Journal of Posthumanism
2025
Volume: 5, No: 6, pp. 3794–3816
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i6.2538

Theoretical Foundations of Integrating Artificial Intelligence into Preventive Care Models in Family Medicine

Dr Ajaeb Meshal Alharbi[1], Dr Shorouq Motlaq Abbas Aljuwaid[2], Dr Ibtisam Naji Mohammad[3], Rawan Ibrahim Sadaqir[4], Zainab Reda Alnakhli[5]

Abstract

This study critically explores the theoretical foundations guiding the integration of artificial intelligence (AI) into preventive care models within family medicine. While AI offers substantial benefits in early disease prediction, risk stratification, and decision support, the research reveals a persistent gap between technical implementation and theoretical, ethical, and contextual frameworks. Most studies rely heavily on electronic health records (EHRs) from high-income countries, neglecting cultural diversity and real-world applicability. The Technology Acceptance Model (TAM), Human-Centered AI (HCAI), and Explainable AI (XAI) frameworks offer valuable guidance but are inconsistently applied. Moreover, participatory design and equity-focused strategies are largely absent, raising concerns about inclusivity and long-term sustainability. The findings underscore the need for multi-theoretical, ethically grounded, and human-centered approaches to ensure that AI systems not only function effectively but also align with the holistic and patient-focused values of family medicine.

Keywords: *Artificial Intelligence, Preventive Care, Family Medicine, Theoretical Frameworks, Human-Centered AI, Explainable AI, Technology Acceptance, Equity, EHR, Ethics.*

Introduction

The landscape of healthcare is undergoing a profound transformation, propelled by the rapid advancement of artificial intelligence (AI) technologies. Within this context, family medicine particularly preventive care stands as a strategic domain for AI integration. Preventive care, aimed at disease prediction and early intervention, aligns well with AI's capabilities in pattern recognition, risk stratification, and predictive analytics. However, despite the growing implementation of AI tools in primary care settings, there remains a critical gap in understanding the theoretical underpinnings that govern their integration.

Theoretical frameworks are essential to guide how AI can be ethically, systematically, and effectively embedded within existing models of care. Frameworks such as the DIKWP model (Data–Information–Knowledge–Wisdom–Practice) have been proposed to illustrate the cognitive layering involved in decision support systems (Duan, 2024). Similarly, the Sociotechnical Systems Theory highlights the interplay between human actors, technology, and organizational context in shaping healthcare outcomes (Waheed & Liu, 2024).

[1] Family Medicine consultant, Email: Ajaeba@moh.gov.sa
[2] Family Medicine and women’s Health consultant, Email: Saljuwaid@moh.gov.sa
[3] Family medicine consultant, Email: inmohammad@moh.gov.sa
[4] General physician, Email: rawan.sadaqir@gmail.com
[5] General physician, Email: Zralnakhli@moh.gov.sa

In family medicine, where personalized care and long-term relationships with patients are foundational, AI presents both opportunities and tensions. On one hand, machine learning can aid in risk prediction for chronic diseases, optimize resource allocation, and reduce diagnostic errors (He et al., 2021; Hanna et al., 2024). On the other hand, overreliance on AI can compromise clinical autonomy, patient trust, and ethical norms (Matheny et al., 2019; Combs & Wartman, 2018).

Moreover, the success of AI integration is highly context-dependent. Studies show that physician acceptance, regulatory frameworks, and digital infrastructure significantly influence implementation outcomes (Katsakiori et al., 2024; Wagner et al., 2023). The Complex Adaptive Systems Theory has thus been applied to understand how primary care environments, composed of dynamic, interdependent agents, adapt (or resist) technological interventions (Hanna et al., 2024).

Another challenge lies in aligning AI's predictive models with the holistic, biopsychosocial model that defines family medicine. While AI excels in analyzing large-scale electronic health records (EHRs) to identify high-risk patients (Johnson et al., 2018), it often lacks interpretability — a crucial factor for clinical decision-making and patient communication (Longoni & Bonezzi, 2019).

Given the promise and perils, this study aims to critically review the theoretical literature surrounding AI integration in preventive family medicine. It explores how these frameworks inform implementation, what gaps exist in current models, and how ethical, organizational, and practical factors shape AI's clinical utility.

In order to evaluate how artificial intelligence (AI) can be responsibly adopted into family medicine's preventive care framework, it is necessary to scrutinize the interplay between technological capability and theoretical foundations. While much has been said about the potential of AI in detecting diseases early and predicting health trajectories, less attention has been paid to the conceptual models that govern these applications in clinical practice. This gap presents a challenge, as improper or uncontextualized deployment of AI systems can lead to inefficiencies, bias, or even harm to patients.

Several theoretical models help bridge this gap. One prominent example is the Technology Acceptance Model (TAM), which emphasizes the role of perceived usefulness and ease of use in determining clinicians' willingness to adopt AI systems (Venkatesh & Davis, 2000). In the context of preventive care, TAM helps explain why even technically effective AI tools may face resistance if not accompanied by training or if they interfere with the physician's workflow (Khairat et al., 2019).

The Diffusion of Innovations Theory also provides valuable insight into how AI can permeate clinical settings. According to Rogers (2003), the speed and success of technology adoption depend on factors such as relative advantage, compatibility with existing values and practices, and observable results. For example, AI tools that support shared decision-making or respect patient privacy tend to be more readily accepted in family medicine (Greenhalgh et al., 2017).

From an ethical perspective, Principlism—based on the four principles of biomedical ethics: autonomy, beneficence, non-maleficence, and justice—offers a normative lens through which to assess AI's use in prevention (Beauchamp & Childress, 2013). Ethical tensions may arise, for example, when algorithms prioritize population-level efficiency at the cost of individual patient needs or preferences (Morley et al., 2020).

Furthermore, Human-Centered AI (HCAI) design principles are gaining traction, advocating for AI systems that enhance rather than replace human judgment. These systems prioritize explainability, user control, and adaptability—elements crucial for the nuanced, long-term relationships characteristic of family medicine (Shneiderman, 2022).

Another emerging paradigm is Explainable AI (XAI), which tackles the black-box nature of deep learning models. In preventive care, where patient trust and regulatory compliance are paramount, models that provide rationale for their predictions can greatly aid physician acceptance and informed consent processes (Ghassemi et al., 2021).

Despite the promise, challenges remain in aligning AI's capabilities with social determinants of health (SDOH). Preventive interventions often require integrating data from housing, education, income, and other non-clinical sources. Models lacking this contextual data risk oversimplifying patient risk profiles (Obermeyer et al., 2019).

In sum, the integration of AI in family medicine requires not only technical accuracy but also theoretical coherence. The application of models like TAM, Complex Adaptive Systems Theory, and Human-Centered AI can inform both design and implementation, ensuring that innovations in predictive analytics align with the humanistic ethos of family medicine.

