Est. 1841 | YORK ST JOHN UNIVERSITY

Ojo, Adeyinka, Osborne, Augustus, Soladoye, Afeez A., Matope, Euphemia Murateng, Bamidele, Doyin Blessing and Olawade, David (2026) The Application of Artificial Intelligence in Research Protocol Development: A Critical Analysis of Benefits, Limitations, and Political-Economic Contexts. Intelligent Hospital, 2 (3). p. 100053.

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/14058/

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version: https://doi.org/10.1016/j.inhs.2026.100053

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repositories Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Intelligent Hospital 2 (2026) 100053

Contents lists available at ScienceDirect

Intelligent Hospital

journal homepage: www.keaipublishing.com/en/journals/intelligent-hospital/

Review

The application of artificial intelligence in research protocol development: A critical analysis of benefits, limitations, and political-economic contexts

Adeyinka Ojo [a], Augustus Osborne [b], Afeez A. Soladoye [c,d], Euphemia Murateng Matope [e], Doyin Blessing Bamidele [f], David B. Olawade [g,h,i,*]

[a] Department of Artificial Intelligence and Data Factory, Capgemini plc, London EC4V 4HN, United Kingdom
[b] Institute for Development, Western Area, Freetown, Sierra Leone
[c] Department of Computer Engineering, Federal University, Oye, Ekiti, Nigeria
[d] Department of Computer Engineering, Adeleke University, Ede, Nigeria
[e] Department of Biomedical Science, Middlesex University, London, United Kingdom
[f] Department of Biomedical Science, University of Wolverhampton, Wolverhampton, United Kingdom
[g] Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham ME7 5NY, United Kingdom
[h] Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, United Kingdom
[i] Department of Public Health, York St John University, London E14 2BA, United Kingdom

ARTICLE INFO

Keywords:
Artificial intelligence
Research protocol
Machine learning
Natural language processing
Study design
Research methodology

ABSTRACT

Background: Research protocol development represents a fundamental cornerstone of scientific inquiry, yet traditional approaches relying on domain expertise and manual literature reviews are increasingly inadequate for addressing contemporary research complexity and exponential scientific data growth. The integration of artificial intelligence into this process occurs within contested political, economic, and epistemological contexts that shape what research becomes possible or prioritized.

Objectives: This narrative review systematically examines AI technologies applicable to protocol development, analyses real-world case studies across diverse research domains, identifies benefits and limitations of AI integration, critically evaluates political-economic contexts shaping AI deployment, and proposes recommendations acknowledging power asymmetries and epistemological constraints.

Methods: A comprehensive literature search was conducted across PubMed, IEEE Xplore, ACM Digital Library, Web of Science, and Google Scholar (2015–2025). Inclusion criteria encompassed peer-reviewed articles, conference proceedings, technical reports, and case studies addressing AI applications in research protocol development.

Results: AI applications suggest emerging potential across natural language processing for literature synthesis, machine learning for study design optimisation, and automated ethical compliance checking. However, challenges persist including data quality issues, algorithmic transparency concerns, corporate control of AI systems, political restrictions on certain research domains, and systematic bias toward methodologically conservative designs that may disadvantage novel, theoretically complex, or population-heterogeneous research. Critical examination reveals AI optimization encodes particular values about what constitutes "good science," potentially foreclosing paradigm-challenging research and methodologically innovative approaches. Independent peer-reviewed validation remains limited for many claimed benefits, with substantial evidence deriving from corporate sources rather than independent evaluation.

Conclusion: AI offers reported potential for research protocol development. However, realization occurs within power structures shaped by corporate ownership, political constraints, and epistemological biases. Future success requires collaborative human-AI frameworks maintaining critical awareness of how efficiency metrics may systematically disadvantage certain knowledge production forms, supported by independent evaluation rather than corporate claims.

Peer review under responsibility of Editorial Board of Intelligent Hospital
* Corresponding author at: Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham ME7 5NY, United Kingdom.
E-mail address: d.olawade@uel.ac.uk (D.B. Olawade).

https://doi.org/10.1016/j.inhs.2026.100053
Received 23 January 2026; Received in revised form 23 February 2026; Accepted 23 February 2026
Available online 24 February 2026
3050-8371/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The development of robust research protocols constitutes the foundational framework upon which all scientific investigations are built, serving as the blueprint that guides methodological approaches, data collection strategies, and analytical procedures.[1] The quality and comprehensiveness of research protocols directly influence the validity, reliability, and reproducibility of study outcomes, making protocol development a critical determinant of scientific advancement and knowledge generation.[2,3] Traditional protocol development processes have historically depended on researchers' domain expertise, extensive manual literature reviews, consultation with statistical experts, and iterative refinement through peer review and institutional oversight.

However, the contemporary research landscape presents unprecedented challenges that strain conventional protocol development approaches. The exponential growth of scientific literature, with over 2.5 million new peer-reviewed articles published annually, makes comprehensive manual literature reviews increasingly impractical and potentially incomplete.[4,5] The complexity of modern research questions often requires interdisciplinary expertise that may exceed individual researchers' knowledge domains, while the pressure for rapid research translation and evidence-based decision-making demands more efficient protocol development processes.[6,7] Additionally, the increasing emphasis on reproducibility and transparency in scientific research necessitates more rigorous and standardised approaches to protocol design and validation.

It is important to note that AI deployment in research methodology occurs within contested political and economic contexts. Federal restrictions in certain jurisdictions have constrained diversity-related research and climate change investigation,[8,9] while corporate ownership of major AI platforms (primarily US-based systems) raises questions about scientific sovereignty and whose agendas shape what constitutes "optimal" methodology.[10] The integration of AI into protocol development is not a politically neutral technical advancement but occurs within power structures that may systematically advantage certain research questions while marginalizing others. For example, a hospital research department in a jurisdiction with federal diversity research restrictions may find AI systems automatically flagging equity-focused protocols as "high-risk" or "politically sensitive," effectively encoding political constraints as methodological limitations even when the research design is scientifically sound. Similarly, a multi-site international study on climate health impacts may receive differential AI recommendations across jurisdictions based on local funding environments rather than scientific merit.

The emergence of artificial intelligence technologies offers transformative potential for addressing these challenges in research protocol development. AI encompasses a broad spectrum of computational approaches, including machine learning algorithms that can identify patterns in large datasets, natural language processing systems that can analyse and synthesise textual information, and expert systems that can encode and apply domain-specific knowledge.[11,12] These technologies have demonstrated remarkable capabilities in automating complex cognitive tasks, processing vast amounts of information, and providing data-driven insights that can inform decision-making processes.

Recent advances in AI applications across various scientific domains have demonstrated the technology's potential to enhance research efficiency, accuracy, and innovation. From drug discovery and clinical trial design to environmental monitoring and social science research, AI tools are increasingly being integrated into research workflows to augment human capabilities and address methodological challenges.[13] The application of AI in research protocol development represents a natural extension of these capabilities, offering opportunities to streamline literature synthesis, optimise study designs, ensure compliance with ethical frameworks, and improve the overall quality of research protocols.[14,15]

However, critical examination is required of how AI systems encode particular epistemological assumptions about what constitutes "good science." When AI recommendations derive from historical literature reflecting homogeneous sampling practices justified by statistical power considerations, these systems may systematically recommend protocols that treat population heterogeneity as noise to be eliminated rather than meaningful variance requiring investigation.[16,17] This becomes particularly problematic when identical symptom presentations reflect completely different neurobiological mechanisms across populations, a distinction AI systems trained on methodologically conservative literature may fail to accommodate.

Despite the growing interest in AI applications for research enhancement, there remains a significant gap in comprehensive understanding of how AI can be effectively integrated into research protocol development processes. Current literature lacks a systematic examination of AI tools and techniques specifically designed for protocol development, their practical applications across different research domains, and the challenges and limitations associated with their implementation. More critically, existing literature has not adequately examined whose interests AI systems serve, what research questions become systematically disadvantaged when efficiency becomes the primary value, and how corporate and political pressures shape AI recommendations in ways that may constrain rather than enable scientific inquiry.

