthritis. *BMC Musculoskelet. Disord.* **2022**, *23*, 433. [CrossRef] [PubMed]
88. Venerito, V.; Manigold, T.; Capodiferro, M.; Markham, D.; Blanchard, M.; Iannone, F.; Hugle, T. Single-camera motion capture of finger joint mobility as a digital biomarker for disease activity in rheumatoid arthritis. *Rheumatol. Adv. Pract.* **2025**, *9*, rkae143. [CrossRef]
89. Guo, L.; Chang, R.; Wang, J.; Narayanan, A.; Qian, P.; Leong, M.C.; Kundu, P.P.; Senthilkumar, S.; Garlapati, S.C.; Yong, E.C.K.; et al. Artificial intelligence-enhanced 3D gait analysis with a single consumer-grade camera. *J. Biomech.* **2025**, *187*, 112738. [CrossRef] [PubMed]
90. Hadjileontiadis, L.J.; Charisis, V.; Hadjidimitriou, S.; Dias, S.B.; Apostolidis, G.; Dimaridis, G.; Kitsas, I.; Karlas, A.; Fasoula, N.A.; Levi-Schaffer, F.; et al. European advances in digital rheumatology: Explainable insights and personalized digital health tools for psoriatic arthritis. *EClinicalMedicine* **2025**, *84*, 103243. [CrossRef]

91. Santosa, A.; Li, J.W.; Tan, T.C. Digital Health for Equitable Rheumatic Care: Integrating Real-World Experiences to Guide Policy Pathways. *Healthcare* **2025**, *13*, 438. [CrossRef]
92. Yang, Y.; Liu, Y.; Chen, Y.; Luo, D.; Xu, K.; Zhang, L. Artificial intelligence for predicting treatment responses in autoimmune rheumatic diseases: Advancements, challenges, and future perspectives. *Front. Immunol.* **2024**, *15*, 1477130. [CrossRef]
93. Zhao, J.; Li, L.; Li, J.; Zhang, L. Application of artificial intelligence in rheumatic disease: A bibliometric analysis. *Clin. Exp. Med.* **2024**, *24*, 196. [CrossRef]
94. Nelson, A.E.; Arbeeva, L. Narrative Review of Machine Learning in Rheumatic and Musculoskeletal Diseases for Clinicians and Researchers: Biases, Goals, and Future Directions. *J. Rheumatol.* **2022**, *49*, 1191–1200. [CrossRef] [PubMed]
95. McMaster, C.; Bird, A.; Liew, D.F.L.; Buchanan, R.R.; Owen, C.E.; Chapman, W.W.; Pires, D.E.V. Artificial Intelligence and Deep Learning for Rheumatologists. *Arthritis Rheumatol.* **2022**, *74*, 1893–1905. [CrossRef]
96. Hurez, V.; Gauderat, G.; Soret, P.; Myers, R.; Dasika, K.; Sheehan, R.; Friedrich, C.; Reed, M.; Laigle, L.; Riquelme, M.A.; et al. Virtual patients inspired by multiomics predict the efficacy of an anti-IFNalpha mAb in cutaneous lupus. *iScience* **2025**, *28*, 111754. [CrossRef] [PubMed]
97. Ebadi Jalal, M.; Emam, O.S.; Castillo-Olea, C.; Garcia-Zapirain, B.; Elmaghraby, A. Abnormality detection in nailfold capillary images using deep learning with EfficientNet and cascade transfer learning. *Sci. Rep.* **2025**, *15*, 2068. [CrossRef]
98. Lledo-Ibanez, G.M.; Saez Comet, L.; Freire Dapena, M.; Mesa Navas, M.; Martin Cascon, M.; Guillen Del Castillo, A.; Simeon, C.P.; Martinez Robles, E.; Todoli Parra, J.; Varela, D.C.; et al. CAPI-Detect: Machine learning in capillaroscopy reveals new variables influencing diagnosis. *Rheumatology* **2025**, *64*, 3667–3675. [CrossRef]
99. Knevel, R.; Liao, K.P. From real-world electronic health record data to real-world results using artificial intelligence. *Ann. Rheum. Dis.* **2023**, *82*, 306–311. [CrossRef] [PubMed]
100. Tonner, C.; Schmajuk, G.; Yazdany, J. A new era of quality measurement in rheumatology: Electronic clinical quality measures and national registries. *Curr. Opin. Rheumatol.* **2017**, *29*, 131–137. [CrossRef]
101. Francisco, M.; Johansson, T.; Kazi, S. Overview of the American College of Rheumatology's Electronic Health Record-Enabled Registry: The Rheumatology Informatics System for Effectiveness. *Clin. Exp. Rheumatol.* **2016**, *34*, S102–S104.
102. Oatis, C.A.; Konnyu, K.J.; Franklin, P.D. Generating consistent longitudinal real-world data to support research: Lessons from physical therapists. *ACR Open Rheumatol.* **2022**, *4*, 771–774. [CrossRef]
103. Izadi, Z.; Schmajuk, G.; Gianfrancesco, M.; Subash, M.; Evans, M.; Trupin, L.; Yazdany, J. Significant Gains in Rheumatoid Arthritis Quality Measures Among RISE Registry Practices. *Arthritis Care Res.* **2022**, *74*, 219–228. [CrossRef] [PubMed]
104. Tabatabaei Hosseini, S.A.; Kazemzadeh, R.; Foster, B.J.; Arpali, E.; Susal, C. New Tools for Data Harmonization and Their Potential Applications in Organ Transplantation. *Transplantation* **2024**, *108*, 2306–2317. [CrossRef]
105. Carbonaro, A.; Giorgetti, L.; Ridolfi, L.; Pasolini, R.; Pagliarani, A.; Cavallucci, M.; Andalo, A.; Gaudio, L.D.; De Angelis, P.; Vespignani, R.; et al. From raw data to research-ready: A FHIR-based transformation pipeline in a real-world oncology setting. *Comput. Biol. Med.* **2025**, *197*, 111051. [CrossRef]
106. Omar, M.; Naffaa, M.E.; Glicksberg, B.S.; Reuveni, H.; Nadkarni, G.N.; Klang, E. Advancing rheumatology with natural language processing: Insights and prospects from a systematic review. *Rheumatol. Adv. Pract.* **2024**, *8*, rkae120. [CrossRef] [PubMed]
107. Benavent, D.; Madrid-Garcia, A. Large language models and rheumatology: Are we there yet? *Rheumatol. Adv. Pract.* **2025**, *9*, rkae119. [CrossRef]
108. Humbert-Droz, M.; Izadi, Z.; Schmajuk, G.; Gianfrancesco, M.; Baker, M.C.; Yazdany, J.; Tamang, S. Development of a Natural Language Processing System for Extracting Rheumatoid Arthritis Outcomes from Clinical Notes Using the National Rheumatology Informatics System for Effectiveness Registry. *Arthritis Care Res.* **2023**, *75*, 608–615. [CrossRef] [PubMed]
109. Maghsoudi, A.; Sada, Y.H.; Nowakowski, S.; Guffey, D.; Zhu, H.; Yarlagadda, S.R.; Li, A.; Razjouyan, J. A Multi-Institutional Natural Language Processing Pipeline to Extract Performance Status from Electronic Health Records. *Cancer Control* **2024**, *31*, 10732748241279518. [CrossRef]
110. Zhang, F.; Kreuter, D.; Chen, Y.; Dittmer, S.; Tull, S.; Shadbahr, T.; Schut, M.; Asselbergs, F.; Kar, S.; Sivapalaratnam, S.; et al. Recent methodological advances in federated learning for healthcare. *Patterns* **2024**, *5*, 101006. [CrossRef]
