OREGON HEALTH & SCIENCE UNIVERSITY
SCHOOL OF MEDICINE – GRADUATE STUDIES

ETHICAL AND OPERATIONAL CONSIDERATIONS IN CLINICIAN ADOPTION OF PREDICTIVE AI-ENABLED CLINICAL DECISION SUPPORT TOOLS

By

Shanya San

A CAPSTONE PROJECT

Presented to the Department of Medical Informatics and Clinical Epidemiology
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Master of Science

June 16, 2025

OREGON HEALTH & SCIENCE UNIVERSITY
SCHOOL OF MEDICINE – GRADUATE STUDIES

Department of Medical Informatics and Clinical Epidemiology

CERTIFICATE OF APPROVAL

This is to certify that the Master's capstone project of

Shanya San

has been approved

Supervisory Committee:

Eilis Boudreau, MD, Ph.D., Advisor
Department of Medical Informatics and Clinical Epidemiology
Oregon Health & Science University

Eran Klein, MD, Ph.D., Co-Advisor
Department of Neurology
Oregon Health & Science University

ETHICAL AND OPERATIONAL CONSIDERATIONS IN CLINICIAN ADOPTION OF PREDICTIVE AI-ENABLED CLINICAL DECISION SUPPORT TOOLS

Shanya San

Oregon Health & Science University, 2025

ABSTRACT

Predictive artificial intelligence (AI) tools are increasingly incorporated into clinical decision support (CDS) systems, promising improved risk detection and personalized care. However, clinician adoption of these tools remains slow due to ethical and operational challenges. This capstone examines how autonomy, privacy, fairness, and accountability shape both clinician and patient trust and influence the integration of predictive AI-CDS in healthcare settings. The project outlines a qualitative research design to explore clinician perspectives on ethical risks, training needs, and workflow barriers. The analysis emphasizes strategies for building trust through transparency, robust data governance, and clear accountability frameworks. Recommendations include the implementation of AI ethics review boards, the development of communication tools like question prompt lists (QPLs) for clinicians and patients, and user-centered design to facilitate seamless workflow integration. By addressing the ethical and operational needs of both clinicians and patients, healthcare organizations can better harness the potential of predictive AI-CDS tools while upholding high standards of patient care.

Acknowledgement

I would like to extend my deepest gratitude to my advisor and mentor, Dr. Eilis Boudreau, whose guidance, expertise, and continuous support were essential in bringing this capstone to life. The insightful perspectives and thoughtful feedback provided by both you and Dr. Eran Klein significantly enriched this project.

To my peers and colleagues, thank you for your enthusiasm, insightful discussions, and unwavering support throughout my academic journey. Your camaraderie and input have truly made all the difference.

I also extend sincere appreciation to the Oregon Health & Science University and the Department of Medical Informatics and Clinical Epidemiology. Your exceptional learning environment provided the resources and support needed to successfully complete this work and my Masters.

Finally, my heartfelt thanks go to my husband, Saro San. Your patience, love, and endless encouragement have been my constant source of strength and inspiration. I genuinely don't know where I'd be without you—thank you for being my anchor through every step of this journey.

Dedicated to my daughters,

Civana and Catanna San

Table of Contents

Table of Figures

1. Introduction

Artificial intelligence (AI) is rapidly transforming healthcare by revolutionizing diagnostic accuracy, streamlining clinical decision-making, and optimizing patient outcomes. By analyzing extensive datasets, including electronic health records (EHRs), genomic data, and patient-reported outcomes, AI-driven predictive tools uncover critical insights and subtle patterns that traditional computational approaches might overlook (Obermeyer et al., 2019; Topol, 2019). Predictive AI, integrated into clinical decision support (CDS) systems, presents considerable potential to enhance clinical efficiency by flagging at-risk patients for timely interventions, such as early detection of sepsis, prevention of hospital readmissions, and identification of mental health crises (Kelly et al., 2019).

Despite these significant technological advancements, the adoption of predictive AI tools in clinical practice has not kept pace. The hesitation among healthcare professionals is rooted primarily in concerns surrounding ethics, trust, and practical implementation challenges. Clinicians remain wary of the ethical implications of AI technologies, including potential biases, adverse effects on patient autonomy, data privacy issues, and the opaque nature of algorithmic decision-making (Dorr et al., 2023; Gianfrancesco et al., 2018). Among these concerns, clinician trust has emerged as a central factor shaping adoption behavior. Trust mediates how clinicians assess not only the technical performance of predictive AI tools but also their alignment with core clinical and ethical values. Without confidence in the fairness, transparency, and reliability of AI recommendations, healthcare providers are unlikely to fully integrate these tools into daily clinical practice (Panch et al., 2019). Ensuring fairness requires that predictive AI tools are built on representative, unbiased data and supported by robust data governance practices that protect patient information and uphold ethical standards. Transparency in how models' function, clarity about their limitations, and well-defined accountability frameworks for AI-influenced decisions are also essential to reinforcing clinician trust. Without these safeguards, concerns about bias, inequity, and ethical uncertainty remain significant barriers to adoption.

Operational challenges further exacerbate the gap between technological potential and clinical utilization. Workflow disruptions, the complexity of integrating new technologies into established clinical routines, and the necessity for specialized training represent tangible barriers clinicians face in adopting AI-CDS tools (Pumplun et al., 2021; Scott et al., 2024; Shaheen, 2021). Additionally, ambiguity surrounding accountability—both legal and ethical—complicates clinicians' comfort with and confidence in AI-supported decision-making processes (Dorr et al., 2023; Greenes, 2014).

This capstone project addresses these critical challenges by examining the ethical dimensions influencing clinician trust and operational integration of predictive AI in healthcare. Trust is used as the organizing construct through which both ethical and operational issues are explored. In particular, four key domains autonomy, privacy, fairness, and accountability are treated as intervention targets that shape clinician perceptions of AI trustworthiness. These domains reflect concrete clinician concerns: control over decision-making (autonomy); how personal information is collected and used (privacy); whether outputs are unbiased and equitably applied (fairness); and who is held responsible when AI tools underperform or fail (accountability). Specifically, it aims to critically explore ethical concerns related to these domains in predictive AI adoption; develop a comprehensive framework to capture clinician perspectives on trust, transparency, and operational feasibility; present a qualitative research methodology designed to investigate clinicians' ethical apprehensions, trust factors, and operational barriers; and recommend strategies for ethically informed, user-centric design and effective implementation of predictive AI-CDS systems.

This study explores clinicians' primary ethical concerns about predictive AI-CDS tools, examines how transparency, accountability, and bias mitigation affect clinician trust and adoption, and identifies operational barriers like workflow disruptions and training gaps, along with effective strategies to address them. The remainder of this capstone is structured to first present a comprehensive literature review that contextualizes predictive AI technologies within healthcare, highlights ethical considerations, and delineates challenges regarding clinician trust and operational integration. A subsequent section details the conceptual framework underpinning this inquiry, followed by a proposed research methodology aimed at empirically examining these critical dimensions. Finally, the analysis section will synthesize findings related to ethical and operational factors, culminating in actionable recommendations and a forward-looking conclusion identifying opportunities for future research and application.