The novelty of this review lies in its comprehensive analysis of AI applications specifically focused on research protocol development, providing a structured examination of current capabilities, real-world implementations, and future directions while maintaining critical awareness of power asymmetries, epistemological biases, and political-economic contexts shaping AI deployment. This narrative review aims to provide a comprehensive analysis of artificial intelligence applications in research protocol development, examining current technologies, practical implementations, and future opportunities for enhancing research methodology. The review objectives are: (1) To systematically examine current AI technologies and tools applicable to research protocol development; (2) To analyse real-world case studies and implementations across diverse research domains with critical attention to evidence quality and corporate versus independent evaluation; (3) To identify key benefits, challenges, and limitations of AI integration in protocol development including systematic biases toward particular research paradigms; (4) To critically evaluate political-economic contexts including corporate control, federal research restrictions, and implications for scientific sovereignty; (5) To propose recommendations for future development and implementation of AI-assisted protocol development systems that acknowledge and mitigate power asymmetries.

2. Methods

2.1. Search strategy and information sources

This narrative review employed a comprehensive literature search strategy to identify relevant studies, reports, and case studies related to artificial intelligence applications in research protocol development. The search was conducted across multiple electronic databases including PubMed, IEEE Xplore, ACM Digital Library, Web of Science, and Google Scholar, covering the period from 2015 to 2025. Searches were conducted between March 15–28, 2025.

Core search strategies employed database-specific syntax combining:

- Primary terms: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "natural language processing" OR "AI")
- AND
- Protocol terms: ("research protocol" OR "study design" OR "protocol development" OR "research methodology" OR "clinical trial design" OR "experimental design")

Example. PubMed search string: ("artificial intelligence"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "natural language processing"[Title/Abstract]) AND ("research protocol"[Title/Abstract] OR "protocol development"[Title/Abstract] OR "study design"[Title/Abstract]) AND ("2015"[Date - Publication]: "2025"[Date - Publication])

Additional targeted searches examined specific AI applications (e.g., "GPT-4 protocol development," "Novartis AI clinical trials," "NIH BRAIN Initiative AI") and corporate implementations to capture grey literature and industry developments.

2.2. Inclusion and exclusion criteria

The inclusion criteria encompassed peer-reviewed articles, conference proceedings, technical reports, and documented case studies that specifically addressed AI applications in research protocol development or related methodological processes. Both empirical studies and theoretical frameworks were included to provide a comprehensive overview of the field. Exclusion criteria included studies focusing solely on AI applications in data analysis without protocol development components, non-English publications, and studies published before 2015 to ensure currency of technological applications.

2.3. Evidence quality assessment framework

As shown in Table 1, evidence sources were classified using a four-tier framework based on validation rigor and independence:

Evidence quality for case studies (Table 4) was assessed independently by two reviewers (AO and DBO) using predefined criteria: (1) presence of independent peer-reviewed validation, (2) disclosure of methodology and data, (3) independence from commercial interests, and (4) replicability of claims. Discrepancies were resolved through discussion and consensus. When evaluating conflicting evidence across tiers, Tier 1 sources were prioritized over lower tiers. Corporate sources (Tier 4) were included for completeness but consistently flagged as requiring independent verification, and claims from these sources are presented with appropriate qualifiers ("reported," "claimed," "credited with") throughout the manuscript.

The review acknowledges inherent limitations in available evidence, particularly the predominance of corporate sources (press releases, conference presentations) over peer-reviewed independent evaluation for commercial AI implementations. This evidence limitation is addressed throughout the manuscript with explicit acknowledgment of verification gaps and corporate versus independent sources.

Given the rapidly evolving nature of AI technology and its applications in research methodology, this narrative review approach was deemed most appropriate to synthesise diverse types of evidence, including emerging tools, pilot implementations, and theoretical frameworks that may not yet have extensive empirical validation. The review also incorporated analysis of commercially available AI tools and platforms specifically designed for research protocol development to provide practical insights for researchers and institutions considering AI adoption. However, we maintain critical awareness that commercial promotion may not constitute robust evidence of effectiveness or appropriateness.

Table 1
Evidence quality tiering framework.

Tier	Evidence Type	Characteristics	Quality Level
Tier 1	Independent peer-reviewed validation	Published in academic journals with external peer review; independent validation; reproducible methodology; disclosed conflicts of interest	High
Tier 2	Institutional technical reports	Governmental/academic institution reports with internal review; methodology disclosed but limited external validation	Moderate
Tier 3	Conference presentations with peer review	Conference proceedings with peer review; limited validation; preliminary findings	Low-Moderate
Tier 4	Corporate materials	Press releases, white papers, promotional content; no independent validation; potential commercial bias	Low

3. AI technologies in research protocol development

3.1. Natural language processing for literature synthesis

Natural language processing represents one of the most mature AI applications in research protocol development, offering sophisticated capabilities for automated literature review and synthesis. Modern NLP systems utilise transformer-based architectures and large language models to process, analyse, and synthesise vast quantities of scientific literature with unprecedented speed and accuracy.[18,19] These systems can identify relevant studies, extract key methodological information, and generate comprehensive literature summaries that inform protocol development decisions.

Advanced NLP tools such as Semantic Scholar's AI-powered search capabilities and Connected Papers' visual literature mapping provide researchers with intelligent literature discovery and analysis features. These platforms employ machine learning algorithms to understand semantic relationships between research papers, identify research gaps, and suggest relevant methodological approaches based on similar studies.[20] The integration of citation analysis with content analysis enables more sophisticated understanding of methodological trends and best practices within specific research domains.

Recent developments in generative AI, particularly large language models like GPT-4 and Claude, have introduced new possibilities for automated protocol drafting and refinement. These systems, predominantly owned and controlled by US-based corporations (OpenAI, Anthropic, Google), raise questions about scientific sovereignty when international research communities become dependent on commercial platforms whose priorities may shift with corporate or political pressures.[21,22] These systems can generate initial protocol sections based on research objectives and constraints, suggest appropriate methodological approaches, and identify potential gaps or inconsistencies in protocol design.[11,14] However, the application of these technologies requires careful validation and human oversight to ensure accuracy and appropriateness of generated content. Critical examination is needed of what agendas are encoded in these corporate systems and how commercial interests shape what counts as "optimal" methodology.

The effectiveness of NLP in literature synthesis is particularly evident in systematic review preparation, where automated screening and data extraction can significantly reduce the time and effort required for comprehensive literature analysis. Tools like Rayyan and Covidence have integrated AI features to assist with title and abstract screening, while more advanced systems can perform full-text analysis and methodological quality assessment.[23]

However, when NLP systems are trained predominantly on literature reflecting homogeneous sampling practices, they may systematically recommend protocols that perpetuate these patterns under the guise of "best practice" optimization. If AI identifies that studies with restricted demographic variance achieve higher citation rates or publication success, the system may recommend population restriction as methodological optimization, thereby encoding existing biases as technical imperatives.[24,25]

3.2. Machine learning for study design optimisation

Machine learning algorithms offer powerful capabilities for optimising study design decisions through predictive modelling and pattern recognition in existing research data. These systems can analyse large databases of completed studies to identify factors associated with successful outcomes, optimal sample sizes, and effective methodological approaches.[26] Furthermore, Bayesian networks and decision tree models are particularly useful for modelling complex relationships between study design parameters and expected outcomes.[27,28]

Reinforcement learning approaches have shown promise in adaptive study design, where protocols can be modified in real-time based on accumulating data and interim analyses. This approach is particularly valuable in clinical trial design, where adaptive protocols can improve efficiency and reduce costs while maintaining statistical validity.[29,30] The integration of machine learning with simulation studies enables researchers to explore multiple design scenarios and identify optimal configurations before study implementation.

Critical questions arise regarding what constitutes "successful outcomes" in training data. When machine learning systems optimize based on historical publication success, citation metrics, or commercial viability, they may systematically disadvantage paradigm-challenging research, theoretically complex designs, or investigation of understudied populations that face structural barriers to traditional success metrics.[31,32] Efficiency optimization is not value-neutral, it encodes particular assumptions about what research is worth doing.

Table 2 provides a systematic overview of the primary AI technologies employed in protocol development, detailing their specific applications, benefits, example tools, and inherent limitations across natural language processing, machine learning, expert systems, and deep learning approaches. Critical limitations include corporate control, epistemological biases, and systematic disadvantaging of certain research paradigms.

The application of machine learning in study design optimisation extends beyond traditional statistical considerations to include practical factors such as recruitment feasibility, resource allocation, and timeline optimisation. Predictive models can analyse historical recruitment data, demographic patterns, and institutional capabilities to provide realistic estimates of study feasibility and duration.[33,34]

However, when AI systems recommend "feasible" recruitment strategies based on historical success with easily accessible populations, they may systematically discourage research with understudied or harder-to-reach communities, thereby perpetuating research inequities under the label of optimization.[8,17]

3.3. AI-driven sample size and statistical planning

Traditional sample size calculations rely on pre-specified effect sizes, power levels, and statistical assumptions that may not accurately reflect the complexity of real-world research scenarios. AI-enhanced approaches to sample size determination leverage historical data, simulation studies, and adaptive algorithms to provide more accurate and flexible sample size recommendations.[35] These systems can account for multiple outcomes, complex statistical models, and uncertainty in effect size estimates.