111. Austin, J.A.; Lobo, E.H.; Samadbeik, M.; Engstrom, T.; Philip, R.; Pole, J.D.; Sullivan, C.M. Decades in the Making: The Evolution of Digital Health Research Infrastructure Through Synthetic Data, Common Data Models, and Federated Learning. *J. Med. Internet Res.* **2024**, *26*, e58637. [CrossRef]
112. Stoel, B. Use of artificial intelligence in imaging in rheumatology—Current status and future perspectives. *RMD Open* **2020**, *6*, e001063. [CrossRef]
113. Bird, A.; Oakden-Rayner, L.; McMaster, C.; Smith, L.A.; Zeng, M.; Wechalekar, M.D.; Ray, S.; Proudman, S.; Palmer, L.J. Artificial intelligence and the future of radiographic scoring in rheumatoid arthritis: A viewpoint. *Arthritis Res. Ther.* **2022**, *24*, 268. [CrossRef]

114. Zhang, N.; Yang, S.; Zwagemaker, A.F.; Huo, A.; Li, Y.J.; Zhou, F.; Hilliard, P.; Squire, S.; Bouskill, V.; Mohanta, A.; et al. A semiquantitative color Doppler ultrasound scoring system for evaluation of synovitis in joints of patients with blood-induced arthropathy. *Insights Imaging* **2021**, *12*, 132. [CrossRef]
115. Zwanenburg, A.; Vallieres, M.; Abdalah, M.A.; Aerts, H.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [CrossRef]
116. Oettl, F.C.; Zsidai, B.; Oeding, J.F.; Hirschmann, M.T.; Feldt, R.; Fendrich, D.; Kraeutler, M.J.; Winkler, P.W.; Szaro, P.; Samuelsson, K.; et al. Artificial intelligence-assisted analysis of musculoskeletal imaging-A narrative review of the current state of machine learning models. *Knee Surg. Sports Traumatol. Arthrosc.* **2025**, *33*, 3032–3038. [CrossRef]
117. Salvi, M.; Seoni, S.; Campagner, A.; Gertych, A.; Acharya, U.R.; Molinari, F.; Cabitza, F. Explainability and uncertainty: Two sides of the same coin for enhancing the interpretability of deep learning models in healthcare. *Int. J. Med. Inform.* **2025**, *197*, 105846. [CrossRef] [PubMed]
118. Chen, M.; Wang, Y.; Wang, Q.; Shi, J.; Wang, H.; Ye, Z.; Xue, P.; Qiao, Y. Impact of human and artificial intelligence collaboration on workload reduction in medical image interpretation. *NPJ Digit. Med.* **2024**, *7*, 349. [CrossRef] [PubMed]
119. Guermazi, A.; Roemer, F.W.; Crema, M.D.; Jarraya, M.; Mobasheri, A.; Hayashi, D. Strategic application of imaging in DMOAD clinical trials: Focus on eligibility, drug delivery, and semiquantitative assessment of structural progression. *Ther. Adv. Musculoskelet. Dis.* **2023**, *15*, 1759720X231165558. [CrossRef] [PubMed]
120. Verhoeven, M.M.A.; Westgeest, A.A.A.; Schwarting, A.; Jacobs, J.W.G.; Heller, C.; van Laar, J.M.; Lafeber, F.; Tekstra, J.; Triantafyllias, K.; Welsing, P.M.J. Development and Validation of Rheumatoid Arthritis Disease Activity Indices Including HandScan (Optical Spectral Transmission) Scores. *Arthritis Care Res.* **2022**, *74*, 1493–1499. [CrossRef] [PubMed]
121. Park, D.J. Importance of Time-Integrated Cumulative Parameters for Radiographic Progression Prediction of Rheumatoid Arthritis. *J. Rheum. Dis.* **2022**, *29*, 129–131. [CrossRef] [PubMed]
122. Gumber, L.; Rayner, F.; Bigirumurame, T.; Dyke, B.; Melville, A.; Kerrigan, S.; McGucken, A.; Naamane, N.; Prichard, J.; Buckley, C.D.; et al. Patient-reported outcomes as early warning signs of flare following drug cessation in rheumatoid arthritis. *RMD Open* **2025**, *11*, e005442. [CrossRef]
123. Gandrup, J.; Selby, D.A.; van der Veer, S.N.; McBeth, J.; Dixon, W.G. Using patient-reported data from a smartphone app to capture and characterize real-time patient-reported flares in rheumatoid arthritis. *Rheumatol. Adv. Pract.* **2022**, *6*, rkac021. [CrossRef]
124. Momtazmanesh, S.; Nowroozi, A.; Rezaei, N. Artificial Intelligence in Rheumatoid Arthritis: Current Status and Future Perspectives: A State-of-the-Art Review. *Rheumatol. Ther.* **2022**, *9*, 1249–1304. [CrossRef]
125. Gul, H.; Di Matteo, A.; Anioke, I.; Shuweidhi, F.; Mankia, K.; Ponchel, F.; Emery, P. Predicting Flare in Patients with Rheumatoid Arthritis in Biologic Induced Remission, on Tapering, and on Stable Therapy. *ACR Open Rheumatol.* **2024**, *6*, 294–303. [CrossRef]
126. Patharkar, A.; Cai, F.; Al-Hindawi, F.; Wu, T. Predictive modeling of biomedical temporal data in healthcare applications: Review and future directions. *Front. Physiol.* **2024**, *15*, 1386760. [CrossRef]
127. Richardson, S.; Lawrence, K.; Schoenthaler, A.M.; Mann, D. A framework for digital health equity. *NPJ Digit. Med.* **2022**, *5*, 119. [CrossRef] [PubMed]
128. Davis, S.E.; Dorn, C.; Park, D.J.; Matheny, M.E. Emerging algorithmic bias: Fairness drift as the next dimension of model maintenance and sustainability. *J. Am. Med. Inform. Assoc.* **2025**, *32*, 845–854. [CrossRef]
129. Temmoku, J.; Migita, K.; Yoshida, S.; Matsumoto, H.; Fujita, Y.; Matsuoka, N.; Yashiro-Furuya, M.; Asano, T.; Sato, S.; Suzuki, E.; et al. Real-world comparative effectiveness of bDMARDs and JAK inhibitors in elderly patients with rheumatoid arthritis. *Medicine* **2022**, *101*, e31161. [CrossRef]
130. Eberhard, A.; Di Giuseppe, D.; Askling, J.; Bergman, S.; Bower, H.; Chatzidionysiou, K.; Forsblad-d'Elia, H.; Kastbom, A.; Olofsson, T.; Frisell, T.; et al. Effectiveness of JAK Inhibitors Compared with Biologic Disease-Modifying Antirheumatic Drugs on Pain Reduction in Rheumatoid Arthritis: Results from a Nationwide Swedish Cohort Study. *Arthritis Rheumatol.* **2025**, *77*, 253–262. [CrossRef] [PubMed]
131. Efthimiou, O.; Seo, M.; Chalkou, K.; Debray, T.; Egger, M.; Salanti, G. Developing clinical prediction models: A step-by-step guide. *BMJ* **2024**, *386*, e078276. [CrossRef]
132. Liu, D.; Yu, G.; Yuan, N.; Nie, D. The efficacy and safety of biologic or targeted synthetic DMARDs in rheumatoid arthritis treatment: One year of review 2024. *Allergol. Immunopathol.* **2025**, *53*, 140–162. [CrossRef]
133. Favalli, E.G.; Maioli, G.; Caporali, R. Biologics or Janus Kinase Inhibitors in Rheumatoid Arthritis Patients Who are Insufficient Responders to Conventional Anti-Rheumatic Drugs. *Drugs* **2024**, *84*, 877–894. [CrossRef]