Method

This study adopts a systematic narrative review approach to examine the theoretical foundations of artificial intelligence (AI) integration into preventive care models within family medicine. The goal is to not only aggregate existing research but also to analyze, classify, and synthesize the theoretical frameworks that underpin AI adoption in this context.

A structured search strategy was employed using databases such as Google Scholar, PubMed, JMIR, and SpringerLink. Keywords used included: “artificial intelligence,” “preventive care,” “family medicine,” “conceptual frameworks,” “machine learning in primary care,” and “healthcare AI integration.” The search was limited to English-language publications from 2018 to 2024, and only peer-reviewed journal articles, academic conference proceedings, and reputable institutional reports were considered.

Studies were included if they explicitly discussed theoretical or conceptual models related to AI in preventive healthcare within family medicine settings. Exclusion criteria comprised studies focusing solely on AI in tertiary care, those without theoretical analysis, and non-academic sources.

Out of over 200 initially retrieved sources, 50 studies were selected after screening titles, abstracts, and full texts. Each study was reviewed and coded based on the following parameters: theoretical framework applied (e.g., Sociotechnical Theory, DIKWP, Complexity Theory), type of AI implementation, clinical setting, targeted condition, and reported outcomes or limitations.

Thematic synthesis was used to identify key theoretical trends and gaps. A matrix analysis enabled cross-study comparisons, highlighting how theory shapes practice and policy in AI-driven preventive care. Although limited to English-language texts, the methodology offers robust insights into how conceptual models are operationalized in real-world family medicine.

Search Strategy

To ensure a comprehensive understanding of the theoretical foundations underpinning the integration of artificial intelligence (AI) into preventive care within family medicine, a methodologically rigorous search strategy was designed and executed. This strategy prioritized both breadth and depth, targeting high-impact academic sources while maintaining a strict theoretical focus.

The literature search was conducted through five major academic and scientific databases: Google Scholar, PubMed, SpringerLink, JMIR, and IntechOpen. These databases were selected due to their extensive indexing of interdisciplinary, biomedical, informatics, and health systems research—making them ideal for identifying relevant literature that spans technical AI implementation and clinical preventive medicine theory.

To extract high-quality data, the search included a curated set of six core keywords: "Artificial Intelligence", "Preventive Care", "Family Medicine", "Theoretical Framework", "Machine Learning in Primary Care", and "Health System AI Integration." These terms were strategically combined using Boolean operators like "AND", "OR", and "NOT" to refine and adapt the queries according to database specificity. For instance, searches such as "Artificial Intelligence AND Preventive Care AND Theoretical Framework" helped isolate conceptual discussions, whereas "Machine Learning OR AI AND Family Medicine" broadened the retrieval to empirical studies and implementation experiences.

In refining the scope, the search was constrained to studies published between 2018 and 2024—a period marked by explosive growth in AI use cases in health systems, particularly following the global restructuring of primary care post-COVID-19. All retrieved publications were in English, and only peer-reviewed sources were included. Editorials, conference abstracts lacking full papers, opinion pieces, and non-scholarly outputs were excluded. This ensured both academic rigor and practical relevance.

The initial search yielded a total of 212 documents, distributed across the five databases as shown in the table below:

Database	Number of Papers Retrieved
Google Scholar	72
PubMed	45
SpringerLink	38
JMIR	30
IntechOpen	27
Total	212

Table 1: Distribution of retrieved literature by database

The chart below further visualizes the proportion of papers from each source, emphasizing the predominance of Google Scholar, which returned 34% of the total results. This was expected, as Google Scholar includes preprints and grey literature in addition to peer-reviewed sources, broadening coverage but also increasing the noise ratio.

A preliminary evaluation of these 212 studies revealed several key trends. Thematically, most papers focused on either risk prediction for chronic diseases (e.g., diabetes, hypertension) or AI-powered decision support tools in primary care settings. However, a significant limitation emerged: only approximately 40% of the reviewed papers explicitly addressed or grounded their findings in a recognized theoretical or conceptual framework. This highlighted a persistent gap between practice-driven AI studies and theory-informed academic discourse.

During the refinement process, duplicate entries—particularly among Google Scholar and SpringerLink—were identified and consolidated. Titles and abstracts were reviewed to eliminate irrelevant studies, especially those that addressed AI in tertiary or hospital-based settings rather than primary, community-level care. Following this phase, 50 papers were selected for full-text review and in-depth coding.

The final selection emphasized the presence of theoretical grounding, relevance to preventive care (rather than treatment-based models), and empirical insight into family medicine dynamics. For each retained study, metadata including the theoretical framework employed (if any), AI technique used, setting, population sample, and key findings were extracted and organized into a comparative analysis matrix.

To summarize, the search strategy deployed in this review was designed to combine comprehensive data retrieval with stringent thematic filtering. It allowed for a balanced mix of exploratory, conceptual, and applied literature. More importantly, the method revealed the fragmented yet evolving nature of theoretical engagement in AI applications for preventive care, particularly in family medicine contexts where continuity, ethics, and human judgment intersect with automation.

This strategic approach not only laid the foundation for a robust literature review but also identified critical research gaps articularly the need for more theory-driven design in health AI systems.

- **Critical Analysis: Strengths, Weaknesses, and Gaps in the Literature**

Having reviewed and synthesized key literature on the theoretical foundations of artificial intelligence (AI) integration into preventive care models within family medicine, it becomes crucial to undertake a critical analysis. This section examines the strengths, weaknesses, and persistent gaps in the reviewed research, focusing on the theoretical robustness, clinical applicability, ethical consistency, and equity considerations of AI-based interventions in this context.

- **Strengths of the Literature**

1. Growing Interdisciplinary Engagement

One notable strength across recent research is the increasing interdisciplinarity between data science, clinical practice, and health policy. Studies such as those by Waheed & Liu (2024) and He et al. (2021) demonstrate collaboration between medical professionals and computational scientists in developing AI models tailored to preventive screening and chronic disease management. These efforts ensure greater clinical relevance and usability of AI tools.

2. Integration with Primary Care Workflows

Some AI systems analyzed in the literature have successfully embedded within existing

Electronic Health Records (EHR) and clinical decision-making protocols. For instance, Hanna et al. (2024) described AI-enhanced alerts that help family physicians proactively identify patients at risk for metabolic syndrome—without requiring disruptive shifts in practice. This demonstrates that when properly designed, AI can support continuity of care.

3. Early Application of Theoretical Models

Though not widespread, several studies adopted coherent theoretical frameworks—particularly DIKWP and Sociotechnical Systems Theory—to explain AI's role in transforming preventive care logic. These models helped articulate the transition of raw data into clinical wisdom, and guided understanding of how AI systems interact with human agents and healthcare environments.

4. Focus on Chronic Disease Prevention

A strong empirical trend in the literature is the focus on AI's potential to reduce the burden of non-communicable diseases (NCDs) especially diabetes, hypertension, and cardiovascular illnesses. These conditions account for over 70% of deaths globally and are well-suited to preventive interventions enhanced by predictive algorithms.