Machine learning models trained on large databases of completed studies can predict likely effect sizes and variance parameters based on study characteristics, population demographics, and methodological approaches. This data-driven approach to sample size planning can reduce the risk of underpowered studies while avoiding unnecessary over-recruitment.[35,36] Additionally, AI systems can recommend optimal allocation ratios, stratification strategies, and interim analysis plans based on study objectives and constraints.

Critical examination reveals that when AI systems optimize for statistical power using historical effect sizes derived from homogeneous samples, they may recommend sample restrictions that enhance statistical efficiency while undermining external validity and mechanistic understanding across heterogeneous populations.[16,37] The fundamental distinction between population diversity (demographic representation) and brain diversity (neurobiological heterogeneity) is rarely accommodated in current AI systems, which may treat within-group variance as methodological noise rather than meaningful biological signal.

4. Applications across research domains

4.1. Biomedical and clinical research

The biomedical research domain has been at the forefront of AI adoption in protocol development, driven by the complexity of clinical trials and the high stakes of medical research. AI platforms specifically designed for clinical research, such as Protocol Builder and TrialPath, provide comprehensive support for generating ethics-compliant, regulatory-ready protocols.[38,39] These systems integrate multiple AI technologies to assist with literature review, study design optimisation, sample size calculation, and regulatory compliance checking.

Clinical trial protocol development presents unique challenges, including complex inclusion and exclusion criteria, multiple endpoints, adaptive design considerations, and stringent regulatory requirements. AI systems have demonstrated particular value in optimising patient selection criteria by analysing electronic health records and identifying biomarkers or clinical characteristics associated with treatment response.[37,40] This capability enables more precise targeting of study

Table 2
AI technologies and their applications in research protocol development (with critical limitations).

AI Technology	Primary Application	Key Benefits	Example Tools	Limitations	Critical Concerns
Natural Language Processing[18–20]	Literature synthesis, text analysis	Rapid literature review, gap identification	Semantic Scholar, Connected Papers, Elicit	Quality dependent on training data, potential bias in source selection	Corporate ownership (US-based), encodes historical biases as "best practice," may systematically recommend homogeneous sampling
Machine Learning[26–28]	Study design optimisation, predictive modelling	Pattern recognition, outcome prediction	AutoML platforms, Bayesian networks	Requires large training datasets, limited interpretability	Optimizes for conventional success metrics, may disadvantage novel/complex designs, treats heterogeneity as noise
Expert Systems[29,30]	Protocol validation, compliance checking	Rule-based validation, consistency checking	Protocol Builder, TrialPath	Limited to programmed rules, may miss novel scenarios	Rules reflect corporate/regulatory priorities, may enforce political restrictions on research domains
Deep Learning[20]	Complex pattern recognition, multimodal analysis	Advanced feature extraction, integration of diverse data types	TensorFlow, PyTorch applications	Black box problem, computational requirements	Opacity prevents understanding whose interests are served, concentrates power in technically sophisticated actors
Computer Vision[29,30]	Image-based protocol elements, data collection planning	Automated image analysis planning, quality control	OpenCV, medical imaging AI	Limited to visual data applications	Training data biases, commercial platforms dominate

populations and improved likelihood of trial success.

However, "treatment response" optimization based on historical data may systematically exclude populations with different neurobiological mechanisms underlying identical symptom presentations. When AI recommends homogeneous samples for statistical efficiency, it may produce findings with limited mechanistic validity across diverse populations, a validity threat distinct from mere generalizability limitations.[8,16]

The integration of AI with clinical data management systems has enabled real-time protocol monitoring and adaptive modifications based on accumulating evidence. Machine learning algorithms can identify trends in recruitment, safety signals, and efficacy outcomes that may warrant protocol amendments or early stopping decisions.[33,34] This capability is particularly valuable in oncology trials, where patient populations may be heterogeneous and treatment responses can vary significantly.

4.2. Environmental and social sciences

Environmental research presents unique methodological challenges related to spatial and temporal variability, complex ecological interactions, and multi-scale phenomena. However, we acknowledge that climate change research faces political restrictions in certain jurisdictions, with federal funding limitations affecting what questions can be investigated and how findings can be disseminated.[41,42] AI applications in environmental protocol development focus on optimising sampling strategies, predicting environmental conditions, and designing monitoring networks.[14,41] Machine learning models can analyse historical environmental data, satellite imagery, and climate projections to recommend optimal sampling locations, timing, and frequency.

Critical questions arise: Do AI systems trained during periods of less restrictive research environments recommend protocols that current political climates render unfundable or unpublishable? How do corporate AI platforms respond when political pressures demand limitations on climate-related research? The treatment of climate modeling and environmental monitoring as politically neutral technical exercises (as in earlier AI literature) is quite naive given 2026 realities.[14,15]

Social science research benefits from AI applications in survey design, participant recruitment strategies, and bias detection.[43] Natural language processing can analyse existing survey instruments to identify potential sources of bias, suggest improved question wording, and recommend appropriate statistical analysis approaches.[44] AI systems can also assist with power analysis for complex social science designs, including multilevel models, longitudinal studies, and quasi-experimental approaches.

However, federal restrictions on diversity-related research in certain jurisdictions mean that AI recommendations for studying equity, discrimination, or demographic disparities may produce protocols that are technically sound but politically unfundable.[9,25] Furthermore, when AI systems flag diversity-focused research as "controversial" or "high-risk" based on current funding climates, they encode political restrictions as methodological imperatives.

Table 3 examines domain-specific implementations, comparing how AI applications vary across clinical research, environmental science, social sciences, and basic sciences, highlighting the unique benefits and challenges encountered in each field including political-economic constraints shaping what research becomes possible.

4.3. Interdisciplinary and translational research

Interdisciplinary research protocols require integration of methodological approaches from multiple disciplines, creating unique challenges for protocol development. AI systems can facilitate this integration by identifying compatible methodological approaches, suggesting appropriate statistical analysis plans for multi-domain data, and ensuring consistency across different disciplinary perspectives.[45] Natural language processing can analyse literature from diverse fields to identify successful interdisciplinary methodological approaches and potential areas of methodological conflict.

However, when AI systems optimize based on historical publication success, genuinely novel interdisciplinary combinations may be flagged as "high-risk" or "methodologically inconsistent" simply because they lack precedent. The pressure to make research "look conventional" to receive AI approval may systematically disadvantage the most innovative work.[15,31]

4.4. Ethical considerations and compliance framework

The integration of AI in research protocol development raises important ethical considerations that must be carefully addressed to ensure responsible implementation. AI systems used in protocol development must themselves adhere to principles of transparency, accountability, and fairness, while also ensuring that the protocols they help develop meet appropriate ethical standards.[10,46] This dual ethical responsibility creates unique challenges for researchers and institutions adopting AI-assisted protocol development.

Ethical AI tools have been developed to assess protocol drafts against established ethical frameworks such as the Declaration of Helsinki,[47] the Belmont Report,[48] and GDPR requirements for data protection. These systems can automatically flag potential ethical concerns, suggest appropriate consent procedures, and recommend safeguards for vulnerable populations. The Ethical AI Toolkit, for example, provides automated screening for missing consent procedures, inadequate data security measures, and potential risks to participant welfare.

Table 3
Domain-specific AI applications in research protocol development (with political-economic context).