134. Wang, S.S.; Lewis, M.J.; Pitzalis, C. DNA Methylation Signatures of Response to Conventional Synthetic and Biologic Disease-Modifying Antirheumatic Drugs (DMARDs) in Rheumatoid Arthritis. *Biomedicines* **2023**, *11*, 1987. [CrossRef]
135. Bhasin, S.; Cheung, P.P. The Role of Power Doppler Ultrasonography as Disease Activity Marker in Rheumatoid Arthritis. *Dis. Markers* **2015**, *2015*, 325909. [CrossRef] [PubMed]

136. Subash, M.; Liu, L.H.; DeQuattro, K.; Choden, S.; Jacobsohn, L.; Katz, P.; Bajaj, P.; Barton, J.L.; Bartels, C.; Bermas, B.; et al. The Development of the Rheumatology Informatics System for Effectiveness Learning Collaborative for Improving Patient-Reported Outcome Collection and Patient-Centered Communication in Adult Rheumatology. *ACR Open Rheumatol.* **2021**, *3*, 690–698. [CrossRef] [PubMed]
137. Abernethy, A.; Adams, L.; Barrett, M.; Bechtel, C.; Brennan, P.; Butte, A.; Faulkner, J.; Fontaine, E.; Friedhoff, S.; Halamka, J.; et al. The Promise of Digital Health: Then, Now, and the Future. *NAM Perspect.* **2022**. [CrossRef]
138. 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]
139. Chen, R.J.; Wang, J.J.; Williamson, D.F.K.; Chen, T.Y.; Lipkova, J.; Lu, M.Y.; Sahai, S.; Mahmood, F. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **2023**, *7*, 719–742. [CrossRef] [PubMed]
140. Kore, A.; Abbasi Bavil, E.; Subasri, V.; Abdalla, M.; Fine, B.; Dolatabadi, E.; Abdalla, M. Empirical data drift detection experiments on real-world medical imaging data. *Nat. Commun.* **2024**, *15*, 1887. [CrossRef]
141. Liu, L.H.; Choden, S.; Yazdany, J. Quality improvement initiatives in rheumatology: An integrative review of the last 5 years. *Curr. Opin. Rheumatol.* **2019**, *31*, 98–108. [CrossRef]
142. Liu, T.; Gu, Y.; Chen, H.; Zhang, Y.; Zheng, L.; Huang, X.; Xu, Y.; Wen, C.; Chen, M.; Lin, J.; et al. A foundational triage system for improving accuracy in moderate acuity level emergency classifications. *Commun. Med.* **2025**, *5*, 322. [CrossRef]
143. Portela, A.; Banga, J.R.; Matabuena, M. Conformal prediction for uncertainty quantification in dynamic biological systems. *PLoS Comput. Biol.* **2025**, *21*, e1013098. [CrossRef]
144. Chen, D.; He, E.; Pace, K.; Chekay, M.; Raman, S. Concordance with SPIRIT-AI guidelines in reporting of randomized controlled trial protocols investigating artificial intelligence in oncology: A systematic review. *Oncologist* **2025**, *30*, oyaf112. [CrossRef]
145. Bodnari, A.; Travis, J. Scaling enterprise AI in healthcare: The role of governance in risk mitigation frameworks. *NPJ Digit. Med.* **2025**, *8*, 272. [CrossRef]
146. Palaniappan, K.; Lin, E.Y.T.; Vogel, S. Global Regulatory Frameworks for the Use of Artificial Intelligence (AI) in the Healthcare Services Sector. *Healthcare* **2024**, *12*, 562. [CrossRef]
147. Martin, A.R.; Kanai, M.; Kamatani, Y.; Okada, Y.; Neale, B.M.; Daly, M.J. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **2019**, *51*, 584–591. [CrossRef] [PubMed]
148. Jain, A.; Brooks, J.R.; Alford, C.C.; Chang, C.S.; Mueller, N.M.; Umscheid, C.A.; Bierman, A.S. Awareness of Racial and Ethnic Bias and Potential Solutions to Address Bias with Use of Health Care Algorithms. *JAMA Health Forum* **2023**, *4*, e231197. [CrossRef] [PubMed]
149. Norori, N.; Hu, Q.; Aellen, F.M.; Faraci, F.D.; Tzovara, A. Addressing bias in big data and AI for health care: A call for open science. *Patterns* **2021**, *2*, 100347. [CrossRef] [PubMed]
150. Cope, A.P.; Jasenecova, M.; Vasconcelos, J.C.; Filer, A.; Raza, K.; Qureshi, S.; D'Agostino, M.A.; McInnes, I.B.; Isaacs, J.D.; Pratt, A.G.; et al. Abatacept in individuals at high risk of rheumatoid arthritis (APIPPRA): A randomised, double-blind, multicentre, parallel, placebo-controlled, phase 2b clinical trial. *Lancet* **2024**, *403*, 838–849. [CrossRef] [PubMed]
151. Rech, J.; Tascilar, K.; Hagen, M.; Kleyer, A.; Manger, B.; Schoenau, V.; Hueber, A.J.; Kleinert, S.; Baraliakos, X.; Braun, J.; et al. Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk (ARIAA): A randomised, international, multicentre, double-blind, placebo-controlled trial. *Lancet* **2024**, *403*, 850–859. [CrossRef]
152. Jin, S.; Zhao, J.; Li, M.; Zeng, X. New insights into the pathogenesis and management of rheumatoid arthritis. *Chronic Dis. Transl. Med.* **2022**, *8*, 256–263. [CrossRef]
153. McDonald, S.M.; Felfeliyan, B.; Hassan, A.; Kupper, J.C.; El-Hajj, R.; Wichuk, S.; Aneja, A.; Kwok, C.; Zhang, C.X.Y.; Jans, L.; et al. Evaluating potential for AI automation of quantitative and semi-quantitative MRI scoring in arthritis, especially at the knee: A systematic literature review. *Skeletal Radiol.* **2025**, *54*, 2339–2349. [CrossRef]
154. Mao, Y.; Imahori, K.; Fang, W.; Sugimori, H.; Kiuch, S.; Sutherland, K.; Kamishima, T. Artificial Intelligence Quantification of Enhanced Synovium Throughout the Entire Hand in Rheumatoid Arthritis on Dynamic Contrast-Enhanced MRI. *J. Magn. Reson. Imaging* **2025**, *61*, 771–783. [CrossRef] [PubMed]
155. Nicoara, A.I.; Sas, L.M.; Bita, C.E.; Dinescu, S.C.; Vreju, F.A. Implementation of artificial intelligence models in magnetic resonance imaging with focus on diagnosis of rheumatoid arthritis and axial spondyloarthritis: Narrative review. *Front. Med.* **2023**, *10*, 1280266. [CrossRef]
156. Schlereth, M.; Mutlu, M.Y.; Utz, J.; Bayat, S.; Heimann, T.; Qiu, J.; Ehring, C.; Liu, C.; Uder, M.; Kleyer, A.; et al. Deep learning-based classification of erosion, synovitis and osteitis in hand MRI of patients with inflammatory arthritis. *RMD Open* **2024**, *10*, e004273. [CrossRef] [PubMed]
157. Kumar, R.; Sporn, K.; Prabhakar, V.; Alnemri, A.; Khanna, A.; Paladugu, P.; Gowda, C.; Clarkson, L.; Zaman, N.; Tavakkoli, A. Computational and Imaging Approaches for Precision Characterization of Bone, Cartilage, and Synovial Biomolecules. *J. Pers. Med.* **2025**, *15*, 298. [CrossRef] [PubMed]