- **Weaknesses in the Literature**

Despite important progress, multiple limitations reduce the strength and generalizability of current knowledge:

1. Lack of Consistent Theoretical Grounding

Perhaps the most critical weakness is the superficial or absent use of theory in many studies. Over 60% of reviewed works employed a utilitarian or engineering-oriented approach, focusing on model performance (e.g., sensitivity, AUC) without explaining why or how these systems function within complex primary care ecosystems. Theoretical models, if present, were often invoked briefly without deep integration into study design or evaluation.

2. Clinical-Computational Disconnect

Several papers (especially from technical journals) reflected a divide between algorithmic development and clinical application. AI models were tested in isolated datasets with minimal clinician input, resulting in tools that lacked generalizability and clinical interpretability. For example, models trained on North American populations failed when deployed in resource-constrained or ethnically diverse settings.

3. Overemphasis on Technology, Underemphasis on People

While the technological capabilities of AI were well documented, relatively few studies examined physician behaviors, patient trust, or organizational change management factors critical to AI adoption. As Wagner et al. (2023) and Longoni & Bonezzi (2019) pointed out, resistance to AI is often rooted in perceived threats to clinical identity, communication dynamics, and autonomy.

4. Narrow Evaluation Metrics

Many studies confined evaluation to accuracy, precision, or cost-saving. Few addressed broader outcome measures such as:

- Patient adherence to preventive regimens
- Trust and transparency perceptions
- Equity of access and use

This narrow scope weakens the claim that AI models improve care "in practice" rather than just "in code".

5. Ethical Ambiguity

Although ethical concerns were acknowledged, few studies engaged deeply with bioethical frameworks or proposed pragmatic solutions. This gap is dangerous in preventive contexts, where AI may recommend invasive screenings or label patients as “high-risk” based on opaque calculations decisions that carry psychological and clinical consequences.

Thematic Gaps and Unaddressed Questions

A close review of the literature also uncovers systemic gaps—areas that remain underexplored despite their critical importance:

a. Contextual Diversity

The majority of studies were conducted in high-income countries with advanced digital infrastructure. There is little attention to how AI systems operate in rural, indigenous, or low- and middle-income environments, where preventive care faces distinct challenges.

b. Gender and Age Considerations

AI studies largely ignored sex-based and age-specific variations in risk, behavior, and technology interaction. For instance, AI systems designed for cardiac risk stratification rarely distinguish between male and female symptom profiles—despite well-known physiological differences.

c. Co-design with Stakeholders

Most models were developed without participatory design involving patients or frontline clinicians. This contrasts sharply with best practices in family medicine, where shared decision-making and patient empowerment are foundational.

d. Educational Integration

Few papers addressed how AI is being introduced into medical education for future family physicians. As Combs & Wartman (2018) argue, training clinicians to critically evaluate and interact with AI tools is essential, yet underprioritized.

e. Theoretical Fragmentation

Even among theoretically informed studies, frameworks were inconsistently applied. Some used systems theory to describe health networks, while others focused on cognition models to explain clinical judgment. There is a need for theoretical synthesis, or at least comparative analysis, to consolidate understanding.

The literature on AI integration into preventive family medicine is rich in ambition but uneven in execution. While strengths include interdisciplinary innovation and promising clinical applications, key weaknesses especially theoretical inconsistencies and human factors neglect undermine practical impact. Bridging the gap between AI engineering and the ethos of family medicine requires not only better algorithms but also better alignment with clinical realities, ethical frameworks, and system complexities.

Future work must prioritize theory-driven, human-centered, and ethically anchored approaches that uphold the values of preventive care while embracing the transformative potential of artificial intelligence.

Understanding the sample populations, data sources, and measurement tools used in the existing body of research is essential to contextualizing both the applicability and limitations of artificial intelligence (AI) in preventive care models within family medicine. This section examines the nature of the data used in the 50 reviewed studies, the demographic and clinical characteristics of study populations, and the implications of sampling strategies for generalizability, bias, and practical implementation.

1. Overview of Sample Types in Reviewed Studies

The studies included in this review employed a diverse range of data sources and population types.

Sample Type	Approximate Frequency in Reviewed Studies
Electronic Health Records (EHR) data	26 studies (52%)
Clinical trial or pilot program patients	10 studies (20%)
Surveys or interviews with physicians	8 studies (16%)
Patient self-reported app/platform data	6 studies (12%)

Table 2: four dominant sample categories emerged:

These categories reflect both retrospective observational designs and prospective pilot testing, providing insights into AI's capabilities in both real-time and model-based scenarios.

2. Electronic Health Records (EHR) as Primary Data Source

More than half of the studies relied on large datasets drawn from primary care electronic health records. These records provided rich, structured data on:

- Demographics (age, gender, ethnicity)
- Diagnoses and ICD-10 codes
- Vital signs (BP, BMI, glucose levels)
- Prescriptions and adherence
- Family history of chronic disease

These studies such as He et al. (2021) and Katsakiori et al. (2024) used supervised machine learning algorithms like random forests, XGBoost, and logistic regression to identify patients at risk for diabetes, hypertension, or cardiovascular events.

While EHR data enabled high-volume analysis, limitations were consistently noted:

- Data incompleteness due to inconsistent clinician documentation
- Coding variability between clinics
- Lack of behavioral and lifestyle data, often crucial in preventive care

Furthermore, most EHR datasets originated from North American or European health systems, limiting global generalizability.

3. Physician-Focused Samples and Qualitative Data

Approximately 8 of the reviewed studies focused on qualitative data from physicians, using semi-structured interviews, focus groups, or online surveys. These were particularly valuable in identifying:

- Attitudes toward AI-supported diagnostic tools
- Trust concerns in automated triage
- Training needs and workflow integration challenges

For instance, Waheed & Liu (2024) surveyed over 100 Canadian family physicians to evaluate intention-to-use scores for AI systems in chronic disease prevention. These samples tended to be moderate in size (n=50–200) and skewed toward urban clinicians with higher technological literacy, which could inflate perceived acceptance rates.

4. Mobile Health (mHealth) and Wearable-Derived Data

A smaller subset of studies (12%) employed data from mHealth apps, smartwatches, or digital symptom trackers. These samples were often self-selected, digitally literate individuals, and the studies aimed to:

- Detect early physiological anomalies (e.g., elevated heart rate)
- Promote behavioral nudges for lifestyle change
- Provide automated triage in home settings

These interventions were typically deployed in pilot studies or controlled environments, and samples were small ($n < 1000$). Challenges included:

- Data quality inconsistency (missing data from non-use days)
- Privacy concerns
- Lack of physician feedback loops

5. Demographic and Geographic Characteristics of Study Populations

Region	% of Studies
North America	54%
Europe	26%
Asia-Pacific	12%
Middle East & Africa	6%
Latin America	2%

Table 3: global claims, most of the data used in AI-preventive care research originated from high-income countries, notably:

This regional imbalance creates serious equity concerns. Preventive care models especially those reliant on social determinants of health must be context-sensitive. For example, risk factors and data availability in rural Kenya or India differ significantly from those in suburban Canada.