Research Domain	AI Applications	Specific Benefits	Implementation Examples	Challenges	Political-Economic Context
Clinical Research[37,40]	Patient stratification, adaptive designs, regulatory compliance	Improved trial efficiency, reduced costs, faster approvals	Protocol Builder, TrialPath, IBM Watson Clinical	Regulatory acceptance, data privacy concerns	Corporate platform dominance (US-based), optimization may exclude diverse populations, commercial interests shape "optimal" methodology
Environmental Science[14,41]	Spatial sampling optimisation, climate modelling integration	Improved spatial coverage, climate-informed designs	GIS-AI integration, satellite data analysis	Data availability, model validation	Federal restrictions on climate research in certain jurisdictions, AI systems may recommend unfundable protocols, political constraints treated as technical limitations
Social Sciences[43,44]	Survey design optimisation, bias detection, sampling strategies	Reduced bias, improved generalisability	NLP survey analysis, demographic prediction models	Cultural sensitivity, ethical considerations	Federal restrictions on diversity-related research, AI may flag equity studies as "controversial," political unfundability encoded as methodological risk
Basic Sciences[45]	Experimental design optimisation, hypothesis generation	Enhanced reproducibility, novel insights	Laboratory automation AI, hypothesis generation systems	Limited domain-specific training data	Pressure toward commercially viable research questions, paradigm-challenging work may be flagged as "high-risk"

The transparency and explainability of AI recommendations in protocol development present particular challenges. Researchers and ethics review boards need to understand the basis for AI-generated recommendations to properly evaluate their appropriateness and potential limitations. This requirement has led to the development of explainable AI approaches specifically designed for research applications, which provide clear justifications for their recommendations and highlight areas of uncertainty.[10]

However, explainability remains fundamentally limited when AI systems encode complex corporate and political pressures. Even when AI provides "explanations" for recommendations, the deeper question of whose interests the system serves and what research it systematically marginalizes often remains opaque.[21,22] Furthermore, when AI tools enforce political restrictions by making non-compliant designs seem methodologically inferior, the ethical implications extend beyond traditional research ethics frameworks.

4.5. Bias detection and mitigation strategies

Research bias represents a fundamental threat to the validity and reliability of scientific findings, making bias detection and mitigation crucial components of protocol development. AI systems offer sophisticated capabilities for identifying potential sources of bias early in the protocol development process, before they can impact study implementation and outcomes.[22,25] These systems can analyse protocol components against large databases of completed studies to identify patterns associated with various forms of bias.

Machine learning algorithms can detect sampling bias by analysing proposed recruitment strategies and comparing them with successful approaches in similar studies. They can identify potential confounding variables based on study design characteristics and suggest appropriate control strategies.[9,17] Additionally, AI systems can flag potential measurement bias by analysing proposed outcome measures and data collection procedures against established validity and reliability criteria.

However, fundamental misunderstanding persists regarding what constitutes "bias" versus meaningful variance. Current AI systems predominantly treat demographic diversity as a threat to statistical power (sampling bias) rather than recognizing that homogeneous sampling introduces systematic validity threats when neurobiological heterogeneity underlies surface-level homogeneity.[16,37] The critical distinction: When identical symptom presentations reflect completely different neurobiological mechanisms across populations, studying only homogeneous samples produces invalid conclusions about mechanisms, not merely limited generalizability.

The application of fairness-aware machine learning techniques to protocol development ensures that study designs do not inadvertently introduce or perpetuate disparities across different population groups. These approaches explicitly consider equity and representation in sample selection, outcome measurement, and analysis planning. Tools adapted from AI Fairness 360 and similar platforms provide automated bias auditing for research protocols with recommendations for bias mitigation strategies.

However, political climate with federal restrictions on diversity-related research, "fairness-aware" AI systems face impossible contradictions: recommend equitable sampling and risk political unfundability, or optimize for current funding climates and perpetuate systemic exclusions.[9,25] The framing of diversity as an uncontroversial priority ignores political realities shaping what research becomes possible.

5. Real-world case studies and implementation examples

5.1. NIH BRAIN initiative protocol development

The National Institutes of Health BRAIN Initiative represents one of the largest implementations of AI-assisted protocol development in neuroscience research. Institutional documentation claims the initiative has developed AI systems that synthesise decades of experimental data to assist in drafting and refining multi-centre neuroscience protocols.[49] However, no independent peer-reviewed publications documenting system performance or validation have been identified (Evidence Tier: 2-Moderate; institutional report without external validation). These systems reportedly integrate literature analysis, experimental design optimisation, and resource allocation planning to support large-scale collaborative research efforts.

The AI system developed for the BRAIN Initiative combines natural language processing for literature synthesis with machine learning models trained on historical neuroscience research data. This integration reportedly enables automated identification of optimal experimental conditions, sample size requirements, and analytical approaches based on research objectives and available resources.[21,42,50] Claims of development time reduction lack independent verification; claim predates the AI system implementation (2015 versus 2021–2024 implementation), raising methodological concerns about evidence validity.

5.2. COVID-19 protocol automation at stanford

Stanford University's response to the COVID-19 pandemic included the development of AI-powered tools for rapid protocol development and modification in response to changing pandemic conditions. The system utilised natural language processing to continuously monitor emerging COVID-19 literature and automatically suggest protocol modifications based on new evidence.[51,52] This capability has been described in independent academic sources (Evidence Tier: 1-High; peer-reviewed validation), though quantitative effectiveness metrics and comparative data remain limited. This approach enabled rapid adaptation of research protocols as understanding of the virus and effective interventions evolved.

The NLP-based system developed at Stanford could process hundreds of new COVID-19 research papers daily, extracting key methodological insights and identifying implications for ongoing studies.[53] This detail provides peer-reviewed documentation of system capabilities (Evidence Tier: 1-High). Machine learning algorithms analysed patterns in successful COVID-19 studies to recommend optimal study designs for new research questions. This approach enabled researchers to maintain high-quality protocols while responding rapidly to urgent research needs.

However, "successful studies" during pandemic conditions may reflect publication speed and institutional resources rather than methodological quality or population representation. AI optimization based on rapid publication metrics may systematically disadvantage more rigorous or inclusive approaches that require longer timelines.[15,42]

5.3. Pharmaceutical industry applications

Several major pharmaceutical companies have implemented AI systems for clinical trial protocol development, with claimed improvements in trial success rates and development timelines. Novartis presented findings at the AWS Life Sciences Symposium describing an AI platform that integrates patient data analysis, regulatory requirement checking, and protocol optimisation to streamline clinical development programs.[54] This evidence derives from a conference presentation (Evidence Tier: 3-Low-Moderate; conference presentation without peer review or independent validation). The system reportedly analyses historical trial data to predict optimal patient selection criteria and endpoint selection strategies.

Pfizer's corporate communications claim the company's AI-assisted protocol development system focuses on adaptive trial design, using machine learning to optimise interim analysis plans and stopping rules based on accumulating trial data.[55] The system has been credited with reducing clinical development timelines by an average of 18 months while maintaining regulatory compliance and scientific rigor. However, this claim derives exclusively from a corporate press release (Evidence

Tier: 4-Low; no peer review, no independent validation, potential commercial bias).

Evidence for pharmaceutical AI implementations comes primarily from corporate sources (press releases,[55] conference presentations[54]) rather than peer-reviewed independent evaluation. Commercial interests in demonstrating AI value may bias outcome reporting. Independent verification of claimed efficiency gains is notably absent.[21,42] The 18-month reduction claim lacks methodological documentation, comparative data, or independent audit (Evidence Tier: 4-Low). When corporate promotional materials constitute the evidence base, critical evaluation is essential.

Furthermore, pharmaceutical AI optimization prioritizes commercially viable indications and patient populations with high probability of regulatory approval. This corporate logic may systematically disadvantage rare diseases, complex heterogeneous populations, or novel therapeutic approaches with higher uncertainty.[10,13] Efficiency optimization in commercial contexts encodes profit maximization, not necessarily scientific advancement or public health benefit.

Table 4 synthesises real-world case studies from major organisations including the NIH BRAIN Initiative, Stanford University, Novartis, and Pfizer, documenting key outcomes, implementation timelines, and critical lessons learned from practical AI deployment in research protocol development with explicit attention to evidence quality and verification gaps.

6. Challenges and limitations

6.1. Data quality and availability

The effectiveness of AI systems in research protocol development is fundamentally dependent on the quality and comprehensiveness of training data. Many AI applications require large datasets of completed research studies with detailed methodological information and outcome data, which may not be readily available across all research domains.[21] The variability in reporting standards and data accessibility across different fields creates challenges for developing robust AI models that can generalise across diverse research contexts.

Historical bias in research literature presents another significant challenge, as AI systems trained on existing studies may perpetuate past methodological limitations or demographic biases. The underrepresentation of certain populations in published research can lead to AI recommendations that are not appropriate for diverse study populations.[8,16] Addressing these limitations requires careful curation of training data and ongoing validation of AI recommendations across different contexts.

More fundamentally, when training data predominantly reflects homogeneous sampling justified by statistical power considerations, AI systems learn to encode these practices as "best practice" optimization. The result: AI recommendations perpetuate systematic exclusions while appearing to provide neutral technical guidance.[24,25] Data quality issues are not merely technical problems but reflect epistemological biases about whose experiences constitute valid knowledge.