158. Currie, G.; Rohren, E. The deep radiomic analytics pipeline. *Vet. Radiol. Ultrasound* **2022**, *63* (Suppl. S1), 889–896. [CrossRef]
159. Ou, Y.; Ambalathankandy, P.; Furuya, R.; Kawada, S.; Zeng, T.; An, Y.; Kamishima, T.; Tamura, K.; Ikebe, M. A Sub-Pixel Accurate Quantification of Joint Space Narrowing Progression in Rheumatoid Arthritis. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 53–64. [CrossRef]
160. Ichikawa, S.; Kamishima, T.; Sutherland, K.; Okubo, T.; Katayama, K. Radiographic quantifications of joint space narrowing progression by computer-based approach using temporal subtraction in rheumatoid wrist. *Br. J. Radiol.* **2016**, *89*, 20150403. [CrossRef]
161. Ou, J.; Zhang, J.; Alswadeh, M.; Zhu, Z.; Tang, J.; Sang, H.; Lu, K. Advancing osteoarthritis research: The role of AI in clinical, imaging and omics fields. *Bone Res.* **2025**, *13*, 48. [CrossRef]
162. Woelfle, T.; Bourguignon, L.; Lorscheider, J.; Kappos, L.; Naegelin, Y.; Jutzeler, C.R. Wearable Sensor Technologies to Assess Motor Functions in People with Multiple Sclerosis: Systematic Scoping Review and Perspective. *J. Med. Internet Res.* **2023**, *25*, e44428. [CrossRef]
163. Lee, S.; Kang, S.; Eun, Y.; Won, H.H.; Kim, H.; Lee, J.; Koh, E.M.; Cha, H.S. Machine learning-based prediction model for responses of bDMARDs in patients with rheumatoid arthritis and ankylosing spondylitis. *Arthritis Res. Ther.* **2021**, *23*, 254. [CrossRef]
164. Tao, W.; Concepcion, A.N.; Vianen, M.; Marijnissen, A.C.A.; Lafeber, F.; Radstake, T.; Pandit, A. Multiomics and Machine Learning Accurately Predict Clinical Response to Adalimumab and Etanercept Therapy in Patients with Rheumatoid Arthritis. *Arthritis Rheumatol.* **2021**, *73*, 212–222. [CrossRef]
165. Benavent, D.; Carmona, L.; Garcia Llorente, J.F.; Montoro, M.; Ramirez, S.; Oton, T.; Loza, E.; Gomez-Centeno, A. Artificial intelligence to predict treatment response in rheumatoid arthritis and spondyloarthritis: A scoping review. *Rheumatol. Int.* **2025**, *45*, 91. [CrossRef]
166. Hameed, M.; Exarchou, S.; Eberhard, A.; Sharma, A.; Bergstrom, U.; Cagnotto, G.; Einarsson, J.T.; Turesson, C. Predictors at diagnosis for start of biologic disease-modifying antirheumatic drugs in patients with early rheumatoid arthritis: A cohort study. *BMJ Open* **2024**, *14*, e076131. [CrossRef]
167. Postal, M.; Vivaldo, J.F.; Fernandez-Ruiz, R.; Paredes, J.L.; Appenzeller, S.; Niewold, T.B. Type I interferon in the pathogenesis of systemic lupus erythematosus. *Curr. Opin. Immunol.* **2020**, *67*, 87–94. [CrossRef]
168. Miyachi, K.; Iwamoto, T.; Kojima, S.; Ida, T.; Suzuki, J.; Yamamoto, T.; Mimura, N.; Sugiyama, T.; Tanaka, S.; Furuta, S.; et al. Relationship of systemic type I interferon activity with clinical phenotypes, disease activity, and damage accrual in systemic lupus erythematosus in treatment-naive patients: A retrospective longitudinal analysis. *Arthritis Res. Ther.* **2023**, *25*, 26. [CrossRef] [PubMed]
169. Kaan, E.D.; Brunekreef, T.E.; Drylewicz, J.; van den Hoogen, L.L.; van der Linden, M.; Leavis, H.L.; van Laar, J.M.; van der Vlist, M.; Otten, H.G.; Limper, M. Association of autoantibodies with the IFN signature and NETosis in patients with systemic lupus erythematosus. *J. Transl. Autoimmun.* **2024**, *9*, 100246. [CrossRef] [PubMed]
170. Felten, R.; Scher, F.; Sagez, F.; Chasset, F.; Arnaud, L. Spotlight on anifrolumab and its potential for the treatment of moderate-to-severe systemic lupus erythematosus: Evidence to date. *Drug Des. Devel Ther.* **2019**, *13*, 1535–1543. [CrossRef] [PubMed]
171. Furie, R.; Khamashta, M.; Merrill, J.T.; Werth, V.P.; Kalunian, K.; Brohawn, P.; Illei, G.G.; Drappa, J.; Wang, L.; Yoo, S.; et al. Anifrolumab, an Anti-Interferon-alpha Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. *Arthritis Rheumatol.* **2017**, *69*, 376–386. [CrossRef]
172. Baker, T.; Sharifian, H.; Newcombe, P.J.; Gavin, P.G.; Lazarus, M.N.; Ramaswamy, M.; White, W.I.; Ferrari, N.; Muthas, D.; Tummala, R.; et al. Type I interferon blockade with anifrolumab in patients with systemic lupus erythematosus modulates key immunopathological pathways in a gene expression and proteomic analysis of two phase 3 trials. *Ann. Rheum. Dis.* **2024**, *83*, 1018–1027. [CrossRef]
173. Cleanthous, S.; Strzok, S.; Haier, B.; Cano, S.; Morel, T. The Patient Experience of Fatigue in Systemic Lupus Erythematosus: A Conceptual Model. *Rheumatol. Ther.* **2022**, *9*, 95–108. [CrossRef] [PubMed]
174. Mai, L.; Asaduzzaman, A.; Noamani, B.; Fortin, P.R.; Gladman, D.D.; Touma, Z.; Urowitz, M.B.; Wither, J. The baseline interferon signature predicts disease severity over the subsequent 5 years in systemic lupus erythematosus. *Arthritis Res. Ther.* **2021**, *23*, 29. [CrossRef]
175. Ruscitti, P.; Allanore, Y.; Baldini, C.; Barilaro, G.; Bartoloni Bocci, E.; Bearzi, P.; Bellis, E.; Berardicurti, O.; Biaggi, A.; Bombardieri, M.; et al. Tailoring the treatment of inflammatory rheumatic diseases by a better stratification and characterization of the clinical patient heterogeneity. Findings from a systematic literature review and experts' consensus. *Autoimmun. Rev.* **2024**, *23*, 103581. [CrossRef]
176. Oliveira, J.J.; Karrar, S.; Rainbow, D.B.; Pinder, C.L.; Clarke, P.; Rubio Garcia, A.; Al-Assar, O.; Burling, K.; Morris, S.; Stratton, R.; et al. The plasma biomarker soluble SIGLEC-1 is associated with the type I interferon transcriptional signature, ethnic background and renal disease in systemic lupus erythematosus. *Arthritis Res. Ther.* **2018**, *20*, 152. [CrossRef] [PubMed]
177. Perng, Y.C.; Lenschow, D.J. ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol.* **2018**, *16*, 423–439. [CrossRef]
178. Liu, W.; Zhang, S.; Wang, J. IFN-gamma, should not be ignored in SLE. *Front. Immunol.* **2022**, *13*, 954706. [CrossRef]

179. Gomez-Banuelos, E.; Goldman, D.W.; Andrade, V.; Darrah, E.; Petri, M.; Andrade, F. Uncoupling interferons and the interferon signature explains clinical and transcriptional subsets in SLE. *Cell Rep. Med.* **2024**, *5*, 101569. [CrossRef]