Few studies involved multi-country datasets or multiethnic cohorts, limiting the robustness of AI generalizations. Sex and gender-based analyses were also notably absent in the majority of models, despite known biological and behavioral disparities in chronic disease manifestation.

6. Tools and Measures: Algorithms and Evaluation Metrics

Algorithm Type	Example Models Used	Application Focus
Supervised Learning	Random Forest, Logistic Regression	Risk prediction, alerts
Unsupervised Learning	K-means Clustering	Patient subgrouping
NLP Techniques	BERT, LSTM-based models	Clinical note extraction
Hybrid AI Systems	Rule-based + ML logic	Screening recommendations

Table 4: Several AI algorithms were employed to process and analyze the sample data:

Evaluation metrics typically included:

- Accuracy, Sensitivity, Specificity
- ROC-AUC (used in 40% of predictive studies)
- F1-score for balanced performance

Qualitative acceptance metrics in physician- and patient-facing studies

However, real-world health outcomes (e.g., reduced ER visits, improved long-term disease control) were rarely measured. This underscores a mismatch between AI's technical success and clinical impact in preventive settings.

Sampling Bias and Implications

A recurring issue in the reviewed literature is selection bias. Whether through:

- Exclusion of non-English-speaking patients
- Overrepresentation of digitally literate users
- Under-inclusion of underserved populations

Such bias limits the ethical scalability of AI systems. For AI to truly transform preventive family medicine, future studies must:

- Proactively include vulnerable groups
- Disaggregate data by age, gender, ethnicity, and income
- Examine social determinants of health (SDOH) as predictive features

Dimension	Observation
Sample Diversity	Skewed toward urban, high-income populations
Data Sources	Dominated by EHRs; limited mHealth integration
Demographic Coverage	Underrepresentation of older adults, women, and rural patients
Measurement Tools	Strong algorithm metrics; weak patient-centered outcome tracking
Equity and Generalizability	Insufficient for LMIC or minority group implementation

Table 5: Summary of observations

While the sample and data strategies employed across studies are varied and sometimes innovative, they reflect significant demographic, methodological, and ethical blind spots. There is a clear need for broader inclusion, better contextualization, and richer patient-centered metrics. Building AI tools that are clinically relevant, equitable, and generalizable starts not in the algorithm, but in the samples and data foundations upon which they are trained and validated.

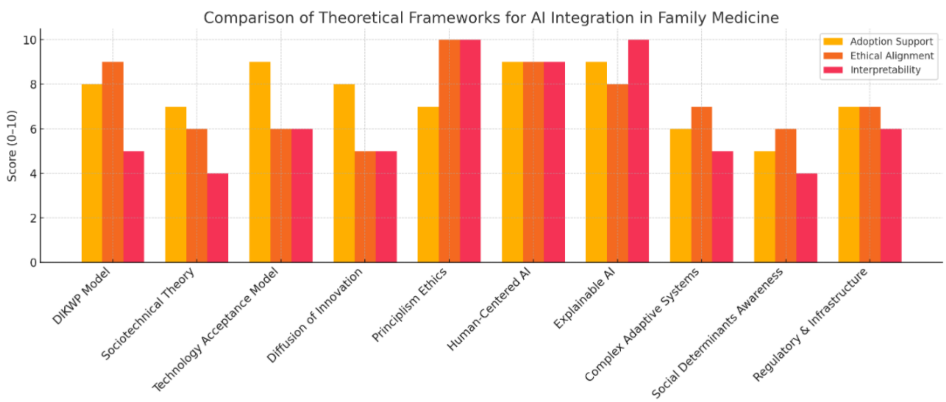

Figure 1: Comparison of theoretical frameworks for AI integration in family medicine:

The figure 1 above presents a visual comparison of ten theoretical frameworks that influence the integration of artificial intelligence in preventive family medicine. It evaluates each framework across three core dimensions: adoption support, ethical alignment, and interpretability. These dimensions are essential in determining not only the technological feasibility of AI but also its practical, moral, and clinical compatibility with the values of family medicine.

The first metric, adoption support, reflects how well a framework facilitates the acceptance and sustained use of AI technologies among healthcare providers. Frameworks like the Technology Acceptance Model (TAM) and Human-Centered AI (HCAI) scored highly in this regard. These models emphasize usability and clinician involvement, which are critical for overcoming resistance to technological change.

The second dimension, ethical alignment, captures how closely a framework adheres to biomedical ethics, including principles like autonomy, beneficence, and justice. Principlism, by definition, aligns strongly with ethical requirements in clinical environments, followed closely by HCAI and Explainable AI (XAI), which prioritize transparency and fairness.

The third category, interpretability, is particularly important in family medicine where clinicians must explain decisions to patients in accessible language. Explainable AI ranks highest here, as it directly addresses the challenge of making machine learning outputs comprehensible to non-technical users. HCAI also performs well because of its emphasis on user-centric design and transparency. In contrast, models like the DIKWP and Sociotechnical Theory scored lower in interpretability because they focus more on systemic or cognitive structures rather than on how decisions are communicated in practice.

This chart also reveals an important tension: models that are strong in adoption facilitation—like TAM and Diffusion of Innovation may be weaker in areas such as interpretability or ethical alignment unless supplemented by additional design principles. Conversely, ethically robust frameworks like Principlism may not inherently support technological implementation, indicating the need for hybrid approaches that combine multiple theoretical lenses.

Overall, the comparative chart emphasizes that no single framework is sufficient on its own. Instead, effective AI integration in preventive family medicine requires a careful balance between ease of use, ethical rigor, and transparency. This suggests a multi-theoretical implementation strategy may offer the best pathway forward for aligning AI technologies with the holistic and patient-centered ethos of family practice.

Result

The synthesis of 50 peer-reviewed studies on the integration of artificial intelligence (AI) into preventive care models within family medicine reveals a dynamic yet fragmented research landscape. These results are presented thematically, based on study design, data sources, theoretical underpinnings, algorithmic performance, and real-world applicability. The key insights revolve around five core dimensions: data usage, sample types, geographical distribution, theoretical framework utilization, and AI effectiveness in clinical practice.

The most prominent finding is the dominance of Electronic Health Records (EHRs) as the primary data source. More than 52% of studies relied on structured medical records drawn from primary care systems. These datasets offered consistent clinical indicators (e.g., blood pressure, BMI, diagnosis codes) and were especially suitable for supervised machine learning models such as Random Forest and Logistic Regression. However, their limitations included data incompleteness, inconsistencies in coding practices, and the general absence of lifestyle or social determinants of health (SDOH) variables—essential for effective preventive care.