6.2. Interpretability and transparency

The "black box" nature of many machine learning algorithms poses significant challenges for research protocol development, where transparency and justification of methodological decisions are essential. Researchers, institutional review boards, and regulatory agencies need to understand the basis for AI-generated recommendations to properly evaluate their appropriateness.[24] This requirement has led to increased focus on explainable AI approaches, but these often come with trade-offs in terms of accuracy or computational efficiency.

The complexity of research protocol development, with its multiple interconnected components and domain-specific considerations, makes it difficult to develop AI systems that can provide clear, understandable explanations for their recommendations. Current explainable AI techniques may be insufficient for capturing the nuanced reasoning required for high-quality protocol development decisions.[32,46]

However, the deeper opacity concerns corporate and political interests encoded in AI systems. Even when AI provides technical "explanations," understanding whose agendas the system serves, what commercial pressures shape recommendations, and how political restrictions become encoded as methodological constraints often remains fundamentally opaque.[10,21] Transparency about algorithmic processes does not address transparency about power structures shaping AI deployment.

6.3. Regulatory and institutional barriers

The adoption of AI-assisted protocol development faces significant regulatory and institutional barriers, particularly in highly regulated fields such as clinical research. Regulatory agencies may be unfamiliar with AI technologies and hesitant to accept protocols developed with AI assistance without extensive validation and documentation.[46] This uncertainty can slow adoption and create additional validation requirements that may offset some of the efficiency gains from AI implementation.

Institutional review boards and ethics committees may also require additional training and guidance to effectively evaluate protocols developed with AI assistance. The need to assess both the appropriateness of the research protocol and the validity of the AI system used in its development creates additional complexity for ethics review

Table 4
Case study summary of AI implementation in research protocol development (with evidence quality assessment).

Organisation	AI Application	Key Outcomes	Implementation Timeline	Lessons Learned	Evidence Quality & Critical Concerns
NIH BRAIN Initiative[49,50]	Multi-centre protocol synthesis and optimisation	Reduction in development time, improved consistency	2021–2024	Importance of domain expertise integration, need for validation frameworks	No independent verification; citation predates described system; evidence comes from institutional claims not peer review; efficiency metrics unsubstantiated
Stanford COVID Research[51–53]	Rapid protocol adaptation based on emerging evidence	Real-time protocol updates, maintained quality under pressure	2020–2023	Value of automated literature monitoring, challenges of rapid implementation	"Successful" defined by publication speed not quality; may disadvantage rigorous approaches; pandemic urgency may not generalize
Novartis Clinical Development[54]	Patient selection and endpoint optimisation	Reported improved trial success rates, reduced development costs	2019-ongoing	Critical role of data quality, regulatory acceptance challenges	Evidence from conference presentation not peer review; commercial bias in outcome reporting; independent verification absent; optimization for profitable indications
Pfizer Adaptive Trials[55]	Adaptive design optimisation and real-time modifications	Claimed 18-month average timeline reduction	2020-ongoing	Need for regulatory education, importance of statistical validation	Evidence from press release not peer review; corporate promotional claims; no independent evaluation; profit-driven optimization may disadvantage non-commercial research

processes.[56]

Furthermore, regulatory frameworks developed in different political contexts may have incompatible requirements, creating challenges when US-based corporate AI platforms become global standards. Questions of scientific sovereignty arise: what happens when international research communities depend on AI tools controlled by entities subject to US federal restrictions or corporate priorities?.[13,42]

6.4. Human-AI interaction and over-reliance

The integration of AI in research protocol development raises concerns about potential over-reliance on AI systems and the corresponding reduction in human expertise and critical thinking skills. There is a risk that researchers may become overly dependent on AI recommendations without developing the skills necessary to critically evaluate and modify these suggestions.[21,24] This over-reliance could lead to reduced methodological innovation and the perpetuation of existing approaches rather than the development of novel methodological solutions.

Effective human-AI collaboration requires researchers to maintain appropriate levels of AI literacy and critical evaluation skills while leveraging the computational capabilities of AI systems. Developing this balance requires ongoing training and education programs that help researchers understand both the capabilities and limitations of AI tools.[21,22]

However, when junior researchers are trained with AI assistance as standard practice, they may never develop independent protocol development expertise. The erosion of human methodological knowledge creates dependencies on corporate AI systems while reducing capacity for critical evaluation of AI recommendations. Furthermore, pressure to conform to AI-optimizable protocols may cause researchers to avoid asking genuinely novel questions that lack clear AI-supported pathways.

6.5. Systematic disadvantaging of certain research paradigms

A fundamental limitation inadequately addressed in existing literature: AI optimization based on historical success systematically disadvantages certain forms of knowledge production.[15,31] Specifically:

1. Paradigm-Challenging Research: When AI recommends based on past methodological patterns, genuinely novel approaches appear "methodologically inconsistent" or "high-risk." Researchers face pressure to make work look conventional to receive AI approval, constraining scientific innovation.[42,45]
2. Understudied Populations: When AI training data over-represents easily accessible populations, protocols for hard-to-reach communities are flagged as "recruitment-infeasible" or "high-variance," perpetuating research inequities under technical optimization.[8,17]
3. Theoretically Complex Designs: Research requiring sophisticated theoretical frameworks or methodological innovation may appear unnecessarily complicated compared to standardized approaches AI recognizes as successful.[32,45]
4. Slow-Developing Knowledge: When efficiency becomes primary value, research domains requiring decades-long observation, incremental theory-building, or patient ethnographic engagement become systematically disadvantaged compared to rapidly publishable designs.[6,7]

The fundamental question AI literature fails to address: What research becomes impossible when AI-optimization becomes gatekeeping? Which forms of knowledge are foreclosed when corporate AI platforms determine what constitutes "methodologically sound" protocol development?.[15,31]

7. Future directions and recommendations

7.1. Human-AI collaborative frameworks

The future of AI in research protocol development lies in developing effective human-AI collaborative frameworks that leverage the complementary strengths of human expertise and artificial intelligence. These frameworks should position AI as a "co-pilot" that augments human capabilities while maintaining human oversight and final decision-making authority.[24,31] Successful collaboration requires clear delineation of roles and responsibilities, with AI systems providing data-driven insights and recommendations while humans provide domain expertise, ethical judgment, and creative problem-solving.

Developing effective collaborative frameworks requires research into optimal human-AI interaction designs, including user interface development, explanation generation, and feedback incorporation mechanisms. These systems should be designed to enhance rather than replace human expertise, providing researchers with powerful tools for protocol development while maintaining their central role in the scientific process.

However, collaborative frameworks must explicitly acknowledge power asymmetries. When corporate AI platforms determine "best practices," human collaborators operate within constraints they did not establish and cannot modify. True collaboration requires transparent governance, diverse stakeholder input in system design, and mechanisms for challenging AI recommendations when they conflict with scientific values.[10,13] Furthermore, frameworks must protect researchers' capacity to pursue questions AI systems flag as "unconventional" or "high-risk," ensuring methodological innovation remains possible.[42,45]

7.2. Interoperable platform development

The development of interoperable AI platforms that integrate multiple aspects of protocol development represents a significant opportunity for improving research efficiency and quality. These platforms should combine literature search and synthesis, protocol drafting, ethical compliance checking, statistical planning, and validation tools in a seamless, user-friendly interface.[14] Integration with existing research infrastructure, including institutional systems and regulatory databases, is essential for widespread adoption.

Standardised data formats and API development will be crucial for enabling interoperability between different AI tools and research systems. This standardisation will allow researchers to combine the best features of different AI systems while maintaining data consistency and reducing implementation complexity.[42]

However, interoperability must address questions of control and sovereignty. When standardization means adopting US corporate platform specifications, international research communities cede autonomy to commercial entities whose priorities may shift with political or economic pressures.[10,13] Interoperable platforms should include open-source alternatives, multi-jurisdictional governance, and mechanisms preventing single-vendor lock-in.[21,42]

7.3. Training and capacity building

The successful integration of AI in research protocol development requires comprehensive training and capacity building programs for researchers, institutional staff, and regulatory personnel. These programs should cover both technical aspects of AI systems and the methodological considerations for their appropriate use in research contexts.[46] Training should emphasise critical evaluation skills, enabling researchers to effectively assess and modify AI-generated recommendations.