180. Jupe, E.R.; Lushington, G.H.; Purushothaman, M.; Pautasso, F.; Armstrong, G.; Sorathia, A.; Crawley, J.; Nadipelli, V.R.; Rubin, B.; Newhardt, R.; et al. Tracking of Systemic Lupus Erythematosus (SLE) Longitudinally Using Biosensor and Patient-Reported Data: A Report on the Fully Decentralized Mobile Study to Measure and Predict Lupus Disease Activity Using Digital Signals-The OASIS Study. *BioTech* **2023**, *12*, 62. [CrossRef]
181. Li, Y.; Yao, L.; Lee, Y.A.; Huang, Y.; Merkel, P.A.; Vina, E.; Yeh, Y.Y.; Li, Y.; Allen, J.M.; Bian, J.; et al. A fair machine learning model to predict flares of systemic lupus erythematosus. *JAMIA Open* **2025**, *8*, ooaf072. [CrossRef] [PubMed]
182. Huang, S.; Chen, Y.; Song, Y.; Wu, K.; Chen, T.; Zhang, Y.; Jia, W.; Zhang, H.T.; Liang, D.D.; Yang, J.; et al. Deep learning model to predict lupus nephritis renal flare based on dynamic multivariable time-series data. *BMJ Open* **2024**, *14*, e071821. [CrossRef] [PubMed]
183. Kamudoni, P.; Lyden, K.; Gunther, O.; Jaitely, V.; Araujo, T.D.; Spies, E.; Park, J.; Thomas, E.; Buie, J.; Blankenship, J.M.; et al. Identifying meaningful aspects of health and concepts of interest for assessment in systemic lupus erythematosus: Implications for digital clinical measure development. *J. Patient Rep. Outcomes* **2024**, *8*, 154. [CrossRef]
184. Brzezinska, O.E.; Rychlicki-Kicior, K.A.; Makowska, J.S. Automatic assessment of nailfold capillaroscopy software: A pilot study. *Reumatologia* **2024**, *62*, 346–350. [CrossRef]
185. Emam, O.S.; Ebadi Jalal, M.; Garcia-Zapirain, B.; Elmaghraby, A.S. Artificial Intelligence Algorithms in Nailfold Capillaroscopy Image Analysis: A Systematic Review. *medRxiv* **2024**. [CrossRef]
186. Adams, L.C.; Bressem, K.K.; Poddubnyy, D. Artificial intelligence and machine learning in axial spondyloarthritis. *Curr. Opin. Rheumatol.* **2024**, *36*, 267–273. [CrossRef]
187. Pons, M.; Georgiadis, S.; Hetland, M.L.; Ahmadzay, Z.F.; Rasmussen, S.; Christiansen, S.N.; Di Giuseppe, D.; Wallman, J.K.; Pavelka, K.; Zavada, J.; et al. Predictors of Secukinumab Treatment Response and Continuation in Axial Spondyloarthritis: Results from the EuroSpA Research Collaboration Network. *J. Rheumatol.* **2025**, *52*, 572–582. [CrossRef]
188. Dalix, E.; Marcelli, C.; Bejan-Angoulvant, T.; Finckh, A.; Rancon, F.; Akrour, M.; De Araujo, L.; Presles, E.; Marotte, H.; ROC-SpA study group. Rotation or change of biotherapy after TNF blocker treatment failure for axial spondyloarthritis: The ROC-SpA study, a randomised controlled study protocol. *BMJ Open* **2024**, *14*, e087872. [CrossRef]
189. Cozzi, G.; Scagnellato, L.; Lorenzin, M.; Collesei, A.; Oliviero, F.; Damasco, A.; Cosma, C.; Basso, D.; Doria, A.; Ramonda, R. Predictors of response to bDMARDs and tsDMARDs in psoriatic arthritis: A pilot study on the role of musculoskeletal ultrasound. *Front. Med.* **2024**, *11*, 1482894. [CrossRef]
190. Tillett, W.; Ogdie, A.; Passey, A.; Gorecki, P. Impact of psoriatic arthritis and comorbidities on ustekinumab outcomes in psoriasis: A retrospective, observational BADBIR cohort study. *RMD Open* **2023**, *9*, e002533. [CrossRef] [PubMed]
191. Kunzler, T.; Bamert, M.; Sprott, H. Factors predicting treatment response to biological and targeted synthetic disease-modifying antirheumatic drugs in psoriatic arthritis—A systematic review and meta-analysis. *Clin. Rheumatol.* **2024**, *43*, 3723–3746. [CrossRef]
192. Guo, H.; Gao, J.; Gong, L.; Wang, Y. Multi-omics analysis reveals novel causal pathways in psoriasis pathogenesis. *J. Transl. Med.* **2025**, *23*, 100. [CrossRef]
193. Shi, Z.; Ding, Y.; Dong, X.; Li, G.; Li, B.; Hou, J.; Xue, L. The diagnostic value and clinical relevance of salivary gland ultrasound in patients with highly suspected Sjogren's Disease: A prospective monocentric study. *Arthritis Res. Ther.* **2025**, *27*, 175. [CrossRef]
194. Yang, J.; Park, Y.; Lee, J.J.; Kim, W.U.; Park, S.H.; Kwok, S.K. Clinical value of salivary gland ultrasonography in evaluating secretory function, disease activity, and lymphoma risk factors in primary Sjogren's syndrome. *Clin. Rheumatol.* **2025**, *44*, 1643–1652. [CrossRef] [PubMed]
195. De Vita, S.; Isola, M.; Baldini, C.; Goules, A.V.; Chatzis, L.G.; Quartuccio, L.; Zabotti, A.; Giovannini, I.; Donati, V.; Ferro, F.; et al. Predicting lymphoma in Sjogren's syndrome and the pathogenetic role of parotid microenvironment through precise parotid swelling recording. *Rheumatology* **2023**, *62*, 1586–1593. [CrossRef] [PubMed]
196. Umapathy, V.R.; Natarajan, P.M.; Swamikannu, B. Review Insights on Salivary Proteomics Biomarkers in Oral Cancer Detection and Diagnosis. *Molecules* **2023**, *28*, 5283. [CrossRef] [PubMed]
197. Sembler-Moller, M.L.; Belstrom, D.; Locht, H.; Pedersen, A.M.L. Proteomics of saliva, plasma, and salivary gland tissue in Sjogren's syndrome and non-Sjogren patients identify novel biomarker candidates. *J. Proteomics* **2020**, *225*, 103877. [CrossRef]
198. Hu, S.; Gao, K.; Pollard, R.; Arellano-Garcia, M.; Zhou, H.; Zhang, L.; Elashoff, D.; Kallenberg, C.G.; Vissink, A.; Wong, D.T. Preclinical validation of salivary biomarkers for primary Sjogren's syndrome. *Arthritis Care Res.* **2010**, *62*, 1633–1638. [CrossRef]
199. Bonroy, C.; Piette, Y.; Allenbach, Y.; Bossuyt, X.; Damoiseaux, J. Positioning of myositis-specific and associated autoantibody (MSA/MAA) testing in disease criteria and routine diagnostic work-up. *J. Transl. Autoimmun.* **2022**, *5*, 100148. [CrossRef]
200. McLeish, E.; Slater, N.; Mastaglia, F.L.; Needham, M.; Coudert, J.D. From data to diagnosis: How machine learning is revolutionizing biomarker discovery in idiopathic inflammatory myopathies. *Brief. Bioinform.* **2023**, *25*, bbad514. [CrossRef]

201. Wang, H.; Chen, X.; Du, Y.; Wang, L.; Wang, Q.; Wu, H.; Liu, L.; Xue, J. Mortality risk in patients with anti-MDA5 dermatomyositis is related to rapidly progressive interstitial lung disease and anti-Ro52 antibody. *Arthritis Res. Ther.* **2023**, *25*, 127. [CrossRef] [PubMed]