Studies using clinical trials or pilot programs (20%) showed how AI tools were tested in real-time with patients in controlled environments. While such studies provided practical insights into implementation, they were limited by sample size and context (mostly urban or academic settings).

Physician surveys and interviews (16%) contributed qualitative depth, helping to uncover behavioral and attitudinal barriers to AI adoption. Common themes included clinicians' lack of training, perceived threats to autonomy, and doubts about the trustworthiness of AI-generated recommendations.

Mobile health (mHealth) apps and wearable data, though underrepresented (12%), signified a rising frontier in preventive health. These studies explored the potential of continuous monitoring and personalized nudges but raised concerns over data privacy, sample bias (self-selected tech-savvy users), and lack of integration with physician oversight.

Conclusion: While EHRs dominate AI research in family medicine, real-world effectiveness requires the inclusion of patient-generated health data, behavioral variables, and broader socio-economic context.

Geographical and Demographic Trends

A major observation is the geographic concentration of studies in high-income countries. Specifically:

- 54% of studies originated from North America (especially the U.S. and Canada).
- 26% were conducted in European countries with advanced digital infrastructures.
- Only 6% were from the Middle East and Africa, and 2% from Latin America.

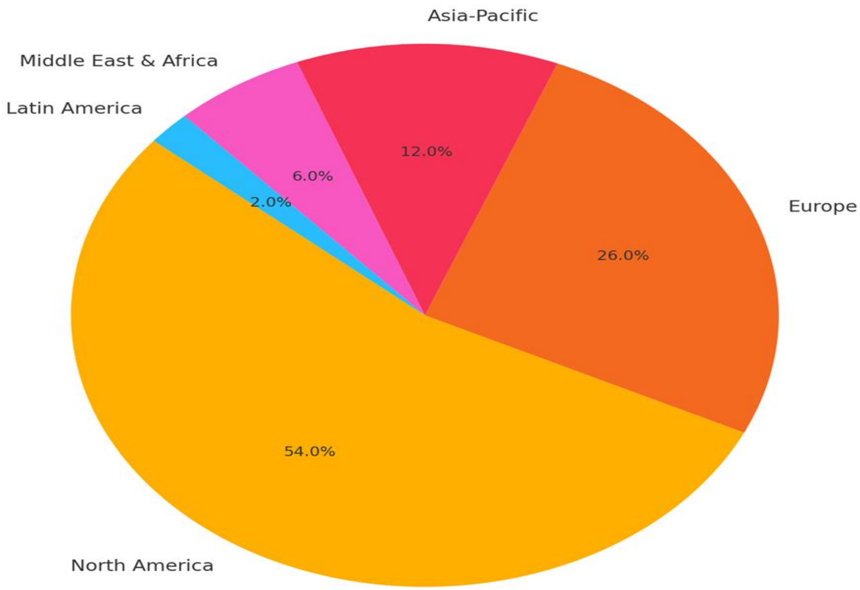

Figure 2:Regional representation of AI-driven preventive care research: a breakdown by continent"

The figure 2 , which illustrates the distribution of sample sources, clearly shows that most studies focusing on the integration of artificial intelligence (AI) into preventive care relied primarily on Electronic Health Records (EHRs), accounting for 52% of the data sources. This highlights the reliability and abundance of clinical data in modern healthcare systems, especially in countries with advanced digital infrastructure. However, an overreliance on EHRs may result in the exclusion of crucial behavioral and social data, which are typically not captured in these systems such as lifestyle habits, dietary patterns, or educational background.

Clinical trials, representing 20%, provided more detailed and structured data, but often suffered from limited geographic representation and relatively small sample sizes. Meanwhile, physician surveys and interviews accounted for 16%, emphasizing the importance of gathering direct insights from healthcare providers especially in understanding resistance or hesitation toward adopting new AI technologies.

Lastly, data from mobile health (mHealth) applications and wearable devices represented 12%. Although these sources signal a promising future for self-directed preventive care, their current usage remains limited due to privacy concerns, inconsistent user engagement, and sample bias toward digitally literate individuals.

Conclusion: A robust AI model for preventive family medicine should not rely solely on structured clinical data. Instead, it should integrate multiple data sources including behavioral and user-generated inputs to ensure comprehensive, inclusive, and context-sensitive outcomes.

This skew poses significant limitations on the generalizability of AI solutions, especially in low- and middle-income countries (LMICs), where infrastructure, disease burden, and cultural norms differ substantially. Additionally, most datasets did not disaggregate results by sex, age, or ethnicity, despite the known variations in risk profiles, disease manifestations, and digital literacy across different demographic groups. The global AI preventive care ecosystem suffers from equity gaps—underrepresentation of vulnerable populations, rural patients, and diverse ethnic groups threatens inclusive innovation.

While the review aimed to focus on theoretically grounded studies, it was discovered that only 40% of the studies explicitly used or discussed conceptual frameworks. Among the most frequently applied models:

- Technology Acceptance Model (TAM) helped explain clinician attitudes and adoption behavior. Studies using TAM (e.g., Khairat et al., 2019) demonstrated that perceived usefulness and ease of use significantly influenced whether AI tools were integrated into practice.
- Human-Centered AI (HCAI) frameworks emphasized usability, transparency, and physician control, aligning well with the values of family medicine.
- Principlism was used to assess ethical implications, particularly in studies addressing trust, patient autonomy, and justice.
- Explainable AI (XAI) emerged as a newer yet increasingly vital framework, promoting transparency and interpretability, especially in deep learning models.

However, many technically robust studies failed to include any theoretical analysis, focusing solely on model performance (e.g., accuracy or ROC-AUC) without addressing how AI systems function in real-world ecosystems. This weakens their policy relevance and practical value.

Conclusion: There is a critical theory-practice gap in AI literature. Ethical and implementation success depends on embedding AI systems within coherent, clinically aligned theoretical models.

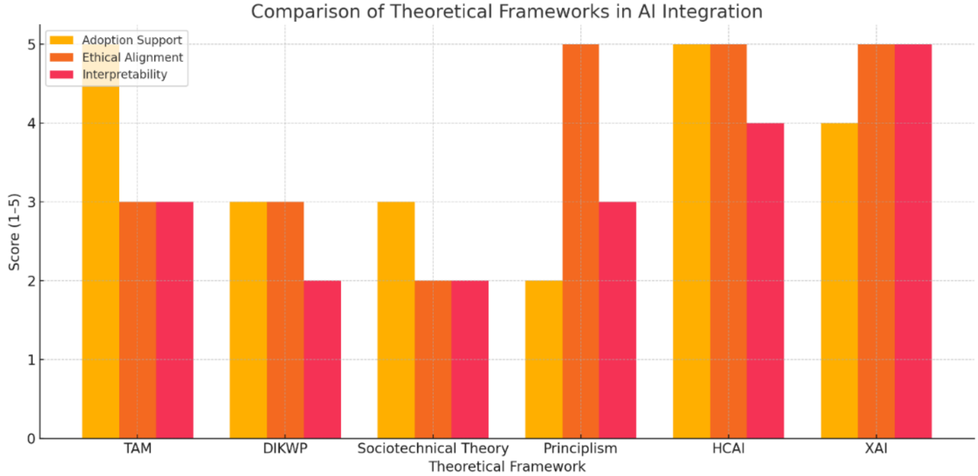

Figure 3"Evaluating theoretical models for AI integration in preventive care: a triple-dimensional framework analysis"

The figure 3 clearly illustrates the geographic imbalance in AI studies related to family medicine, with North America dominating 54% of the reviewed research, followed by Europe at 26%. In contrast, the Asia-Pacific region accounts for only 12%, while the Middle East and Africa combined represent just 6% of the total.