2. Literature Review

2.1 Predictive AI in Clinical Decision Support

Predictive AI tools integrated within clinical CDS systems leverage machine learning algorithms to predict patient risk and suggest clinical interventions based on comprehensive data inputs, including EHRs, medical imaging, and real-time physiological monitoring (Shortliffe & Cimino, 2014). Across various specialties such as cardiology, psychiatry, and oncology, predictive AI promises to significantly contributes to clinical effectiveness by forecasting hospitalization risks, potential disease progression, and patient deterioration, thereby allowing timely interventions (Kelly et al., 2019). However, the efficacy of predictive AI systems depends not solely on the technical accuracy of algorithms but also heavily on effective clinical integration, acceptance by healthcare providers, and meaningful engagement by users (Topol, 2019). Without sufficient user acceptance, even the most sophisticated AI tools can fall short of their intended clinical benefits.

2.2 Clinician Adoption of CDS Tools

Historically, clinical decision support systems have experienced variable adoption due to barriers such as alert fatigue, limited interoperability with existing healthcare platforms, and uncertainty regarding predictive accuracy (Greenes, 2014). Predictive AI-CDS systems specifically encounter additional skepticism rooted in "black-box" algorithms, which provide recommendations without clear explanations or justifications that clinicians can readily understand. Explainability, or the transparent articulation of how AI models arrive at their recommendations, is thus crucial to fostering clinician trust and acceptance (Pumplun et al., 2021). Clinicians value explainability not only to better understand the tools they are being asked to adopt, but also because they bear responsibility for communicating AI-driven decisions to their patients. In the absence of a clear understanding, clinicians may feel ethically and professionally compromised when attempting to explain or justify AI-influenced care pathways.

Clinician trust is further impacted by the perceived reliability of AI predictions, transparency regarding algorithmic limitations, and the extent to which AI systems align with clinicians' existing workflows and professional judgment. Thus, promoting acceptance requires clear, consistent communication of AI functionalities, limitations, and reliability (Panch et al., 2019).

2.3 Ethical Considerations in Predictive AI

2.3.1 Autonomy

A primary ethical concern surrounding predictive AI-CDS involves effects on patient autonomy. AI systems can strongly influence clinical decision-making and patient care plans, potentially causing clinicians and patients to defer to algorithmic recommendations without fully understanding their basis or implications (Dorr et al., 2023). Ethical practice demands that AI recommendations be transparently communicated to ensure patients retain informed decision-making capabilities. According to Beauchamp and Childress (1994), meaningful informed consent requires not only a voluntary choice but also adequate disclosure of relevant information (Beauchamp & Childress, 1994). In the context of predictive AI, this disclosure includes explaining how the AI functions, its limitations, the degree of uncertainty involved, and its integration into clinical decision-making. Without sufficient disclosure, patients may consent to treatment pathways without fully appreciating the AI's role, thereby undermining their autonomy. The clinician-patient partnership must be maintained, with AI serving as a supplementary tool rather than an authoritative determinant of clinical action (Vayena et al., 2018).

2.3.2 Privacy

Data privacy represents another significant ethical challenge due to the extensive datasets required by predictive AI systems, often collected from multiple healthcare institutions and possibly managed by third-party vendors. Such aggregation raises concerns regarding data security, unauthorized data usage, and compliance with privacy laws, such as the Health Insurance Portability and Accountability Act (HIPAA). Robust measures including advanced data encryption, secure data-sharing protocols, and transparent data usage policies must be systematically employed to preserve confidentiality and patient trust (Beauchamp & Childress, 1994; Dorr et al., 2023).

2.3.3 Fairness

Algorithmic fairness is critical to ensure equitable healthcare outcomes. Predictive AI models can inadvertently preserve or worsen existing biases, particularly when built on datasets reflective of historical disparities in healthcare delivery. For instance, Obermeyer et al. (2019) revealed racial bias in widely used population health management algorithms, where Black patients were significantly disadvantaged in predictive assessments. The algorithm used healthcare costs as a proxy for health status, resulting in Black patients—who historically incurred lower healthcare expenditures despite comparable levels of illness—being less likely to be identified for high-risk care programs compared to White patients. After adjusting the algorithm to predict actual health outcomes instead of healthcare costs, the disparity in risk identification between Black and White patients was substantially reduced (Obermeyer et al., 2019). Achieving fairness requires rigorous data audits, thoughtful dataset curation, continuous monitoring of predictive outcomes, and an iterative refinement process to identify and eliminate biases effectively (Chen et al., 2023).

2.3.4 Accountability

The deployment of predictive AI-CDS raises complex accountability issues, particularly concerning adverse patient outcomes resulting from AI-driven decisions. Determining responsibility, whether resting with clinicians, algorithm developers, healthcare institutions, or shared among multiple stakeholders, is challenging and can significantly influence clinician confidence and willingness to adopt AI tools (Beauchamp & Childress, 1994). Establishing clear frameworks for legal and ethical accountability is thus critical. These frameworks should explicitly define roles and responsibilities, ensuring that AI use supplements clinical judgment without replacing clinician accountability (Ghassemi et al., 2021).

2.3.5 Trust and Transparency

Transparency is foundational to building clinician trust in predictive AI systems. Clinicians who clearly understand algorithmic decision processes and predictive accuracy are far more likely to adopt these technologies into routine practice (Panch et al., 2019). Transparency encompasses implementing explainable AI techniques, providing straightforward algorithmic rationales, and openly discussing AI performance metrics and limitations.

Patient trust in clinicians typically aligns with clinicians' trust in predictive AI systems; hence, clinicians' openness and confidence in AI-based recommendations directly influence patients' willingness to trust those recommendations and feel comfortable with their use in care. Effective clinician-patient communication about AI recommendations reinforces patient confidence, encourages active participation in care decisions, and sustains a therapeutic relationship that leverages AI as a beneficial tool rather than a source of confusion or mistrust (Gianfrancesco et al., 2018).

3. Conceptual Framework

3.1 Ethical Frameworks

Principlism (Beauchamp & Childress, 1994) provides a foundational lens through four core tenets: autonomy, beneficence, justice, and non-maleficence. These principles map onto issues of informed consent (autonomy), improved outcomes (beneficence), fairness (justice), and risk mitigation (non-maleficence).

1. Principle of respect for autonomy emphasizes the importance of informed consent, ensuring that patients retain control over healthcare decisions involving predictive AI tools. These tools must transparently communicate predictive outcomes, enabling patients and clinicians to collaboratively determine the best course of action without undue influence or coercion from automated recommendations.
2. Principle of beneficence underscores the obligation to improve patient outcomes. Predictive AI must demonstrably enhance healthcare quality, accurately identifying risks and providing actionable recommendations. Clinical validation, continuous performance monitoring, and ongoing model refinement ensure these systems deliver tangible patient benefits.
3. Principle of justice addresses fairness and equity, requiring predictive AI systems to avoid perpetuating or exacerbating existing healthcare disparities. Equitable AI models require careful dataset selection, diverse representation, and consistent auditing processes to identify and eliminate biases that could disadvantage marginalized populations.
4. Principle of non-maleficence mandates risk mitigation and harm avoidance. Predictive AI tools must minimize potential harm from incorrect predictions, data privacy breaches, or unintended clinical consequences. Ethical implementation involves robust data security, transparent risk communication, and clear accountability structures for adverse outcomes.