202. Nagawa, K.; Suzuki, M.; Yamamoto, Y.; Inoue, K.; Kozawa, E.; Mimura, T.; Nakamura, K.; Nagata, M.; Niitsu, M. Texture analysis of muscle MRI: Machine learning-based classifications in idiopathic inflammatory myopathies. *Sci. Rep.* **2021**, *11*, 9821. [CrossRef]
203. Wang, F.; Zhou, S.; Hou, B.; Santini, F.; Yuan, L.; Guo, Y.; Zhu, J.; Hilbert, T.; Kober, T.; Zhang, Y.; et al. Assessment of idiopathic inflammatory myopathy using a deep learning method for muscle T2 mapping segmentation. *Eur. Radiol.* **2023**, *33*, 2350–2357. [CrossRef]
204. Danieli, M.G.; Brunetto, S.; Gammeri, L.; Palmeri, D.; Claudi, I.; Shoenfeld, Y.; Gangemi, S. Machine learning application in autoimmune diseases: State of art and future prospectives. *Autoimmun. Rev.* **2024**, *23*, 103496. [CrossRef]
205. Moingeon, P. Artificial intelligence-driven drug development against autoimmune diseases. *Trends Pharmacol. Sci.* **2023**, *44*, 411–424. [CrossRef]
206. Wu, H.; Li, X.; Xu, H.; Li, Z.; Feng, F.; Zhang, J.; Xu, Z.; Ni, H.; Guo, Y.; Li, Y. Malignancy in Idiopathic Inflammatory Myopathies: Recent Insights. *Clin. Rev. Allergy Immunol.* **2025**, *68*, 83. [CrossRef]
207. van der Geest, K.S.M.; Treglia, G.; Glaudemans, A.; Brouwer, E.; Sandovici, M.; Jamar, F.; Gheysens, O.; Slart, R. Diagnostic value of [18F]FDG-PET/CT for treatment monitoring in large vessel vasculitis: A systematic review and meta-analysis. *Eur. J. Nucl. Med. Mol. Imaging* **2021**, *48*, 3886–3902. [CrossRef]
208. Wilk, B.; Wisenberg, G.; Dharmakumar, R.; Thiessen, J.D.; Goldhawk, D.E.; Prato, F.S. Hybrid PET/MR imaging in myocardial inflammation post-myocardial infarction. *J. Nucl. Cardiol.* **2020**, *27*, 2083–2099. [CrossRef]
209. Brilland, B.; Riou, J.; Quemeneur, T.; Vandenbussche, C.; Merillon, N.; Boizard-Moracchini, A.; Roy, M.; Despre, M.; Piccoli, G.B.; Djema, A.; et al. Identification of Renal Transcripts Associated with Kidney Function and Prognosis in ANCA-Associated Vasculitis. *J. Am. Soc. Nephrol.* **2025**. [CrossRef] [PubMed]
210. Jia, M.; Han, S.; Li, L.; Fu, Y.; Zhou, D. Interferon-Stimulated Genes: Novel Targets in Renal Pathogenesis. *Kidney Dis.* **2025**, *11*, 390–401. [CrossRef] [PubMed]
211. Omar, M.; Agbareia, R.; Naffaa, M.E.; Watad, A.; Glicksberg, B.S.; Nadkarni, G.N.; Klang, E. Applications of Artificial Intelligence in Vasculitides: A Systematic Review. *ACR Open Rheumatol.* **2025**, *7*, e70016. [CrossRef]
212. Maarseveen, T.D.; Glas, H.K.; Veris-van Dieren, J.; van den Akker, E.; Knevel, R. Improving musculoskeletal care with AI enhanced triage through data driven screening of referral letters. *NPJ Digit. Med.* **2025**, *8*, 98. [CrossRef]
213. Knitza, J.; Janousek, L.; Kluge, F.; von der Decken, C.B.; Kleinert, S.; Vorbruggen, W.; Kleyer, A.; Simon, D.; Hueber, A.J.; Muehlensiepen, F.; et al. Machine learning-based improvement of an online rheumatology referral and triage system. *Front. Med.* **2022**, *9*, 954056. [CrossRef] [PubMed]
214. Guo, L.; Wang, J.; Li, J.; Yao, J.; Zhao, H. Biomarkers of rheumatoid arthritis-associated interstitial lung disease: A systematic review and meta-analysis. *Front. Immunol.* **2024**, *15*, 1455346. [CrossRef]
215. Frederiksen, B.A.; Hammer, H.B.; Terslev, L.; Ammitzboll-Danielsen, M.; Savarimuthu, T.R.; Weber, A.B.H.; Just, S.A. Automated ultrasound system ARTHUR V.2.0 with AI analysis DIANA V.2.0 matches expert rheumatologist in hand joint assessment of rheumatoid arthritis patients. *RMD Open* **2025**, *11*, 005805. [CrossRef] [PubMed]
216. Nigro, A. Fast-track capillaroscopic progression in systemic sclerosis: A case-based review of active pattern emerging within 3 months of raynaud's phenomenon onset. *Rheumatol. Int.* **2025**, *45*, 203. [CrossRef] [PubMed]
217. Giansanti, D. Revolutionizing Medical Imaging: The Transformative Role of Artificial Intelligence in Diagnostics and Treatment. *Diagnostics* **2025**, *15*, 1557. [CrossRef]
218. Kocak, B.; Ponsiglione, A.; Stanzione, A.; Bluethgen, C.; Santinha, J.; Ugga, L.; Huisman, M.; Klontzas, M.E.; Cannella, R.; Cuocolo, R. Bias in artificial intelligence for medical imaging: Fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. *Diagn. Interv. Radiol.* **2025**, *31*, 75–88. [CrossRef]
219. Hasanzadeh, F.; Josephson, C.B.; Waters, G.; Adedinsewo, D.; Azizi, Z.; White, J.A. Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *NPJ Digit. Med.* **2025**, *8*, 154. [CrossRef]
220. Koo, B.S.; Eun, S.; Shin, K.; Yoon, H.; Hong, C.; Kim, D.H.; Hong, S.; Kim, Y.G.; Lee, C.K.; Yoo, B.; et al. Machine learning model for identifying important clinical features for predicting remission in patients with rheumatoid arthritis treated with biologics. *Arthritis Res. Ther.* **2021**, *23*, 178. [CrossRef]
221. Bellocchi, C.; Favalli, E.G.; Maioli, G.; Agape, E.; Rossato, M.; Paini, M.; Severino, A.; Vigone, B.; Biggioggero, M.; Trombetta, E.; et al. Whole-Blood RNA Sequencing Profiling of Patients with Rheumatoid Arthritis Treated with Tofacitinib. *ACR Open Rheumatol.* **2025**, *7*, e11761. [CrossRef] [PubMed]
222. Lee, H.K.; Jung, O.; Hennighausen, L. JAK inhibitors dampen activation of interferon-stimulated transcription of ACE2 isoforms in human airway epithelial cells. *Commun. Biol.* **2021**, *4*, 654. [CrossRef]

223. Bergero, M.A.; Martinez, P.; Modina, P.; Hosman, R.; Villamil, W.; Gudino, R.; David, C.; Costa, L. Artificial intelligence model for predicting early biochemical recurrence of prostate cancer after robotic-assisted radical prostatectomy. *Sci. Rep.* **2025**, *15*, 30822. [CrossRef]
224. Chen, M.M.; Rosenkrantz, A.B.; Nicola, G.N.; Silva, E., 3rd; McGinty, G.; Manchikanti, L.; Hirsch, J.A. The Qualified Clinical Data Registry: A Pathway to Success Within MACRA. *AJNR Am. J. Neuroradiol.* **2017**, *38*, 1292–1296. [CrossRef]