Professional development programs should be developed in collaboration with academic institutions, professional societies, and AI

technology providers to ensure comprehensive and up-to-date training content. These programs should address both current AI capabilities and emerging technologies, preparing researchers for continued technological advancement in the field.[21]

However, training must extend beyond technical operation to include critical analysis of power structures, epistemological biases, and political-economic contexts shaping AI deployment.[10,46] Researchers need capacity to recognize when efficiency optimization systematically disadvantages certain knowledge forms, to identify corporate or political agendas encoded in AI recommendations, and to maintain methodological autonomy despite pressures toward AI-conforming designs.[22,24] Training should explicitly address: How to recognize when AI flags novel approaches as "high-risk" due to lack of precedent rather than genuine methodological problems. How to evaluate whether "optimization" serves scientific advancement or corporate/political interests. How to maintain capacity for paradigm-challenging research when AI systems encode conservative epistemologies.[31,45]

7.4. Standardisation and quality assurance

The development of standardised approaches to AI validation and quality assurance in research protocol development is essential for ensuring reliability and regulatory acceptance. These standards should address training data quality, model validation procedures, explanation requirements, and ongoing monitoring protocols.[2] Industry-wide adoption of these standards will facilitate regulatory acceptance and improve confidence in AI-assisted protocol development.

Quality assurance frameworks should include both technical validation of AI systems and assessment of their impact on research quality and outcomes. Long-term studies evaluating the effectiveness of AI-assisted protocol development in improving research outcomes will be essential for demonstrating the value of these technologies.

However, quality assurance must address whose values define "quality." When standards derive from corporate AI developers or well-resourced institutions, they may encode particular epistemological assumptions about what constitutes good science.[2,13] Quality frameworks should explicitly evaluate: Whether AI systems accommodate paradigm-challenging research or systematically favor conventional approaches. Whether optimization systematically disadvantages understudied populations or theoretically complex designs. Whether efficiency metrics reflect scientific value or commercial viability. Whether corporate claims of effectiveness are independently verified.[10,46]

Quality assurance must include independent evaluation rather than accepting corporate promotional claims as evidence. The predominance of press releases and conference presentations over peer-reviewed assessment in current AI implementation literature represents a fundamental quality failure that standardization efforts must address.[21,42]

7.5. Addressing scientific sovereignty and political-economic constraints

A critical future direction absent from existing AI literature: developing frameworks that maintain scientific autonomy despite corporate platform dominance and political restrictions.[10,13] Specific recommendations:

1. Multi-Jurisdictional AI Governance: International research communities should develop AI systems with distributed governance preventing single-jurisdiction political control or corporate monopoly.[13,42]
2. Open-Source Alternatives: Investment in open-source AI platforms that can be locally controlled and modified according to diverse scientific priorities, reducing dependence on US corporate systems.[21,42]
3. Epistemological Diversity Protection: Quality standards should explicitly value methodological innovation, paradigm-challenging research, and investigation of understudied populations, preventing AI optimization from systematically foreclosing these approaches.[31,45]
4. Political Context Transparency: AI systems should explicitly acknowledge when political restrictions (e.g., federal limitations on climate or diversity research) constrain recommendations, rather than encoding political unfundability as technical optimization.[15,25]
5. Independent Evaluation Requirements: Mandatory peer-reviewed assessment of AI effectiveness, rejecting corporate promotional claims as evidence.[21,42]
6. Researcher Protection: Mechanisms ensuring researchers can pursue questions AI systems flag as "unconventional" without facing systematic disadvantage in funding, publication, or career advancement.[31,45]

7.6. Hospital implementation and governance framework

The deployment of AI-assisted protocol development in hospital research settings requires comprehensive governance frameworks addressing procurement, operations, and ethics/compliance dimensions. Successful hospital implementation must balance innovation potential with patient safety, data sovereignty, and institutional accountability.

7.6.1. Procurement and adoption requirements

Hospital procurement of AI protocol development systems should mandate contractual provisions addressing:

- Data Sovereignty and Localization: Contracts must specify where protocol data, research metadata, and patient information are stored and processed, with particular attention to cross-border data transfer restrictions and compliance with jurisdictional requirements (GDPR, HIPAA, local data protection laws). Hospitals should negotiate data residency guarantees preventing unilateral vendor relocation of processing infrastructure.[10,13]
- Model Update Notifications and Version Control: Vendors must provide advance notification (minimum 30–90 days) before deploying algorithm updates that may alter protocol recommendations, enabling institutional review of changes and assessment of impacts on ongoing studies. Hospitals should require detailed change logs documenting modifications to AI decision logic, training data, or recommendation algorithms.[42,46]
- Audit Logs and Traceability: Systems must maintain comprehensive audit trails documenting all AI-generated recommendations, user modifications, decision rationales, and protocol versions. Audit capabilities should enable retrospective review for regulatory compliance, quality assurance, and identification of systematic biases.[2,10]
- Explainability and Interpretability Standards: Contracts should mandate that AI systems provide interpretable explanations for recommendations at a level comprehensible to institutional review boards, regulatory agencies, and research teams. Black-box systems without explainability mechanisms should be excluded from procurement consideration.[24,46]
- Third-Party Evaluation Rights: Hospitals should retain rights to commission independent technical audits and performance evaluations, with vendor cooperation requirements for providing necessary system access and documentation. Procurement should prioritize vendors willing to undergo external validation.[21,42]

7.6.2. Operational integration and access control

Effective hospital operations require seamless integration with existing research infrastructure while maintaining appropriate access controls:

- Interoperability with Hospital Systems: AI protocol development platforms must integrate with research data warehouses, electronic

health record systems, regulatory databases, and institutional repositories through standardized APIs and data formats. Integration should enable bidirectional data flow while maintaining data integrity and security.[14,42] Responsibility: Chief Information Officer (CIO) / Chief Technology Officer (CTO) in collaboration with Research IT leadership.

- Role-Based Access Control (RBAC): Implementation should employ granular access controls aligned with institutional roles and responsibilities. Principal investigators, research coordinators, statisticians, regulatory specialists, and IRB members should have differentiated access levels reflecting their protocol development functions. RBAC should prevent unauthorized modification while enabling appropriate collaboration.[10] Responsibility: Data Governance Committee in partnership with Research Administration Office.
- Minimal Access Principles: Systems should implement principle of least privilege, granting users only the minimum access necessary for their specific functions. Access to sensitive protocol components (patient-level data, proprietary methodologies, commercially sensitive information) should require explicit justification and approval.[10,46] Responsibility: Information Security Office in coordination with Research Compliance.

7.6.3. *Ethics, compliance, and quality oversight*

Robust ethical oversight mechanisms are essential for responsible AI deployment:

- Clear Labeling of AI Involvement: All protocols developed with AI assistance must include explicit disclosure statements identifying which protocol components received AI input, which recommendations were accepted or rejected, and the rationale for human decisions overriding AI suggestions. This transparency enables appropriate IRB evaluation and maintains researcher accountability.[46,56] Responsibility: Institutional Review Board (IRB) / Research Ethics Committee.
- IRB/Ethics Review Considerations: Institutional review boards require training and decision frameworks for evaluating AI-assisted protocols. Review criteria should address: (1) appropriateness of AI recommendations for the specific research context, (2) potential biases encoded in AI training data, (3) adequacy of human oversight and modification, (4) implications for participant diversity and inclusion, and (5) compliance with evolving regulatory standards for AI in research.[46,56] Responsibility: IRB Chair and Ethics Committee leadership.
- Version Control and Protocol Traceability: Hospitals must implement rigorous version control systems tracking all protocol iterations, AI-generated drafts, human modifications, approval stages, and amendment cycles. Version control should enable reconstruction of decision histories for regulatory audits and quality reviews.[2,10] Responsibility: Research Administration Office and Quality Assurance team.
- Continuous Monitoring and Quality Metrics: Implementation should include ongoing evaluation of AI system performance through metrics assessing recommendation appropriateness, bias detection, protocol quality outcomes, regulatory approval rates, and researcher satisfaction. Regular quality reviews should identify systematic issues requiring intervention.[2,13] Responsibility: Research Quality Committee with input from Data Governance Committee.

Accountability Boundaries: Clear delineation of responsibilities across hospital leadership (CIO/CTO for technical infrastructure), IRB/ethics committees (protocol appropriateness and ethical oversight), research administration (compliance and quality), and data governance committees (access control and data protection) prevents gaps in oversight while avoiding duplicative review processes. Formal governance charters should specify decision-making authority, escalation pathways, and inter-committee coordination mechanisms.[13,46]

8. Limitations of this review

This narrative review has several limitations that should be acknowledged when interpreting its findings and recommendations. First, the rapidly evolving nature of AI technology means that some tools and capabilities discussed may become outdated quickly, while new developments may emerge that are not captured in this review. The review was conducted using literature available through early 2025, and significant advances in AI technology may have occurred since that time. Furthermore, political contexts continue shifting rapidly; federal research restrictions discussed as 2026 realities may evolve further, requiring ongoing critical evaluation.