Complementary to principlism, guidelines from authoritative bodies such as the European Commission (2019) further expand ethical considerations. Their framework for trustworthy AI emphasizes human agency, technical robustness, transparency, and accountability. These additional principles ensure that predictive AI maintains a human-centered approach, prioritizes robust technical performance, and provides clear mechanisms for accountability and redress in the event of system failures or ethical breaches.

3.2 Adoption and Acceptance Models

Models of technology adoption, such as the Technology Acceptance Model (TAM) (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003), provide critical insights into predictive AI acceptance within healthcare environments. These models identify perceived usefulness, ease of use, and social influences as core determinants of technology adoption.

In the predictive AI context, perceived usefulness relates to clinicians' perceptions of whether AI-driven insights meaningfully contribute to clinical practice efficiency, effectiveness, and improved patient care outcomes. Ease of use addresses the integration of AI into existing workflows without imposing additional cognitive or operational burdens. Social influence pertains to the extent to which clinicians perceive AI adoption as socially endorsed by peers, institutions, or authoritative entities. Building upon these traditional constructs, recent research highlights trust and ethical alignment as crucial extensions for predictive AI adoption in healthcare (Gianfrancesco et al., 2018). Clinician trust, shaped by transparency, accountability, and reliability, determines comfort in integrating AI recommendations into clinical judgments, while ethical alignment ensures that AI systems align with core professional values, fostering acceptance and sustained use.

3.3 Human-Centered Design Principles

Human-centered design (HCD) principles advocate the continuous and direct involvement of end-users, particularly clinicians, throughout predictive AI development and deployment processes.

Within the healthcare system, applying HCD involves clinician input in algorithm training, ensuring that datasets reflect clinical realities and patient diversity. Interface design processes involve clinicians in developing intuitive user interfaces that complement, rather than disrupt, clinical workflows. Continuous feedback mechanisms are vital to human-centered predictive AI, facilitating iterative refinement based on user experiences, clinical validation, and real-world performance data. Such participatory design approaches can substantially reduce workflow disruption, enhance clinician confidence, and ultimately improve the overall acceptance and clinical utility of predictive AI tools (Pumplun et al., 2021).

Figure 1 presents a conceptual model that synthesizes the three foundational influences on clinician adoption of predictive AI-enabled CDS tools: ethical frameworks, technology acceptance theories, and human-centered design principles. Ethical frameworks, such as Principlism and Trustworthy AI guidelines, establish critical standards for transparency, fairness, and accountability. TAM offer insights into how perceived usefulness and ease of use drive adoption behaviors. Human-centered design ensures that predictive AI systems align with clinicians' workflows, cognitive processes, and communication needs. The convergence of these domains highlights that successful adoption of predictive AI-CDS tools depends on their ethical soundness, operational usability, and perceived trustworthiness. This integrated triadic framework provides the foundation for examining the ethical and operational challenges addressed in the following sections.

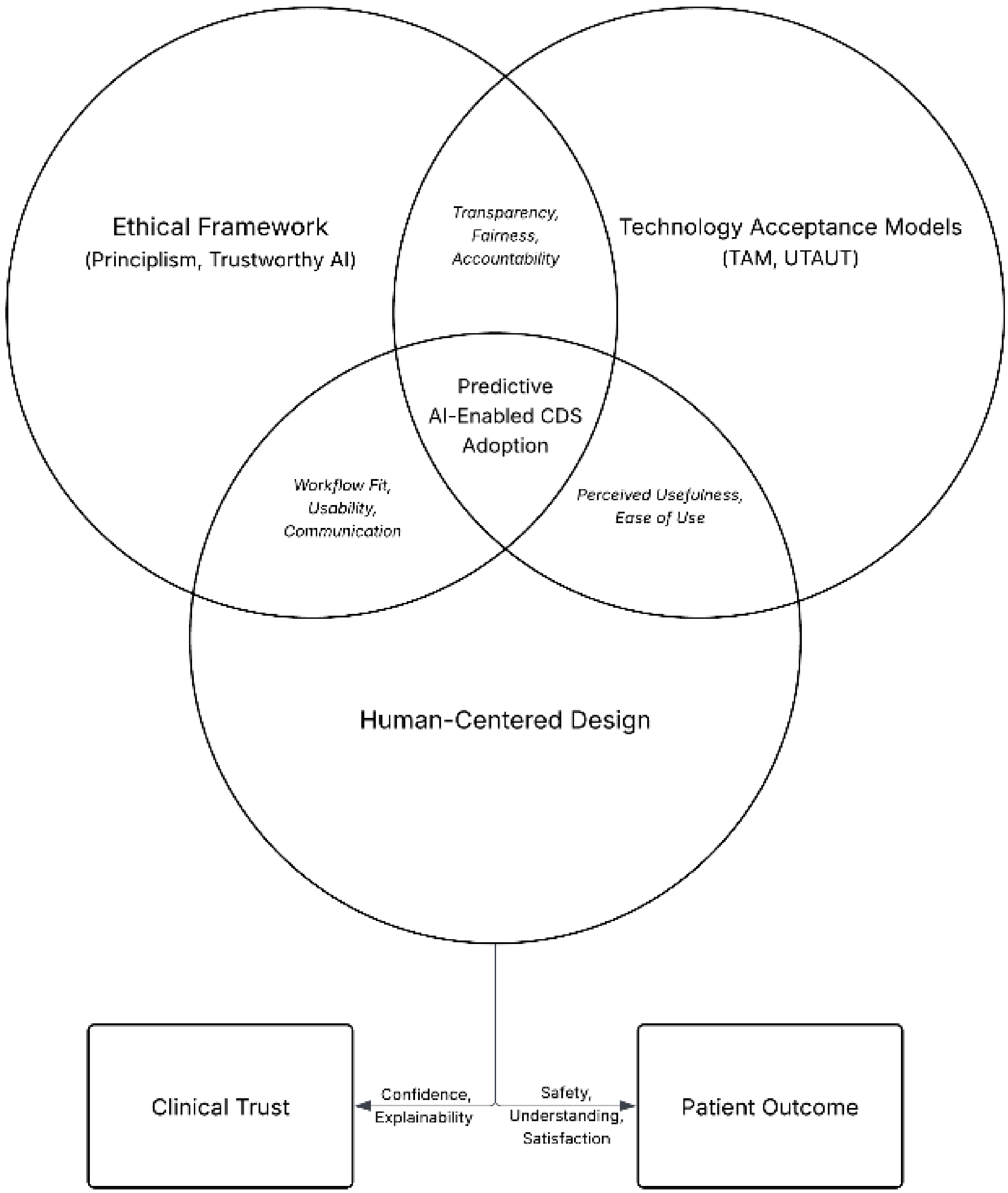

Figure 1 Conceptual Triadic Framework showing the convergence of Ethical Frameworks, Technology Acceptance Models, and Human-Centered Design on Predictive AI-CDS Adoption.

4. Proposed Methodology

4.1 Research Design

This capstone employs a qualitative exploratory research design, a theoretical and empirical approach used to investigate new or poorly understood phenomena by gathering rich, detailed, and subjective data directly from participants (Creswell & Poth, 2016), to examine the ethical and operational considerations surrounding clinician adoption of predictive AI-CDS tools. A qualitative exploratory design is particularly appropriate for this inquiry, as it enables a nuanced examination of clinicians' subjective experiences, perceptions, and ethical apprehensions concerning AI-driven decision-making. This approach allows for a rich, detailed exploration of complex phenomena, especially when addressing areas where empirical research is currently limited or emerging, such as predictive AI in healthcare.