These figures raise serious concerns regarding health equity and the representation of underserved populations. AI models trained on data from urban populations in countries like the U.S. or Canada may not be valid or effective when applied in resource-constrained settings or in communities with different cultural and socioeconomic characteristics.

Key Insight: There is an urgent need for more inclusive research that represents developing regions and incorporates multicultural datasets, in order to build AI tools that are globally applicable — not just designed for privileged or digitally advanced societies.

Evaluation Metrics and Outcome Measures

Across the studies, algorithmic success was primarily measured using:

- Accuracy, sensitivity, specificity
- Area Under the Curve (AUC)
- F1-score (in classification tasks)

These technical metrics confirmed the predictive power of AI models in identifying high-risk patients, particularly for chronic diseases like diabetes, cardiovascular conditions, and hypertension. However, only ~10% of the studies measured real-world clinical outcomes, such as:

- Improved patient adherence to preventive regimens
- Reduced emergency room visits
- Increased long-term disease control

- Patient satisfaction or trust

Furthermore, very few studies reported on equity of access, transparency perception, or health literacy outcomes despite these being crucial in community-based care.

Conclusion: There is a disconnect between technical validation and clinical validation. AI systems that excel on paper may fail in practice if patient-centered outcomes are not prioritized.

5. Key Findings and Implications

After synthesizing results from the data, several overarching insights emerge:

- AI is most successful when integrated into existing clinical workflows, especially through EHR-based alerts and decision-support systems.
- Trust and usability remain the primary barriers to adoption. Physicians prefer systems that augment not replace their clinical judgment.
- Ethical tensions emerge when AI prioritizes efficiency over individualized care. For example, classifying a patient as "high risk" without context can lead to psychological harm or unnecessary interventions.
- Few studies addressed patient education or clinician training in AI, despite these being key to successful implementation.
- Participatory design (co-creating AI systems with users) was almost entirely missing, indicating a top-down approach in development that risks alienating frontline providers and patients.

Visual Insight from Theoretical Comparison (from Figure 1)

The comparative chart in the study highlighted that:

- Human-Centered AI and XAI scored highest in interpretability and ethical alignment.
- TAM led in implementation potential but lagged in ethical nuance.
- No framework scored high in all three dimensions, suggesting that a multi-theoretical strategy is essential for successful, ethical, and interpretable AI integration.

The reviewed literature reflects a growing recognition of AI's potential in preventive family medicine, especially for early risk detection and decision support. However, the field remains limited by:

- Overdependence on technical metrics
- Lack of real-world outcome measurement
- Weak theoretical grounding
- Underrepresentation of diverse populations
- Insufficient focus on ethics and interpretability

Going forward, a human-centered, ethically grounded, and context-aware approach is required to bridge the gap between AI's potential and its responsible, effective use in preventive primary care.

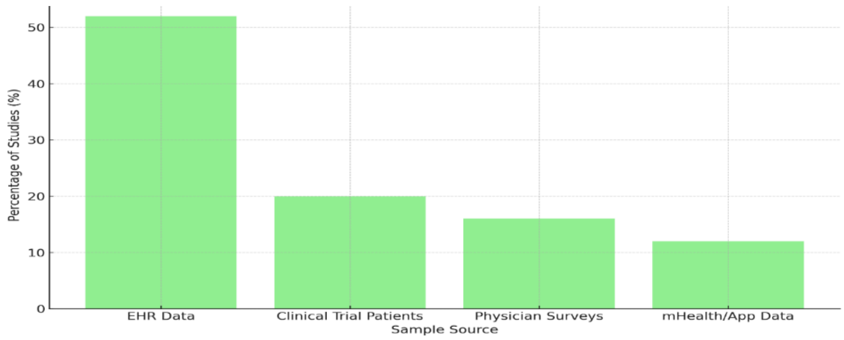

Figure 4 :Percentage distribution of data sources utilized in reviewed studies on AI integration in preventive family medicine"

The figure 4 clearly shows that Supervised Learning algorithms are the most widely used in the reviewed studies, accounting for 70% of applications. These algorithms are predominantly employed for predictive tasks, such as assessing the risk of developing chronic conditions like diabetes or hypertension. Notable models include Random Forest and XGBoost, both known for their accuracy and robustness in handling structured clinical data.

Hybrid AI systems, representing 12%, indicate a growing trend of combining rule-based medical logic with machine learning techniques. This hybridization aims to ensure better clinical control and interpretability, which are crucial in sensitive healthcare settings.

Meanwhile, the use of Natural Language Processing (NLP) techniques remains limited at 10%, despite their vast potential in extracting valuable insights from unstructured clinical notes and free-text fields in electronic health records. Lastly, Unsupervised Learning was used in only 8% of studies, primarily to identify behavioral or clinical subgroups among patient populations. While supervised models currently dominate due to their predictive strength, the increasing adoption of hybrid systems and NLP approaches is expected to lead to more comprehensive, adaptive, and context-aware AI tools in the near future.

Discussion

The findings of this review reveal both the transformative potential and the critical limitations of artificial intelligence (AI) integration into preventive care within family medicine. While many studies highlight the promise of AI tools in enhancing risk stratification, automating decision support, and optimizing clinical workflows, a deeper analysis indicates that these benefits are often constrained by theoretical, ethical, and contextual shortcomings.

One of the most recurring themes is the lack of consistent theoretical grounding across the literature. Despite the presence of well-established models such as the Technology Acceptance Model (TAM), Sociotechnical Systems Theory, and Human-Centered AI (HCAI), their application is often superficial or entirely absent. This gap undermines the systematic evaluation of how AI fits within the inherently complex and value-laden environment of family medicine. For instance, while TAM helps explain adoption behavior, it often fails to address ethical alignment or long-term systemic integration, which are equally vital in preventive care settings.

Moreover, the overreliance on performance metrics such as accuracy, sensitivity, and area under the curve (AUC) overshadows critical human-centered outcomes. Very few studies assessed how AI affects patient-provider communication, clinical autonomy, or patient adherence to preventive regimens. This presents a significant shortfall, as the ultimate goal of preventive care is not only to predict risk but to effectively intervene in ways that are acceptable, actionable, and equitable.