Second, the focus on published literature and documented case studies may miss important developments in proprietary AI systems developed by pharmaceutical companies or other organisations that have not been publicly disclosed. This limitation may result in an incomplete picture of the current state of AI implementation in research protocol development. More critically, corporate opacity about AI system operation prevents full evaluation of whose interests these systems serve and what agendas they encode.[10,21] The predominance of promotional claims over peer-reviewed evaluation in available evidence represents a fundamental limitation constraining robust assessment.

Third, the narrative review methodology, while appropriate for synthesising diverse types of evidence in a rapidly evolving field, does not provide the systematic evaluation of evidence quality that would be possible with a systematic review approach. The selection and interpretation of studies may be subject to reviewer bias, and important studies may have been inadvertently omitted. As a narrative rather than systematic review, this manuscript did not follow PRISMA guidelines or employ formal systematic screening protocols. Consequently, precise quantification of evidence sources by tier classification was not conducted, and the evidence base should be understood as representative rather than exhaustive. Future systematic reviews employing PRISMA methodology would strengthen evidence synthesis and enable meta-analysis of AI effectiveness where sufficient comparable studies exist.

Fourth, many of the AI applications discussed in this review are still in early stages of development or implementation, with limited long-term evaluation of their effectiveness and impact on research quality. The evidence base for many recommendations is therefore preliminary, and more extensive validation studies will be needed to confirm the benefits and identify potential risks of AI integration in research protocol development. Specifically, claimed efficiency gains lack independent verification, relying on institutional claims or corporate press releases rather than peer-reviewed evaluation.[21,42]

Fifth, the exclusion of non-English publications may limit generalizability of findings to global hospital contexts and multi-jurisdictional governance frameworks. AI implementation challenges and solutions may differ substantially across linguistic, cultural, regulatory, and healthcare system contexts. Non-English literature may document alternative approaches to AI governance, different regulatory frameworks, or region-specific implementation barriers not adequately represented in English-language sources. This limitation particularly affects applicability to hospitals in non-English-speaking jurisdictions and international collaborative research networks requiring multi-jurisdictional governance models.[8,13]

Finally, the review may not adequately address the diversity of research contexts and institutional settings where AI might be implemented. The examples and case studies primarily focus on well-resourced institutions and research contexts, which may not be representative of the broader research community's experiences and challenges with AI adoption. Furthermore, the review's scope limits detailed examination of how AI deployment varies across different political-economic contexts, regulatory environments, and power structures shaping research globally.[8,13]

9. Conclusion

Artificial intelligence is fundamentally transforming research protocol development by offering unprecedented capabilities for literature synthesis, study design optimisation, ethical compliance checking, and bias detection. The integration of AI technologies including natural language processing, machine learning, and expert systems suggests emerging potential for improving the efficiency, accuracy, and methodological rigour of research protocols across diverse scientific domains. Real-world implementations in major research initiatives and pharmaceutical development programs have reported improvements in development timelines, protocol quality, and research outcomes, though independent verification of many claims remains limited and evidence derives primarily from corporate sources rather than peer-reviewed evaluation.

However, the successful integration of AI in research protocol development requires careful attention to challenges including data quality, algorithmic transparency, regulatory acceptance, and appropriate human-AI collaboration. More fundamentally, critical examination reveals that AI deployment occurs within contested political-economic contexts shaped by corporate ownership of major platforms (predominantly US-based), federal restrictions on certain research domains (climate change, diversity-related investigation), and epistemological biases encoded in training data that systematically favor methodologically conservative, commercially viable, and politically uncontroversial research.[10,21] The treatment of AI optimization as politically neutral technical advancement obscures how efficiency metrics may systematically disadvantage paradigm-challenging research, understudied populations, theoretically complex designs, and knowledge production requiring long timelines or methodological innovation.[15,31]

The fundamental questions AI literature has failed to adequately address: Whose interests do corporate AI platforms serve when commercial pressures or political restrictions shift? What research becomes impossible when AI-optimizable protocols become gatekeeping requirements? How does dependence on US-based corporate systems affect scientific sovereignty for international research communities? When efficiency becomes primary value, which forms of knowledge production are foreclosed?.[13,42]

The future of AI in research methodology lies not in replacing human expertise but in creating effective collaborative frameworks where AI augments human capabilities while maintaining essential oversight and critical evaluation. However, collaboration requires explicit acknowledgment of power asymmetries, transparent governance preventing corporate monopoly or single-jurisdiction political control, protection for researchers pursuing questions AI systems flag as "unconventional," and mandatory independent evaluation rather than acceptance of promotional claims.[31,46] Hospital implementation requires comprehensive governance frameworks addressing procurement requirements (data sovereignty, explainability standards, third-party evaluation rights), operational integration (interoperability, access control), and ethics oversight (AI disclosure, IRB training, version control), with clear accountability delineation across institutional leadership. This approach requires ongoing investment in technology development, training and capacity building that includes critical analysis of epistemological biases and political-economic contexts, and the establishment of appropriate quality assurance and validation frameworks that value methodological innovation and paradigm-challenging research alongside efficiency optimization.[2,13]

As AI technology continues to advance, its role in research protocol development will likely expand and evolve. The research community must proactively address the challenges and limitations identified in this review while working to realise the significant potential benefits that AI offers for advancing scientific knowledge and improving research efficiency. However, this proactive engagement must include critical evaluation of whose values AI systems encode, what knowledge forms are systematically marginalized when efficiency becomes primary metric, and how to maintain scientific autonomy despite corporate platform dominance and political restrictions.[10,31] The ultimate goal should be the development of AI systems that serve as intelligent partners in the research process, enhancing human creativity and expertise rather than replacing them, while preventing AI optimization from foreclosing the paradigm-challenging, methodologically innovative, and politically contentious research that often produces the most transformative scientific advances.[42,45]

CRediT authorship contribution statement

Adeyinka Ojo: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Augustus Osborne:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation. **Soladoye Afeez:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Matope Euphemia:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Bamidele Doyin:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Olawade David Bamidele:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Khan JA, Raman AM, Sambamoorthy N, Prashanth KC. *Research Methodology (Methods, Approaches And Techniques)*. SAN International Scientific Publications; 2023.
2. Costa C, Pezzetta R, Toffalini E, et al. Enhancing the quality and reproducibility of research: Preferred Evaluation of Cognitive and Neuropsychological Studies - The PECANS statement for human studies. *Behav Res Methods*. 2025;57(7).
3. Mandasini AA. Improving the Validity of Scientific Research Results through Research Methodology Management. *Kontigensi J Ilm Manaj*. 2022;10(2):439–447.
4. Page MJ, Moher D, Fidler FM, et al. The REPRISE project: protocol for an evaluation of REProducibility and Replicability In Syntheses of Evidence. *Syst Rev*. 2021;10(1).
5. Getz KA, Campo RA. New Benchmarks Characterizing Growth in Protocol Design Complexity. *Ther Innov Regul Sci*. 2018;52(1):22–28.
6. Poonam G, Minara C, Sodzi ST. Interdisciplinary collaboration: driving better outcomes and efficiency. *Int J Qual Health Care*. 2025 Jul;37(3), mzaf069.
7. Morgan SE, Mosser A, Ahn S, et al. Developing and evaluating a team development intervention to support interdisciplinary teams. *J Clin Transl Sci*. 2021;5(1), e166.
8. Marko JGO, Neagu CD, Anand PB. Examining inclusivity: the use of AI and diverse populations in health and social care: a systematic review. *BMC Med Inf Decis Mak*. 2025;25(1).
9. Fares NY, Moore SA, Jammal M. *Mitigating Bias in AI Recruitment: Leveraging LIME for Fair and Transparent Hiring Models. In: Lecture Notes in Networks and Systems*. Cham: Springer Nature Switzerland; 2025:448–464.
10. Cheong BC. Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. *Front Hum Dyn*. 2024 Jul 3;6, 1421273.
11. Hagos DH, Battle R, Rawat DB. Recent Advances in Generative AI and Large Language Models: Current Status, Challenges, and Perspectives. *IEEE Trans Artif Intell*. 2024 Aug 1;1(01):1–21.
12. Thatikonda VK, Golla Y, Mudunuri HRV. Developments in Artificial Intelligence and Machine Learning: Recent Advances and Future Prospects. *ESP Int J Adv Sci Technol*. 2024;2(3).
13. Lekadir K, Frangi AF, Porras AR, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*. 2025;388, e081554.
14. Weerasinghe K, Olawade DB, Teke J, Msiska M, Boussios S. The Impact of AI-Driven Chatbot Assistance on Protocol Development and Clinical Research Engagement: An Implementation Report. *J Pers Med*. 2025;15(7):269.
15. Goyanes M, Lopezosa C, Piñeiro-Naval V. The use of artificial intelligence (AI) in research: a review of author guidelines in leading journals across eight social science disciplines. *Scientometrics*. 2025 Jul;130(7):3725–3741.
16. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLOS Digit Health*. 2022;1(3), e0000022.
17. Vaughan-Johnston TI, Imtiaz F, Patro GA, Shang SX, Fabrigar L, Ji LJ. Recruitment strategies bias sampling and shape replicability. *Personal Soc Psychol Bull*. 2024 Dec 19, 01461672241293504.