225. Kersey, E.; Li, J.; Adler-Milstein, J.; Yazdany, J.; Shiboski, S.; Schmajuk, G. Association of Qualified Clinical Data Registry Clinician Dashboard Engagement with Performance on Quality-of-Care Measures: Cross-Sectional Analysis. *J. Med. Internet Res.* **2025**, *27*, e72709. [CrossRef]
226. Tabari, P.; Costagliola, G.; De Rosa, M.; Boeker, M. State-of-the-Art Fast Healthcare Interoperability Resources (FHIR)-Based Data Model and Structure Implementations: Systematic Scoping Review. *JMIR Med. Inform.* **2024**, *12*, e58445. [CrossRef]
227. Marfoglia, A.; Nardini, F.; Arcobelli, V.A.; Moscato, S.; Mellone, S.; Carbonaro, A. Towards real-world clinical data standardization: A modular FHIR-driven transformation pipeline to enhance semantic interoperability in healthcare. *Comput. Biol. Med.* **2025**, *187*, 109745. [CrossRef]
228. Rossander, A.; Lindskold, L.; Ranerup, A.; Karlsson, D. A State-of-the Art Review of SNOMED CT Terminology Binding and Recommendations for Practice and Research. *Methods Inf. Med.* **2021**, *60*, e76–e88. [CrossRef] [PubMed]
229. Bakken, S. Standards and frameworks. *J. Am. Med. Inform. Assoc.* **2024**, *31*, 1629–1630. [CrossRef] [PubMed]
230. Shen, Y.; Yu, J.; Zhou, J.; Hu, G. Twenty-Five Years of Evolution and Hurdles in Electronic Health Records and Interoperability in Medical Research: Comprehensive Review. *J. Med. Internet Res.* **2025**, *27*, e59024. [CrossRef]
231. Gulden, C.; Macho, P.; Reinecke, I.; Strantz, C.; Prokosch, H.U.; Blasini, R. recruIT: A cloud-native clinical trial recruitment support system based on Health Level 7 Fast Healthcare Interoperability Resources (HL7 FHIR) and the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM). *Comput. Biol. Med.* **2024**, *174*, 108411. [CrossRef]
232. Kim, D.; Oh, K.; Lee, Y.; Woo, H. Overview of fair federated learning for fairness and privacy preservation. *Expert Syst. Appl.* **2025**, *293*, 128568. [CrossRef]
233. Ye, H.; Zhang, X.; Liu, K.; Liu, Z.; Chen, W.; Liu, B.; Ngai, E.W.; Hu, Y. A personalized federated learning approach to enhance joint modeling for heterogeneous medical institutions. *Digit. Health* **2025**, *11*, 20552076251360861. [CrossRef] [PubMed]
234. Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **2021**, *27*, 1735–1743. [CrossRef] [PubMed]
235. Solitano, V.; Ahuja, D.; Lee, H.H.; Gaikwad, R.; Yeh, K.H.; Facciorusso, A.; Singh, A.G.; Ma, C.; Ananthakrishnan, A.N.; Yuan, Y.; et al. Comparative Safety of JAK Inhibitors vs TNF Antagonists in Immune-Mediated Inflammatory Diseases: A Systematic Review and Meta-Analysis. *JAMA Netw. Open* **2025**, *8*, e2531204. [CrossRef]
236. Li, H.; Zang, C.; Xu, Z.; Pan, W.; Rajendran, S.; Chen, Y.; Wang, F. Federated target trial emulation using distributed observational data for treatment effect estimation. *NPJ Digit. Med.* **2025**, *8*, 387. [CrossRef]
237. Matta, S.; Lamard, M.; Zhang, P.; Le Guilcher, A.; Borderie, L.; Cochener, B.; Quellec, G. A systematic review of generalization research in medical image classification. *Comput. Biol. Med.* **2024**, *183*, 109256. [CrossRef]
238. Le, J.P.; Shashikumar, S.P.; Malhotra, A.; Nemati, S.; Wardi, G. Making the Improbable Possible: Generalizing Models Designed for a Syndrome-Based, Heterogeneous Patient Landscape. *Crit. Care Clin.* **2023**, *39*, 751–768. [CrossRef]
239. Zhu, H.; Bai, J.; Li, N.; Li, X.; Liu, D.; Buckeridge, D.L.; Li, Y. FedWeight: Mitigating covariate shift of federated learning on electronic health records data through patients re-weighting. *NPJ Digit. Med.* **2025**, *8*, 286. [CrossRef]
240. Wu, Q.; Reps, J.M.; Li, L.; Zhang, B.; Lu, Y.; Tong, J.; Zhang, D.; Lumley, T.; Brand, M.T.; Van Zandt, M.; et al. COLA-GLM: Collaborative one-shot and lossless algorithms of generalized linear models for decentralized observational healthcare data. *NPJ Digit. Med.* **2025**, *8*, 442. [CrossRef]
241. Zhang, F.; Zhai, D.; Bai, G.; Jiang, J.; Ye, Q.; Ji, X.; Liu, X. Towards fairness-aware and privacy-preserving enhanced collaborative learning for healthcare. *Nat. Commun.* **2025**, *16*, 2852. [CrossRef]
242. Curtis, J.R.; Su, Y.; Black, S.; Xu, S.; Langholff, W.; Bingham, C.O.; Kafka, S.; Xie, F. Machine Learning Applied to Patient-Reported Outcomes to Classify Physician-Derived Measures of Rheumatoid Arthritis Disease Activity. *ACR Open Rheumatol.* **2022**, *4*, 995–1003. [CrossRef] [PubMed]
243. Williams, E.; Kienast, M.; Medawar, E.; Reinelt, J.; Merola, A.; Klopfenstein, S.A.I.; Flint, A.R.; Heeren, P.; Poncette, A.S.; Balzer, F.; et al. A Standardized Clinical Data Harmonization Pipeline for Scalable AI Application Deployment (FHIR-DHP): Validation and Usability Study. *JMIR Med. Inform.* **2023**, *11*, e43847. [CrossRef] [PubMed]
244. Xiao, G.; Pfaff, E.; Prud'hommeaux, E.; Booth, D.; Sharma, D.K.; Huo, N.; Yu, Y.; Zong, N.; Ruddy, K.J.; Chute, C.G.; et al. FHIR-Ontop-OMOP: Building clinical knowledge graphs in FHIR RDF with the OMOP Common data Model. *J. Biomed. Inform.* **2022**, *134*, 104201. [CrossRef]
245. El Arab, R.A.; Al Moosa, O.A. Systematic review of cost effectiveness and budget impact of artificial intelligence in healthcare. *NPJ Digit. Med.* **2025**, *8*, 548. [CrossRef] [PubMed]

246. Scipion, C.E.A.; Manchester, M.A.; Federman, A.; Wang, Y.; Arias, J.J. Barriers to and facilitators of clinician acceptance and use of artificial intelligence in healthcare settings: A scoping review. *BMJ Open* **2025**, *15*, e092624. [CrossRef]
247. Wei, Q.; Pan, S.; Liu, X.; Hong, M.; Nong, C.; Zhang, W. The integration of AI in nursing: Addressing current applications, challenges, and future directions. *Front. Med.* **2025**, *12*, 1545420. [CrossRef]
248. Nowell, W.B.; Curtis, J.R. Remote Therapeutic Monitoring in Rheumatic and Musculoskeletal Diseases: Opportunities and Implementation. *Med. Res. Arch.* **2023**, *11*, 3957. [CrossRef]
249. FDA. *Artificial Intelligence and Machine Learning Software as a Medical Device (SaMD)*; FDA: Silver Spring, MD, USA, 2024.