4.2 Research Questions

This qualitative investigation is structured around three central research questions:

1. How do clinicians perceive ethical risks associated with predictive AI-CDS tools?
2. What barriers and facilitators influence clinicians' trust and willingness to adopt these systems?
3. How are autonomy, privacy, fairness, and accountability navigated within clinical settings using predictive AI-CDS?

4.3 Proposed Participants and Sampling

4.3.1 Sampling Method

Purposive sampling will be employed to intentionally select participants who have direct experience and insights relevant to the research objectives. This targeted sampling strategy ensures that collected data directly addresses the outlined research questions.

4.3.2 Participant Criteria

Participants will include licensed clinicians such as physicians, nurse practitioners, and physician assistants. Criteria for inclusion are:

- Active licensure and clinical practice
- Experience using or piloting predictive AI-CDS tools in healthcare settings
- Representation from various healthcare environments, including primary care, behavioral health, and specialty clinics, to ensure diverse perspectives and contextual richness

4.3.3 Sample Size

The sample size will consist of approximately 15–20 participants, a number commonly sufficient to achieve thematic saturation in qualitative studies. Saturation will be assessed iteratively to determine when additional interviews yield no significantly new themes or insights.

4.4 Data Collection Methods

4.4.1 Semi-structured Interviews

Semi-structured, open-ended interviews will be the primary method for data collection, providing flexibility to explore individual clinician experiences deeply. Interviews will be conducted either in person or via secure video conferencing platforms, ensuring participant comfort and confidentiality. The interview guide will include open-ended questions exploring clinicians' ethical concerns, perceptions of trust, experiences with training, and the integration of predictive AI tools within their workflows.

4.4.2 Focus Groups (Optional)

A facilitated focus group using conversation prompts modeled on the semi-structured guide described in Section 4.4.1 may be conducted to supplement individual interviews. The focus group will be conducted in person or via secure video conferencing, depending on participant availability and institutional guidelines. The discussion will last approximately 60–90 minutes and will be audio-recorded and professionally transcribed. The focus group conversation guide will include open-ended prompts designed to elicit collective perspectives on trust, ethical concerns, and specifically for clinicians the workflow integration related to predictive AI-CDS tools. Emphasis will be placed on capturing shared concerns, diverging views, and interpersonal dynamics that may influence attitudes toward adoption. Data from the focus group will be analyzed alongside interview data to enrich thematic interpretation and provide additional context regarding group norms and clinician discourse.

4.4.3 Document Review (Supplementary)

Supplementary document analysis will involve reviewing institutional training materials, predictive AI-CDS user manuals, and policy statements related to AI implementation in clinical practice. This method will offer contextual insights into institutional expectations, guidelines, and the formal framing of AI ethics and operational integration strategies.

4.5 Data Analysis Approach

4.5.1 Thematic Analysis

Thematic analysis, guided by the principles outlined by Braun and Clarke (2006), will be employed for analyzing qualitative data. This approach involves systematically identifying, analyzing, and interpreting patterns within the data related to autonomy, privacy, fairness, and accountability, as informed by ethical frameworks (Principlism) and adoption models TAM/UTAUT.

Coding will be iterative, progressing from initial codes representing broad ideas to refined categories and finally cohesive themes that capture clinicians' perceptions and experiences comprehensively. Qualitative analysis software, such as NVivo, will be used to manage and organize data efficiently, facilitating rigorous analysis and transparent documentation of coding processes and emerging insights.

4.6 Ethical Considerations

4.6.1 Informed Consent

Participants will be provided with a clear and comprehensive explanation of the study objectives, procedures, risks, benefits, and their rights as voluntary participants. Written informed consent will be obtained prior to participation, ensuring that clinicians are fully aware of how their data will be utilized, stored, and reported.

4.6.2 Data Protection

Robust measures will be in place to ensure adherence to relevant privacy regulations, including HIPAA (United States) and GDPR (European Union). This will include anonymizing/deidentifying data, securely storing records, and limiting data access to authorized research personnel only.

4.7 IRB Approval

The research protocol, focus group and interview guides, consent forms, and data protection measures will be submitted to relevant Institutional Review Boards (IRB) for approval prior to data collection. This ensures that the study adheres to the highest ethical standards and is conducted responsibly and ethically.

Through this robust qualitative methodology, this capstone will comprehensively address the ethical and operational considerations critical to clinician adoption of predictive AI-CDS systems, thereby informing future strategies for ethically sound and practically feasible AI implementation in healthcare settings.

5. Ethical Dimensions: Analysis and Application

5.1 Autonomy

Respect for autonomy represents a cornerstone of ethical clinical practice, focusing on patients' rights to make informed decisions regarding their own healthcare. The integration of predictive AI tools introduces unique challenges to preserving patient autonomy, notably the risk that patients may inadvertently defer to AI-generated recommendations without full understanding (Dorr et al., 2023). Ensuring autonomy requires clinicians to thoughtfully navigate the presentation of AI recommendations, fostering a climate of shared decision-making rather than algorithmic determinism.

Several strategies can effectively preserve or enhance patient autonomy in clinical settings involving predictive AI. First, transparent disclosure about the involvement and role of AI in clinical decision-making is essential. Clinicians should clearly articulate how AI recommendations are generated, the data used in deriving predictions, and potential limitations or biases inherent in the algorithms. Transparency empowers patients to assess AI suggestions critically, facilitating genuinely informed consent.

Additionally, providing clear and understandable rationales for AI-driven recommendations strengthens patient engagement and fosters trust. Clinicians should be trained to translate complex AI outputs into lay-friendly explanations that patients can easily understand. Educational materials, simplified visualizations, and structured communication protocols can significantly enhance patient comprehension, empowering them to actively participate in decisions regarding their healthcare.

5.2 Privacy

Predictive AI systems' reliance on extensive datasets raises critical concerns regarding patient privacy, confidentiality, and consent for secondary data use. AI applications in healthcare often require aggregation of data across multiple sources, heightening the risk of unauthorized access, data misuse, or breaches, thus undermining patient trust (Dorr et al., 2023).

Robust institutional data governance policies are indispensable to addressing these privacy concerns effectively. Such policies must explicitly define data collection, storage, usage, and sharing practices, particularly when involving third-party vendors or cross-institutional collaborations. Comprehensive data security frameworks, incorporating encryption technologies and strict access controls, form the foundation for securing patient data.

Regular privacy audits are crucial components of maintaining compliance and trust. Institutions should perform periodic audits to evaluate adherence to privacy regulations like HIPAA and GDPR. Robust consent protocols must be established in the form of transparent consent forms that informs patients about potential secondary uses of their data. Additionally, these forms should clarify how their information is utilized, specify the purpose, and identify the parties involved.

5.3 Fairness

Fairness within predictive AI involves equitable representation in datasets and the continuous monitoring of algorithm performance to identify and rectify biases that disproportionately affect vulnerable populations (Obermeyer et al., 2019). AI systems that fail fairness criteria may exacerbate existing health disparities by systematically disadvantaging certain demographic groups. Ensuring fairness begins with robust and representative training data that adequately reflects patient diversity across demographics such as ethnicity, gender, disability, socioeconomic status, and geographic location. Data collection methodologies must proactively include traditionally underrepresented populations, helping prevent biased algorithmic predictions.