18. Min B, Ross H, Sulem E, et al. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Comput Surv*. 2023 Sep 14;56(2):1–40.
19. Fox DM. Transformer-based Language Models - Architectures and Applications: Analyzing transformer-based language models such as BERT, GPT, and T5, and their applications in NLP tasks such as text generation and classification. *J AIAssist Sci Discov*. 2022;2(1):150–164.
20. Behera PK, Jain SJ, Kumar A. Visual exploration of literature using connected papers: A practical approach. *Issues Sci Technol Librariansh*. 2023 Oct 5;(104).
21. Aldoseri A, Khalifa KNA, Hamouda AM. Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. *Appl Sci*. 2023;13(12):7082.
22. AlMakinah R, Goodarzi M, Tok B, Canbaz MA. Mapping artificial intelligence bias: a network-based framework for analysis and mitigation. *AI Ethics*. 2025 Jun;5(3):1995–2014.
23. Valizadeh A, Moassefi M, Nakhostin-Ansari A, et al. Abstract screening using the automated tool Rayyan: results of effectiveness in three diagnostic test accuracy systematic reviews. *BMC Med Res Method*. 2022;22(1).
24. Gichoya JW, Thomas KJ, Celi LA, et al. AI pitfalls and what not to do: Mitigating bias in AI. *Br J Radio*. 2023;96(1150).
25. Ntoufa S. Bias Starts Early in Research. *So Can Ai Appl Com*; 2025. ⟨https://www.appliedclinicaltrialsonline.com/view/bias-starts-early-in-research-ai⟩.
26. Martinović M, Dokic K, Pudić D. Comparative Analysis of Machine Learning Models for Predicting Innovation Outcomes: An Applied AI Approach. *Appl Sci*. 2025;15(7):3636.
27. Nuti G, Jiménez Rugama LA, Cross AI. An explainable Bayesian decision tree algorithm. *Front Appl Math Stat*. 2021 Mar 22;7, 598833.
28. Jensen FV, Nielsen TD. *Information Science and Statistics*. In: *Bayesian Networks and Decision Graphs*. New York, NY: Springer New York; 2007. ⟨https://link.springer.com/chapter/10.1007/978-0-387-68282-2_9⟩.
29. Kaizer AM, Belli HM, Ma Z, et al. Recent innovations in adaptive trial designs: a review of design opportunities in translational research. *J Clin Transl Sci*. 2023 Jan;7(1), e125.
30. Thorlund K, Haggstrom J, Park JJ, Mills EJ. Key design considerations for adaptive clinical trials: a primer for clinicians. *BMJ*. 2018 Mar 8;360.
31. Gomez C, Cho SM, Ke S, Huang CM, Unberath M. Human-AI collaboration is not very collaborative yet: a taxonomy of interaction patterns in AI-assisted decision making from a systematic review. *Front Comput Sci*. 2025;6:1521066.
32. Coupland H, Scheidwasser N, Katsiferis A, et al. Exploring the potential and limitations of deep learning and explainable AI for longitudinal life course analysis. *BMC Public Health*. 2025;25(1).
33. Villasante-Tezanos A, Kuo YF, Kurinec C, Li Y, Yu X. A non-parametric approach to predict the recruitment for randomized clinical trials: an example in elderly inpatient settings. *BMC Med Res Method*. 2024;24(1).
34. Gkioni E, Rius R, Dodd S, Gamble C. A systematic review describes models for recruitment prediction at the design stage of a clinical trial. *J Clin Epidemiol*. 2019;115:141–149.
35. Rajput D, Wang WJ, Chen CC. Evaluation of a decided sample size in machine learning applications. *BMC Bioinforma*. 2023;24(1).
36. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441.
37. Chen YM, Hsiao TH, Lin CH, Fann YC. Unlocking precision medicine: clinical applications of integrating health records, genetics, and immunology through artificial intelligence. *J Biomed Sci*. 2025;32(1).
38. UNC Research. Protocol Builder - UNC Research. UNC Research; 2025. ⟨https://research.unc.edu/systems/protocol-builder/⟩.
39. Realm Bio. TrialPath – REALM Bio. Realm.bio; 2025. ⟨https://www.realm.bio/white-papers/⟩.
40. Cioffi M. *The Role of AI in Predictive Biomarker Patient Matching*. WCG; 2025. ⟨https://www.centerwatch.com/insights/the-role-of-ai-in-predictive-biomarker-patient-matching/⟩.
41. Alotaibi E, Nassif N. Artificial intelligence in environmental monitoring: in-depth analysis. *Discov Artif Intell*. 2024;4(1).
42. Madanchian M, Taherdoost H. The Impact of Artificial Intelligence on Research Efficiency. *Results Eng*. 2025;26, 104743.
43. Stubbs-Richardson M, Brown L, Paul M, Brenner D. Artificial Intelligence Applications for Social Science Research. *Artif Intell Appl Soc Sci Res*; 2023. ⟨https://scholarsjunction.msstate.edu/ssrc-publications/6/⟩.
44. Hou Y, Huang J. Natural language processing for social science research: A comprehensive review. *Chin J Socio*. 2025;11(1).
45. Pal S. A Paradigm Shift in Research: Exploring the Intersection of Artificial Intelligence and Research Methodology. *Int J Innov Res Eng Multidiscip Phys Sci*. 2023;11(3).
46. Akinrinola O, Okoye CC, Ofodile OC, Ugochukwu CE. Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Adv Res Rev*. 2024;18(3):050–058.
47. World Medical Association. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Participants. Wma.net; 2024. ⟨https://www.wma.net/policies-post/wma-declaration-of-helsinki/⟩.
48. Office for Human Research Protections. The Belmont Report. U.S. Department of Health and Human Services; 2024. ⟨https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html⟩.
49. NIH BRAIN Initiative. NIH BRAIN Initiative; 2023. ⟨https://braininitiative.nih.gov/news-events/blog/brain-initiative-alliance-look-back-brain-initiative-2023-and-whats-coming-2024⟩.
50. Jorgenson LA, Newsome WT, Anderson DJ, et al. The BRAIN Initiative: developing technology to catalyse neuroscience discovery. *Philos Trans R Soc B Biol Sci*. 2015;370(1668), 20140164.
51. Alavi A, Bogu GK, Wang M, et al. Real-time alerting system for COVID-19 and other stress events using wearable data. *Nat Med*. 2022;28(1):175–184.
52. Sinn DI, Muppidi S, Miglis MG, Jaradeh S. Autonomic function test during the COVID-19 pandemic: the Stanford experience. *Clin Auton Res*. 2021;31(1):127–129.
53. Raza S, Schwartz B. Entity and relation extraction from clinical case reports of COVID-19: a natural language processing approach. *BMC Med Inf Decis Mak*. 2023;23(1).
54. Hoh N. *AI-powered Clinical Trials. In: Novartis*. AWS Life Sciences Symposium; 2025. ⟨https://pages.awscloud.com/rs/112-TZM-766/images/Novartis_AI-enhanced%20clinical%20trials.pdf?version=1&trk=8fceb922-1796-48ed-904c-81eea783d4e6&sc_channel=el⟩.
55. Pfizer. Artificial Intelligence: on a Mission to Make Clinical Drug Development Faster and Smarter. Pfizer.com; 2021. ⟨https://www.pfizer.com/news/articles/artificial_intelligence_on_a_mission_to_make_clinical_drug_development_faster_and_smarter⟩.
56. Daoudi M. Ethical limits and suggestions for improving the use of AI in scientific research, academic publishing, and the peer review process, based on deontological and consequentialist viewpoints. *Discov Educ*. 2025 Jul 16;4(1):241.