250. EMA. *Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle (EMA/CHMP/CVMP/83833/2023)*; EMA: Amsterdam, NL, USA, 2024.
251. Hopewell, S.; Chan, A.W.; Collins, G.S.; Hrobjartsson, A.; Moher, D.; Schulz, K.F.; Tunn, R.; Aggarwal, R.; Berkwits, M.; Berlin, J.A.; et al. CONSORT 2025 explanation and elaboration: Updated guideline for reporting randomised trials. *BMJ* **2025**, *389*, e081124. [CrossRef]
252. Collins, G.S.; Dhiman, P.; Ma, J.; Schlussel, M.M.; Archer, L.; Van Calster, B.; Harrell, F.E., Jr.; Martin, G.P.; Moons, K.G.M.; van Smeden, M.; et al. Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ* **2024**, *384*, e074819. [CrossRef]
253. de Hond, A.A.H.; Shah, V.B.; Kant, I.M.J.; Van Calster, B.; Steyerberg, E.W.; Hernandez-Boussard, T. Perspectives on validation of clinical predictive algorithms. *NPJ Digit. Med.* **2023**, *6*, 86. [CrossRef]
254. Piovani, D.; Sokou, R.; Tsantes, A.G.; Vitello, A.S.; Bonovas, S. Optimizing Clinical Decision Making with Decision Curve Analysis: Insights for Clinical Investigators. *Healthcare* **2023**, *11*, 2244. [CrossRef]
255. Vickers, A.J.; Holland, F. Decision curve analysis to evaluate the clinical benefit of prediction models. *Spine J.* **2021**, *21*, 1643–1648. [CrossRef] [PubMed]
256. Kerr, K.F.; Brown, M.D.; Zhu, K.; Janes, H. Assessing the Clinical Impact of Risk Prediction Models with Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *J. Clin. Oncol.* **2016**, *34*, 2534–2540. [CrossRef] [PubMed]
257. Tam, T.Y.C.; Sivarajkumar, S.; Kapoor, S.; Stolyar, A.V.; Polanska, K.; McCarthy, K.R.; Osterhoudt, H.; Wu, X.; Visweswaran, S.; Fu, S.; et al. A framework for human evaluation of large language models in healthcare derived from literature review. *NPJ Digit. Med.* **2024**, *7*, 258. [CrossRef] [PubMed]
258. Venerito, V. Artificial intelligence in rheumatology: Days of a future past. *Rheumatol. Adv. Pract.* **2025**, *9*, rkaf022. [CrossRef]
259. Chopra, H.; Annu; Shin, D.K.; Munjal, K.; Priyanka; Dhama, K.; Emran, T.B. Revolutionizing clinical trials: The role of AI in accelerating medical breakthroughs. *Int. J. Surg.* **2023**, *109*, 4211–4220. [CrossRef]
260. Dolin, P.; Li, W.; Dasarathy, G.; Berisha, V. Statistically Valid Post-Deployment Monitoring Should Be Standard for AI-Based Digital Health. *arXiv* **2025**. [CrossRef]
261. Wang, J.; Tian, Y.; Zhou, T.; Tong, D.; Ma, J.; Li, J. A survey of artificial intelligence in rheumatoid arthritis. *Rheumatol. Immunol. Res.* **2023**, *4*, 69–77. [CrossRef]
262. Moreno-Grau, S.; Vernekar, M.; Lopez-Pineda, A.; Mas-Montserrat, D.; Barrabes, M.; Quinto-Cortes, C.D.; Moatamed, B.; Lee, M.T.M.; Yu, Z.; Numakura, K.; et al. Polygenic risk score portability for common diseases across genetically diverse populations. *Hum. Genomics* **2024**, *18*, 93. [CrossRef]
263. Miao, J.; Guo, H.; Song, G.; Zhao, Z.; Hou, L.; Lu, Q. Quantifying portable genetic effects and improving cross-ancestry genetic prediction with GWAS summary statistics. *Nat. Commun.* **2023**, *14*, 832. [CrossRef]
264. Roschewitz, M.; Khara, G.; Yearsley, J.; Sharma, N.; James, J.J.; Ambrozay, E.; Heroux, A.; Kecskemethy, P.; Rijken, T.; Glocker, B. Automatic correction of performance drift under acquisition shift in medical image classification. *Nat. Commun.* **2023**, *14*, 6608. [CrossRef]
265. Lambert, B.; Forbes, F.; Doyle, S.; Dehaene, H.; Dojat, M. Trustworthy clinical AI solutions: A unified review of uncertainty quantification in Deep Learning models for medical image analysis. *Artif. Intell. Med.* **2024**, *150*, 102830. [CrossRef]
266. Gilbert, S.; Adler, R.; Holoyad, T.; Weicken, E. Could transparent model cards with layered accessible information drive trust and safety in health AI? *NPJ Digit. Med.* **2025**, *8*, 124. [CrossRef] [PubMed]
267. Chinta, S.V.; Wang, Z.; Palikhe, A.; Zhang, X.; Kashif, A.; Smith, M.A.; Liu, J.; Zhang, W. AI-driven healthcare: Fairness in AI healthcare: A survey. *PLoS Digit. Health* **2025**, *4*, e0000864. [CrossRef] [PubMed]
268. Gallifant, J.; Kistler, E.A.; Nakayama, L.F.; Zera, C.; Kripalani, S.; Ntatin, A.; Fernandez, L.; Bates, D.; Dankwa-Mullan, I.; Celi, L.A. Disparity dashboards: An evaluation of the literature and framework for health equity improvement. *Lancet Digit. Health* **2023**, *5*, e831–e839. [CrossRef]
269. Alderman, J.E.; Palmer, J.; Laws, E.; McCradden, M.D.; Ordish, J.; Ghassemi, M.; Pfohl, S.R.; Rostamzadeh, N.; Cole-Lewis, H.; Glocker, B.; et al. Tackling algorithmic bias and promoting transparency in health datasets: The STANDING Together consensus recommendations. *Lancet Digit. Health* **2025**, *7*, e64–e88. [CrossRef] [PubMed]
270. Teikari, P.; Jarrell, M.; Azh, M.; Pesola, H. The Architecture of Trust: A Framework for AI-Augmented Real Estate Valuation in the Era of Structured Data. *arXiv* **2025**. [CrossRef]

271. Davis, S.E.; Embi, P.J.; Matheny, M.E. Sustainable deployment of clinical prediction tools-a 360 degrees approach to model maintenance. *J. Am. Med. Inform. Assoc.* **2024**, *31*, 1195–1198. [CrossRef]
272. Rosenthal, J.T.; Beecy, A.; Sabuncu, M.R. Rethinking clinical trials for medical AI with dynamic deployments of adaptive systems. *NPJ Digit. Med.* **2025**, *8*, 252. [CrossRef]
273. Zerrouk, N.; Auge, F.; Niarakis, A. Building a modular and multi-cellular virtual twin of the synovial joint in Rheumatoid Arthritis. *NPJ Digit. Med.* **2024**, *7*, 379. [CrossRef] [PubMed]

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.