Continuous performance monitoring is integral to fairness. Health systems should conduct systematic algorithmic audits, routinely comparing predictive outcomes across diverse demographic subgroups. Discrepancies or biases identified through these audits should trigger immediate investigation, algorithm recalibration, and, if necessary, retraining using more balanced datasets. Transparent reporting of audit results fosters accountability and demonstrates institutional commitment to fairness.

5.4 Accountability

Ambiguities around accountability in AI-assisted clinical decisions can significantly undermine clinician trust and willingness to adopt predictive technologies (Dorr et al., 2023). Accountability frameworks in predictive AI must clarify the delineation of responsibility across clinicians, developers, and healthcare institutions, particularly in scenarios involving adverse outcomes.

Effective accountability requires clear documentation protocols outlining how clinicians should record AI-assisted recommendations and associated clinical actions. This includes explicit documentation of AI-generated advice, clinical reasoning behind acceptance or rejection of recommendations, and steps taken to mitigate potential risks. Such rigorous documentation practices facilitate transparency and retrospective analyses in case of adverse outcomes. Institutional oversight structures, such as designated AI oversight committees, offer robust mechanisms to ensure algorithm accountability. These committees can monitor ongoing AI performance, address ethical dilemmas or performance concerns, and provide guidance on liability issues arising from AI use. Additionally, clearly defined accountability policies detailing how to manage AI-related errors, including reporting processes and error mitigation protocols, provide clinicians with confidence and clarity in navigating AI integration.

Comprehensive training and education programs are essential to ensure clinicians understand their roles within AI-driven accountability frameworks. Educational initiatives should focus on responsibilities related to AI usage, risks associated with reliance on predictive tools, and ethical considerations, thereby equipping clinicians to confidently and responsibly integrate predictive AI into clinical workflows.

6. Operational Considerations

6.1 Workflow Integration

The successful adoption of predictive AI-CDS systems heavily depends on their seamless integration into existing clinical workflows. Clinicians frequently operate under considerable time constraints and are required to navigate multiple demands associated with EHRs. Predictive AI tools that interrupt or complicate routine workflows can inadvertently reduce clinical efficiency, detract from patient care quality, and generate clinician frustration or resistance (Greenes, 2014).

To address these concerns, predictive AI-CDS systems must be thoughtfully designed with user-centric principles that prioritize seamless integration into existing clinical workflows. Ideally, these tools should integrate directly into existing EHR platforms to minimize workflow disruptions. For example, the Epic Sepsis Model (ESM) integrates sepsis risk scores and alerts directly into Epic Systems' patient dashboards, allowing clinicians to access predictive information without exiting their standard clinical workflow (Wong et al., 2021). Prospective validation studies have shown that, while the predictive performance of ESM may vary, embedding alerts within familiar EHR environments reduces workflow fragmentation and supports clinician awareness. Similarly, Sepsis Watch, developed at Duke University Health System, uses deep learning to monitor patients in real-time and delivers risk alerts to clinicians through the existing EHR infrastructure. Rather than requiring clinicians to log into separate systems, Sepsis Watch assigns predictive risk scores to patients directly within the clinical workflow, enabling earlier recognition and intervention without additional cognitive load (Sendak et al., 2020). Embedding predictive alerts within familiar clinical interfaces, rather than relying on external modules, significantly reduces cognitive burden, preserves clinical efficiency, and fosters greater acceptance and adoption of AI-CDS tools. This aligns with principles from both Human-Centered Design and Technology Acceptance Models, where workflow fits and perceived ease of use converge to support adoption and minimize disruption. Interface design should prioritize clarity and conciseness, limiting unnecessary information or intrusive notifications that can lead to alert fatigue. Moreover, incorporating usability testing involving active clinician participation during the design phase can help identify and mitigate workflow impediments. Real-world scenario testing, cognitive walkthroughs, and clinician feedback during early implementation stages provide critical insights, allowing developers to refine interfaces, streamline interactions, and enhance overall user acceptance. These iterative adjustments, based on clinician input, including how such adjustments might affect patient communication and trust, can further support the ethical and interpersonal dimensions of AI adoption. This reflects the intersection of usability, fairness, and explainability in the triadic framework, ensuring that workflow design decisions are not only efficient but ethically responsive.

6.2 Training and Support

An essential operational consideration for predictive AI-CDS adoption is the provision of robust training and support systems for clinicians. Many healthcare professionals lack formal education or training in AI fundamentals, algorithm interpretability, or AI system limitations, resulting in skepticism, mistrust, or improper application of AI recommendations (Scott et al., 2024; Shaheen, 2021).

To address this gap, healthcare institutions can implement comprehensive educational initiatives, including on-site workshops, continuing medical education (CME) modules, or simulation-based training programs focused explicitly on AI literacy. Where possible, AI education should be integrated into existing mandatory training programs—such as patient safety training, clinical decision-making modules, or annual compliance updates—to minimize additional burden on clinicians' schedules. These training efforts should cover critical topics such as understanding predictive algorithms, recognizing algorithmic bias, accurately interpreting AI-generated recommendations, and effectively communicating AI-informed decisions to patients. Integrating these topics into standard training environments also reinforces the connection between technical literacy and the ethical imperative to ensure informed, autonomous decision-making by patients.

Simulation-based training programs offer a particularly effective avenue for clinicians to practice real-world application scenarios in a controlled environment, reinforcing proper usage and decision-making. Moreover, training should incorporate troubleshooting strategies to address common data quality issues or interpretive ambiguities, empowering clinicians to maintain confidence and accuracy when employing predictive AI-CDS tools. When clinicians are quipped with both technical understanding and ethical communication strategies, they are better positioned to foster patient trust and collaborative decision-making. This supports the central goal shown in the triadic framework: aligning usability, ethics, and acceptance to product explainability, understanding, and ultimately, improved patient outcomes.

6.3 Usability and Feedback Mechanisms

Effective feedback mechanisms are critical to enhancing the usability and efficacy of predictive AI-CDS systems. Establishing a collaborative feedback loop allows clinicians to report inaccurate, confusing, or clinically inappropriate AI recommendations promptly. This active feedback not only contributes to continuous improvement but also fosters a culture of collaboration and mutual learning between AI developers and healthcare providers (Pumplun et al., 2021). Institutions can implement structured feedback processes, such as dedicated online reporting platforms, direct feedback channels embedded within EHR systems, or periodic clinician surveys to systematically capture user experiences. Regularly scheduled feedback meetings between clinicians and developers can also facilitate nuanced discussions on clinical challenges, algorithm performance issues, and potential system enhancements. Embedding ethical and human-centered design considerations into these feedback loops, such as asking how predictive outputs affect patient conversations or perceived fairness, can further refine how AI is used in practice.

This iterative approach ensures that predictive AI systems evolve dynamically, remaining responsive to clinical realities, user experiences, and shifts in the patient population. By inviting feedback on both usability and ethical implications, institutions reinforce their commitment to transparency, fairness, and accountability, which are core elements of clinician and patient trust as highlighted in the triadic framework. By actively engaging clinicians in the feedback process, predictive AI-CDS tools can continually improve their clinical relevance, accuracy, and overall acceptance among healthcare providers.