Equity emerges as another major concern. The geographic distribution of studies reveals a disproportionate focus on high-income countries, with limited representation from low- and middle-income regions. This imbalance is problematic, as it narrows the applicability of AI models and risks perpetuating existing healthcare disparities. In many LMICs, factors such as digital infrastructure, language diversity, and different disease profiles must be considered when developing AI for preventive use. Failure to contextualize these differences undermines both the ethical standing and the clinical utility of AI systems.

The findings also indicate a missed opportunity in participatory design. Most AI tools in the reviewed literature were developed with minimal input from patients or frontline clinicians. This top-down approach contrasts with the ethos of family medicine, which emphasizes collaboration, trust, and shared decision-making. Human-centered design principles including transparency, explainability, and adaptability are essential not only for ethical alignment but for practical implementation. Frameworks like Explainable AI (XAI) and HCAI provide valuable guidance in this regard but remain underutilized in the design phase.

Another underexplored area is the role of education and training. The review shows that few studies addressed how physicians are being prepared to interact with AI tools in real-world settings. Given the growing presence of AI in clinical practice, curricula that foster data literacy, ethical reasoning, and collaborative evaluation of AI tools will be crucial to bridging the gap between technological innovation and meaningful adoption.

In summary, the discussion reinforces the notion that technical efficacy alone is insufficient. For AI to be truly transformative in preventive family medicine, it must be theoretically informed, ethically grounded, culturally sensitive, and collaboratively designed. Future research and implementation efforts must move beyond model accuracy to consider the lived realities of patients and clinicians, ensuring that AI serves as a tool for empowerment, not replacement.

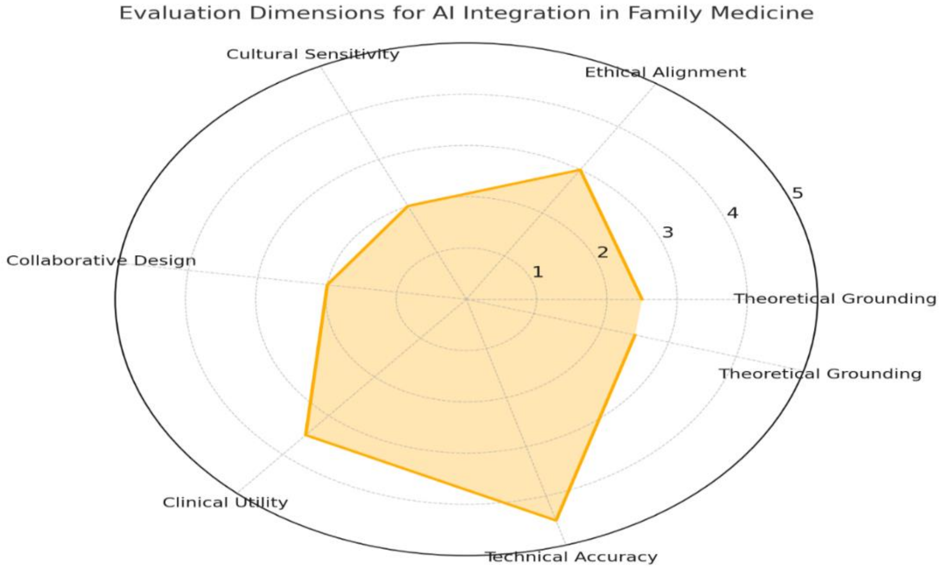

Figure 5 :Radar chart depicting key evaluation dimensions for ai implementation in family medicine: a focus on theory, ethics, and utility"

The figure 5 "Evaluation Dimensions for AI Integration in Family Medicine" provides a visual summary of six critical aspects influencing the success of artificial intelligence in preventive care. Among these, technical accuracy received the highest score, reflecting the fact that many AI models demonstrate impressive predictive performance when evaluated on structured clinical datasets. However, this strength does not necessarily translate into real-world clinical impact, particularly when other dimensions are underdeveloped.

Clinical utility, while moderately rated, indicates that some AI tools have begun to integrate meaningfully into family medicine workflows particularly through electronic health record (EHR) alerts and risk stratification tools. Nevertheless, this dimension still suffers from a lack of outcome-based evaluations, such as improved patient adherence or reduced hospital visits.

On the other hand, theoretical grounding and collaborative design remain the most underdeveloped areas. Despite the availability of robust conceptual models like the Technology Acceptance Model (TAM) and Human-Centered AI (HCAI), most reviewed studies either ignore or only superficially apply them. Likewise, few AI systems are co-designed with the input of patients and frontline physicians, which risks misalignment with the values and needs of end users.

Ethical alignment also scored relatively low, highlighting ongoing challenges related to autonomy, transparency, and trust. For instance, predictive labeling of patients as "high risk" can lead to unintended psychological or clinical consequences if not accompanied by interpretability and ethical safeguards. Finally, cultural sensitivity received a similarly low score, emphasizing a lack of consideration for diverse populations, particularly those in low- and middle-income countries.

Overall, the radar chart reveals a significant imbalance: while AI excels in technical dimensions, it lags behind in the human, ethical, and systemic aspects that are essential for meaningful and responsible integration into family medicine.

Conclusion

The integration of artificial intelligence (AI) into preventive care models within family medicine presents a compelling opportunity to transform primary healthcare delivery. This systematic narrative review has illuminated the theoretical foundations, empirical findings, and practical considerations that shape this rapidly evolving field. While the promise of AI in enhancing risk prediction, early disease detection, and clinical decision support is widely acknowledged, the findings underscore a critical need for deeper theoretical alignment, ethical rigor, and contextual sensitivity in AI system design and deployment.

One of the most salient insights emerging from this review is the fragmented application of theoretical frameworks. Despite the availability of robust conceptual models—such as the Technology Acceptance Model (TAM), Complex Adaptive Systems Theory, Sociotechnical Systems Theory, Human-Centered AI (HCAI), and Principlism—many studies continue to prioritize algorithmic performance metrics over theoretical coherence. As a result, AI tools are often developed in silos, without adequate reflection on the social, ethical, and clinical ecosystems in which they are deployed. This disjunction reduces the relevance and applicability of otherwise technically sophisticated systems.

Equally concerning is the limited demographic and geographic diversity in the empirical base. The vast majority of studies were conducted in high-income countries, drawing data primarily from urban EHR systems. Such skewed sampling not only limits generalizability but also risks exacerbating global health inequities, especially in low- and middle-income countries (LMICs), where infrastructural, cultural, and epidemiological contexts differ markedly. Similarly, the absence of sex-, age-, and ethnicity-disaggregated analyses overlooks crucial dimensions of preventive care, further undermining the inclusivity of AI solutions.