6.4 Evaluation and Improvement

Continuous monitoring and evaluation are fundamental to the sustained success and ethical deployment of predictive AI-CDS systems in healthcare. Systematic assessment enables healthcare organizations to ensure AI tools consistently meet clinical, operational, and ethical objectives while providing measurable value to patient care and outcomes (Panch et al., 2019).

Key metrics for evaluation include clinician adoption rates, the accuracy and relevance of AI-generated recommendations, clinician satisfaction, patient outcomes, and overall impact on clinical workflows. Tracking these indicators through structured data collection, such as usage analytics, clinician surveys, and patient outcome data, provides actionable insights into the system's effectiveness and identifies areas requiring further improvement. Organizations should regularly review evaluation results through multidisciplinary oversight committees tasked with AI governance. These committees can interpret evaluation findings, recommend system refinements, and oversee the implementation of improvements to address identified deficiencies or clinician concerns. Transparency in evaluation processes and results is crucial for maintaining clinician trust and encouraging sustained engagement. When clinicians understand how their experiences contribute to broader efforts to improve fairness, accountability, and usability, they are more likely to support ongoing refinement of AI systems.

Moreover, incorporating longitudinal studies and comparative effectiveness research can offer deeper insights into predictive AI-CDS tools' long-term impacts, particularly concerning patient care quality, healthcare efficiency, and ethical adherence. Institutions that prioritize rigorous evaluation and transparent communication of results foster an environment of continuous learning and improvement, essential for maintaining clinician engagement, patient trust, and operational excellence.

7. Recommendations

7.1 For Policy and Governance

7.1.1 AI Ethics Review Boards/AI Governance Boards

Hospitals and health systems should proactively establish specialized AI Ethics Review Boards tasked with evaluating new AI tools before their deployment. These committees should include multidisciplinary stakeholders, such as clinicians, ethicists, data scientists, patient representatives, and legal advisors. The boards' primary responsibilities include scrutinizing AI systems for ethical compliance, fairness, transparency, and adherence to privacy regulations, such as HIPAA and GDPR. Detailed evaluations should assess how AI recommendations impact patient autonomy, data privacy, algorithmic fairness, and accountability structures. While the structure of such governance bodies may vary across institutions, examples like the AI Governance Board at Oregon Health & Science University (OHSU) demonstrate how early adopters are addressing oversight needs in practice. More broadly, national-level guidance emphasized the importance of multidisciplinary governance frameworks rooted in transparency, accountability, and public trust. Such governance bodies can significantly enhance organizational oversight, foster ethical transparency, and reinforce clinician and patient trust in AI systems (Dorr et al., 2023).

7.1.2 Auditing Standards

Establishing comprehensive algorithm auditing standards is critical to ensuring the continued fairness and effectiveness of predictive AI tools across diverse patient populations. Health systems should mandate regular audits evaluating algorithmic predictions to identify and correct any biases or disparities systematically. Audits should analyze AI outputs by demographic factors such as race, gender, age, and socioeconomic status. These processes should be embedded within internal governance structures and informed by ethical principles such as transparency, accountability, and equity (Dorr et al., 2023).

Auditability must be a core design requirement of AI-CDS tools, allowing institutions to inspect model outputs and underlying logic to ensure that systems remain responsive to clinical realities and do not reinforce structural inequities. Results from these audits must be transparently reported to clinicians and institutional leadership, enabling timely remediation of any identified issues. In line with national guidance, healthy systems should also consider sharing audit findings publicly when feasible, as part of broader commitments to public trust and responsible AI stewardship (Dorr et al., 2023). Institutions should also consider third-party audits or independent evaluations to provide additional assurance of algorithm reliability and ethical compliance.

7.2 For Clinical Practice

7.2.1 Question Prompt Lists (QPLs)

A common challenge in the adoption of predictive AI in clinical care is the communication gap between clinicians and patients, particularly when explaining how AI tools inform decision-making. Many clinicians are unsure how to describe the AI's role, while patients may be unaware that AI was involved at all. Question Prompt Lists (QPLs) address this gap by equipping both parties with language and structure to guide these conversations, thereby promoting transparency, trust, and patient engagement.

Healthcare institutions should consider developing comprehensive QPLs tailored separately for clinicians and patients to enhance communication, ethical transparency, and trust in the use of predictive AI in clinical practice as demonstrated in **Appendices A-D,** interchangeably. Clinician-facing QPLs can serve as structured tools to guide discussions around the application and interpretation of AI-generated recommendations, emphasizing key concerns such as appropriate contexts for AI use, potential biases, interpretive challenges, and ethical considerations (see **Appendix B** for a Sample Clinician QPL). These prompts enable clinicians to critically engage with AI outputs, facilitate transparent communication, and maintain professional judgement when integrating AI into patient care.

Patient-facing QPLs should provide accessible, patient-friendly prompts that motivate patients to ask questions about AI's role in their clinical care (see **Appendix C** for a Sample Patient QPL). Prompts should focus on the transparency of AI predictions, the reliability of the tools, and patients' rights to be informed and involved in decision-making. These QPLs help reinforce patient autonomy by facilitating shared decision-making and ensuring that AI involvement is fully disclosed and understood within clinical encounters. By encouraging more equitable dialogue between clinicians and patients, QPLs may also reduce power irregularities that can inhibit open discussion about new and unfamiliar technologies.

In addition to communication tools, structured interview guides (see **Appendix A**) can support broader research efforts to understand clinician perceptions, ethical concerns, and operational barriers when adopting predictive AI tools. Regular use of these structured tools can help institutions identify ethical gaps early and support continuous improvement in AI integration practices.

Finally, institutions should implement predictive AI ethics evaluation frameworks, such as the Predictive AI Ethics Evaluation Checklist (see **Appendix D**), to ensure that AI-CDS tools meet ethical standards around autonomy, privacy, fairness, accountability, and reliability before and during clinical use.

By systematically incorporating QPLs, structured interviews, and ethical evaluation checklists, healthcare organizations can strengthen clinician and patient trust, promote ethical AI adoption, and foster more transparent and accountable AI-driven clinical environments.

7.2.2 Disclosure Protocols

Clear and standardized disclosure protocols are essential to maintaining ethical standards in AI-supported clinical practice. Disclosure guidelines should specify the types of AI used, purposes of their recommendations, data sources employed, and the inherent limitations of these systems. Protocols should emphasize patient-friendly language, ensuring comprehension and empowering informed consent. Healthcare providers must be trained regularly to effectively implement these protocols, reinforcing ethical standards and patient autonomy within clinical encounters (Dorr et al., 2023).

Clear disclosure is a foundational element of meaningful informed consent, ensuring that patients have the necessary information to autonomously evaluate AI-supported clinical recommendations (Beauchamp & Childress, 1994). Effective disclosure protocols should prioritize the use of patient-friendly language that explains complex AI concepts in a clear, understandable manner, avoiding technical jargon that could obscure and, as a result, undermine meaningful informed consent. Training programs should be developed to help clinicians consistently apply these protocols across different clinical contexts, ensuring that disclosure is not treated as optional but as a fundamental ethical obligation. Such training would reinforce clinician accountability and empower patients to make informed choices regarding AI-assisted clinical recommendations.

Integrating disclosure protocols alongside QPLs and ethical evaluation frameworks strengthens the broader institutional commitment to autonomy, transparency, fairness, and accountability, which are essential pillars for trustworthy AI adoption in healthcare settings.

7.3 For System Designers

7.3.1 Co-Design Processes

Designers of predictive AI-CDS systems must actively involve frontline clinicians throughout iterative design and development phases. Engaging clinicians early in co-design processes ensures systems align effectively with real-world clinical workflows, usability expectations, and clinical judgment processes. Regular clinician feedback sessions, workflow simulations, and iterative design adjustments based on clinician inputs enhance practical utility, minimize workflow disruptions, and increase clinician acceptance and satisfaction.

7.3.2 Explainable Interfaces

System designers should prioritize developing intuitive, explainable interfaces that enhance clinicians' understanding of AI-driven recommendations. User-friendly dashboards, clear visualizations, and interactive components should articulate the underlying algorithmic logic, data sources utilized, and prediction reliability. Enhancing system explainability through these visual tools reduces uncertainty, supports informed clinical decision-making, and bolsters clinician trust and confidence in integrating AI into routine practice.

7.4 For Training and Education

7.4.1 Simulation-Based Learning

Healthcare institutions should implement simulation-based training programs designed to enhance clinicians' practical understanding of predictive AI-CDS tools. Through case-based scenarios, these simulations can effectively demonstrate AI system capabilities, limitations, interpretive challenges, and associated ethical dilemmas. Simulated clinical environments enable clinicians to practice integrating AI recommendations into clinical decision-making safely, reinforcing proper usage, critical thinking, and ethical awareness regarding predictive AI systems.

7.4.2 Continuing Medical Education (CME) Modules

Institutions should offer structured continuing medical education modules focused explicitly on AI literacy, including fundamentals of machine learning, interpretability, limitations, and ethical and regulatory compliance. Regular CME modules should update clinicians on evolving AI technologies, newly identified ethical considerations, changes in regulatory landscapes, and best practices for integrating predictive AI into clinical practice. These educational initiatives promote ongoing professional development, equip clinicians to navigate ethical challenges, and ensure informed, effective utilization of AI-CDS tools in healthcare (Scott et al., 2024; Shaheen, 2021).

These recommendations collectively foster robust ethical governance, enhance clinical integration, support user-centric system design, and prioritize comprehensive clinician training, promoting responsible, effective, and ethical implementation of predictive AI-CDS in healthcare.

8. Conclusion and Future Research

8.1 Summary of Key Themes

This capstone has systematically examined critical ethical and operational considerations that significantly influence clinician adoption and integration of predictive AI-CDS systems. Key ethical considerations encompass autonomy, privacy, fairness, and accountability, each demanding specific strategies for ethical compliance and effective management.

Autonomy emphasizes empowering patients to actively engage in healthcare decisions informed by transparent, comprehensible disclosures about AI involvement. Privacy demands stringent data governance protocols to maintain patient confidentiality, regulatory compliance, and trust. Ensuring fairness involves rigorous monitoring and auditing of algorithms to mitigate bias and promote equitable healthcare delivery. Lastly, accountability frameworks must explicitly define responsibilities and processes to address potential adverse outcomes, thereby sustaining both patient and clinician trust and effective AI integration.

Operationally, this analysis highlights critical elements such as comprehensive clinician training, human-centered AI system design, and seamless workflow integration. These elements collectively enable clinicians to integrate AI-driven tools effectively without significant disruptions, ensuring that AI complements rather than complicates clinical practice.

8.2 Emphasis on Clinician-Centered, Ethically Grounded Implementation

For predictive AI-CDS to achieve its transformative potential within healthcare, systems must be demonstrably trustworthy, transparent, and user-friendly for clinicians. Practical guidelines outlining appropriate AI usage, continuous monitoring and auditing for algorithmic fairness, and active clinician involvement throughout the design and refinement processes are crucial. A clinician-centered approach ensures AI systems effectively support rather than supplant clinical judgment, sustaining the patient-clinician relationship's integrity and quality of care.

Clear policies and standardized disclosure protocols foster transparency, while structured feedback loops and usability testing further align AI tools with clinicians' real-world experiences and expectations. Rigorous education and training programs are essential to cultivating clinicians' AI literacy, empowering informed, confident utilization of predictive AI-CDS. By emphasizing ethically grounded practices, healthcare institutions can foster broader acceptance and more meaningful integration of AI technologies in clinical environments.

9. Future Research Directions

9.1 Patient Perspectives

Future research should prioritize investigating patient perceptions regarding predictive AI-generated recommendations. Understanding how patients interpret AI involvement in their care, the implications for trust in healthcare providers, and the subsequent effects on shared decision-making dynamics are critical areas warranting further exploration. Qualitative studies, including patient interviews and focus groups, can yield invaluable insights into patient concerns, informational needs, and conditions necessary for effective patient engagement with AI-driven recommendations. As part of this future work, researchers should explore the development and implementation of patient-facing QPLs as a potential tool to support these conversations. Evaluating how QPLs influence patient understanding, comfort with AI involvement, and trust in clinician recommendations will be essential to refining communication strategies in AI-assisted care.

9.2 Real-World Outcome Evaluations

Longitudinal, real-world outcome evaluations are essential for comprehensively assessing the clinical effectiveness, cost-efficiency, and patient satisfaction impacts of predictive AI-CDS implementation. Such studies should span diverse healthcare contexts, incorporating quantitative metrics (clinical outcomes, operational efficiency, error rates, financial impact) alongside qualitative assessments (clinician and patient satisfaction, perceived value of AI contributions). Systematic collection of longitudinal data can guide future refinements, policy developments, and targeted training initiatives.

9.3 Comparative Analysis

Comparative analyses of various predictive AI-CDS tools across different healthcare settings constitute another critical area for future research. Investigating differences in ethical compliance, operational performance, clinical effectiveness, and user acceptance among distinct AI systems can illuminate best practices, inform policy guidelines, and drive more targeted improvements. Such analyses should consider diverse organizational structures, clinical specialties, and patient populations, providing broad-based insights into optimal conditions and strategies for AI integration.

In conclusion, addressing the intersection of ethical principles with real-world operational challenges enables healthcare organizations to deploy predictive AI-CDS tools more effectively, ensuring enhanced clinical outcomes and sustained patient trust and autonomy. Through continued research and commitment to clinician-centered, ethically sound practices, predictive AI can fulfill its potential as a transformative force in healthcare.

Reference

Beauchamp, T. L., & Childress, J. F. (1994). *Principles of biomedical ethics*. Edicoes Loyola.

Chen, R. J., Wang, J. J., Williamson, D. F., Chen, T. Y., Lipkova, J., Lu, M. Y., Sahai, S., & Mahmood, F. (2023). Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nature Biomedical Engineering*, *7*(6), 719–742.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

Dorr, D. A., Adams, L., & Embí, P. (2023). Harnessing the Promise of Artificial Intelligence Responsibly. *JAMA*, *329*(16), 1347–1348. https://doi.org/10.1001/jama.2023.2771

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, *3*(11), e745–e750.

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Internal Medicine*, *178*(11), 1544–1547.

Greenes, R. (2014). *Clinical decision support: The road to broad adoption*. Academic Press.

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*, 1–9.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453.

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: Implications for health systems. *Journal of Global Health*, *9*(2), 020318.

Pumplun, L., Fecho, M., Wahl, N., Peters, F., & Buxmann, P. (2021). Adoption of machine learning systems for medical diagnostics in clinics: Qualitative interview study. *Journal of Medical Internet Research*, *23*(10), e29301.

Scott, I., Vegt, A., Lane, P., McPhail, S., & Magrabi, F. (2024). Achieving large-scale clinician adoption of AI-enabled decision support. *BMJ Health & Care Informatics*, *31*, e100971. https://doi.org/10.1136/bmjhci-2023-100971

Sendak, M. P., D’Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A path for translation of machine learning products into healthcare delivery. *EMJ Innov*, *10*, 19–00172.

Shaheen, M. Y. (2021). Adoption of machine learning for medical diagnosis. *ScienceOpen Preprints*.

Shortliffe, E. H., & Cimino, J. J. (2014). *Biomedical Informatics Computer Applications in Health Care and Biomedicine*.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56.

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine*, *15*(11), e1002689.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425–478.

Wong, A., Otles, E., Donnelly, J. P., Krumm, A., McCullough, J., DeTroyer-Cooley, O., Pestrue, J., Phillips, M., Konye, J., & Penoza, C. (2021). External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Internal Medicine*, *181*(8), 1065–1070.

Appendix A: Sample Interview / Focus Group Guide

This guide is intended for use in qualitative interviews or focus groups with clinicians to explore their perceptions, experiences, and concerns related to the use of predictive AI clinical decision support (CDS) tools in healthcare settings.

1. Opening Questions

- What is your current understanding or experience with predictive AI-CDS tools?
- Can you describe a time when you encountered or used such a tool in your practice?

2. Ethical Concerns

- Have you encountered any situations where AI-driven recommendations conflicted with patient autonomy or privacy norms?
- Do you feel that predictive AI tools are fair to all patient populations? Why or why not?
- Who do you think should be accountable when an AI-generated recommendation leads to a poor outcome?

3. Trust Factors

- What factors increase your trust in AI-generated predictions?
- What factors make you skeptical or hesitant to rely on them?
- How do you determine when to override or follow an AI-based recommendation?

4. Operational Workflow

- How have predictive AI tools impacted your clinical workflow or time management?
- Have these tools made decision-making more efficient, more complex, or both?
- Are there any barriers to integrating these tools into your everyday practice?

5. Training Gaps

- What training or educational resources have you received regarding the use of predictive AI-CDS tools?
- What additional support, training, or materials would help you feel more confident using them?
- Would you be comfortable explaining how a predictive AI tool works to a patient? Why or why not?

Appendix B: Sample Clinician Question Prompt List (QPL) for AI Conversations

This Question Prompt List (QPL) is intended to help clinicians critically engage with predictive AI tools used in healthcare. These prompts can guide conversations with developers, administrators, or ethics teams, and prepare clinicians to address patient questions with clarity and confidence.

Understanding the Model

1. How was the AI model trained, and how reliable are its predictions in clinical settings like mine?
2. What kind of clinical data (e.g., lab results, clinical notes, demographics) does the AI use — structured, unstructured, or both?
3. Is the model validated on a patient population similar to the one I treat?

Ethical and Bias Considerations

4. Are there known biases or limitations in the model, especially across different demographic groups? e.g., based on race, age, gender
5. Does the AI tool consider social determinants of health, and could that impact fairness?
6. Has the model undergone bias auditing or external ethical review?

Clinical Integration

7. How is this tool integrated into my workflow, and what are its limitations during real-time care?
8. What happens if I override the AI recommendation — is that tracked, penalized, or reviewed?
9. Does the AI provide explanations or just predictions? Will explanations be provided to help me understand why the tool made a given recommendation?

Patient Communication

10. Will guidance be provided to help me explain the AI-driven recommendations to my patient in a clear, non-technical way?
11. Should I disclose when an AI tool is influencing clinical decisions? If so, how and when?

12. What patient concerns might arise about AI involvement, and how should I respond to them?

Accountability and Legal Considerations

13. Who is accountable if an AI-generated prediction leads to a poor outcome — me, my institution, or the vendor?
14. Are there institutional policies or legal protections in place regarding the use of AI-CDS tools?

Training and Support

15. What training is available to help me understand how this predictive tool works and how to use it appropriately?
16. Is there technical support or a feedback channel if I encounter errors or ethical dilemmas?

Appendix C: Sample Patient Question Prompt List (QPL) for Conversations Involving Predictive AI in Care

This Question Prompt List is designed to help patients engage in meaningful conversations with their healthcare providers when artificial intelligence-enable (AI) tools are used to support medical decisions. Patients can use these questions to better understand how AI influences their care and ensure that their values and preferences are respected.

> "**Patient Question Prompt List (QPL): Talking About AI in Your Healthcare**
>
> *You have the right to understand how decisions about your care are made. If your doctor is using an AI-enabled (Artificial Intelligence) tool to help guide decisions, you can use the questions below to ask your provider to learn more about your care and make sure your voice is heard."*

Understanding the Role of AI

1. Is an AI or computer program being used to help make decisions about my care?
2. What is the AI-enabled tool predicting about my health or condition?
3. How accurate is this prediction, and how does it compare to what a doctor might decide without AI?

Knowing What's Behind the Prediction

4. What kind of information did the AI-enabled tool use to make this prediction?
5. Is the AI-enabled tool using data from people like me (my age, gender, background, or condition)?
6. Can you explain how the AI came to this conclusion in a way I can understand?

Ethical and Personal Concerns

7. Are there any risks or downsides to using AI in my case?
8. Is my personal health information safe when AI tools are used?
9. Will the final decision be made by a person or a computer?

Your Preferences and Involvement

10. Can I choose not to have AI involved in decisions about my care?
11. What are my other options if I'm not comfortable with what the AI recommends?
12. How will you make sure my values and preferences are still part of the decision-making process?

Appendix D: Predictive AI Ethics Evaluation Checklist

This checklist provides a quick-reference tool for clinicians, developers, or institutional review teams to evaluate whether a predictive AI-enabled tool aligns with key ethical principles before implementation or use in practice.

Autonomy & Transparency

- ☐ Are clinicians informed when AI is influencing a decision?
- ☐ Are patients notified or given a chance to ask about AI's role in their care?
- ☐ Is the tool's purpose, function, and limitations clearly documented?

Privacy & Data Use

- ☐ Does the tool comply with HIPAA and local data protection laws?
- ☐ Is patient data anonymized or de-identified before model training?
- ☐ Are there clear data governance and consent policies?

Fairness & Bias Mitigation

- ☐ Has the model been tested for demographic bias (e.g., race, gender, age)?
- ☐ Are underrepresented populations adequately included in the training data?
- ☐ Has the tool undergone a bias audit or fairness impact assessment?

Accountability & Governance

- ☐ Is there a clear chain of responsibility when AI is used in patient care?
- ☐ Are decisions tracked to determine whether clinicians follow or override AI recommendations?
- ☐ Is there a process for clinicians to report concerns or adverse outcomes related to the AI?

Reliability & Clinical Validity

- ☐ Has the tool been externally validated in a setting similar to its intended use?

- Are performance metrics (e.g., area under the curve (AUC), positive predictive value (PPV), sensitivity) publicly available or shared with end users?
- Is the model regularly updated to reflect new data and reduce drift?

Training & Support

- Are clinicians given sufficient training on how to interpret and use the AI tool?
- Is there documentation or support available for technical and ethical questions?
- Is explainability built in to help clinicians understand predictions?