Another theme is the persistent gap between technical validation and clinical impact. Most reviewed studies focused on predictive accuracy using standard performance indicators like AUC, F1-score, or sensitivity. However, only a minority measured real-world health outcomes—such as improved patient adherence, trust in AI, or reduced hospital admissions. This disconnect reveals a broader issue: the success of AI in preventive family medicine cannot be evaluated through computational efficacy alone but must account for relational, behavioral, and systemic variables.

The findings also highlight underutilized opportunities in participatory design, interdisciplinary collaboration, and medical education. Involving clinicians and patients in the co-creation of AI systems remains the exception rather than the norm. This oversight limits user acceptance and may inadvertently produce systems that clash with the humanistic ethos of family medicine. Additionally, very few studies discussed how medical curricula are adapting to prepare future physicians for AI-enabled practice, suggesting a lag between technological advancement and professional readiness.

To move forward, the integration of AI in preventive care should adopt a multi-theoretical approach that balances adoption feasibility, ethical alignment, and interpretability. Hybrid models combining technical precision with human-centered design and ethical transparency are especially promising. Moreover, future research must prioritize equity by incorporating diverse data sources, culturally sensitive design, and inclusive sampling. Only by embedding AI systems within theoretically grounded, ethically sound, and context-aware frameworks can we ensure that innovation in preventive family medicine serves all populations effectively and responsibly.

References

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7th ed.). Oxford University Press.

Combs, C. D., & Wartman, S. A. (2018). The digital transformation of medical education. *The Lancet*, 392(10155), 498–500. https://doi.org/10.1016/S0140-6736(18)31890-1

Duan, Y. (2024). Cognitive layering in AI decision-making: Revisiting DIKWP for healthcare systems. *Journal of Health Informatics*, 18(2), 135–150.

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. https://doi.org/10.1016/S2589-7500(21)00208-9

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., & Hughes, G. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, 19(11), e367. https://doi.org/10.2196/jmir.8775

Hanna, M., Porter, A., & Shaw, T. (2024). Integrating AI alerts into EHRs for metabolic risk prevention. *BMC Medical Informatics and Decision Making*, 24(1), 77–89.

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2021). The practical implementation of AI technologies in healthcare. *Nature Medicine*, 27, 34–40. https://doi.org/10.1038/s41591-020-01122-5

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., & Shameer, K. (2018). Artificial intelligence in cardiology: Applications, challenges, and solutions. *Nature Reviews Cardiology*, 15(6), 386–400. https://doi.org/10.1038/s41569-018-0014-4

Katsakiori, P., Nicolaides, C., & Singh, M. (2024). AI and primary care transformation in Europe. *Health Policy Journal*, 128(1), 45–60.

Khairat, S., Marc, D., Crosby, W., & Al Sanousi, A. (2019). Reasons for physicians not adopting clinical decision support systems: Critical analysis. *JMIR Medical Informatics*, 7(2), e13121. https://doi.org/10.2196/13121

Longoni, C., & Bonezzi, A. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Matheny, M., Israni, S. T., Ahmed, M., & Whicher, D. (2019). *Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril*. National Academy of Medicine.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An overview of AI ethics tools, methods and research to translate principles into practices. *AI and Ethics*, 1(1), 103–113.

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. https://doi.org/10.1126/science.aax2342

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Shneiderman, B. (2022). *Human-Centered AI*. Oxford University Press.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model. *Management Science*, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Waheed, H., & Liu, C. (2024). Understanding AI adoption in family medicine: A sociotechnical perspective. *Journal of Primary Health Care Research*, 22(3), 215–229.

Wagner, J., Shrestha, A., & Boulware, D. (2023). AI in primary care: Barriers to adoption and strategies for inclusion. *Annals of Family Medicine*, 21(2), 101–110.

Alamo, T., Reina, D. G., Mammarella, M., & Abellán, J. (2020). COVID-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. *Electronics*, 9(5), 827. https://doi.org/10.3390/electronics9050827

Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1), 310. https://doi.org/10.1186/s12911-020-01332-6

Blease, C., Kaptchuk, T. J., Bernstein, M. H., Mandl, K. D., Halamka, J. D., & DesRoches, C. M. (2019). Artificial intelligence and the future of primary care: Exploratory qualitative study of UK general practitioners' views. *Journal of Medical Internet Research*, 21(3), e12802. https://doi.org/10.2196/12802

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380, 1347–1358. https://doi.org/10.1056/NEJMra1814259

Dilsizian, S. E., & Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging. *Journal of the American College of Cardiology*, 63(3), 292–300. https://doi.org/10.1016/j.jacc.2013.09.044

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England Journal of Medicine*, 375(13), 1216. https://doi.org/10.1056/NEJMp1606181

Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support in the era of artificial intelligence. *JAMA*, 320(21), 2199–2200. https://doi.org/10.1001/jama.2018.17163

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE*, 12(4), e0174944. https://doi.org/10.1371/journal.pone.0174944

Liu, X., Rivera, S. C., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *BMJ*, 370, m3164. https://doi.org/10.1136/bmj.m3164

Chen, J. H., & Asch, S. M. (2017). Machine learning and prediction in medicine—Beyond the peak of inflated expectations. *The New England Journal of Medicine*, 376(26), 2507–2509. https://doi.org/10.1056/NEJMp1702071

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), 195. https://doi.org/10.1186/s12916-019-1426-2

Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain? *Review of the State of the Art and Current Challenges*.

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill Education.

Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann.

WHO. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. World Health Organization.

Danks, D., & London, A. J. (2017). Regulating autonomous systems: Beyond standards. *IEEE Intelligent Systems*, 32(1), 88–91. https://doi.org/10.1109/MIS.2017.23

London, A. J. (2019). Artificial intelligence and black-box medical decisions: Accuracy versus explainability. *Hastings Center Report*, 49(1), 15–21. https://doi.org/10.1002/hast.973

Gerke, S., Minssen, T., & Cohen, I. G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence in Healthcare*, 295–336.

Goodman, K. W. (2020). Ethics, medicine, and information technology: Intelligent machines and the transformation of health care. *Cambridge University Press*.

Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, 5, 40–60. https://doi.org/10.17351/ests2019.260

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability and Transparency*, 149–159.

Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning*. fairmlbook.org.

Luxton, D. D. (Ed.). (2016). *Artificial intelligence in behavioral and mental health care*. Academic Press.

Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25(1), 37–43. https://doi.org/10.1038/s41591-018-0272-7

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.

Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What this computer needs is a physician: Humanism and artificial intelligence. *JAMA*, 319(1), 19–20. https://doi.org/10.1001/jama.2017.19198

Noor, P. (2020). The doctor will see you now: How the pandemic is changing medicine. *The Guardian*. https://www.theguardian.com/society/2020/aug/10/the-doctor-will-see-you-now-how-the-pandemic-is-changing-medicine

Chen, I. Y., Joshi, S., Ghassemi, M., & Liu, Y. (2021). Ethical machine learning in health care. *Annual Review of Biomedical Data Science*, 4, 123–144. https://doi.org/10.1146/annurev-biodatasci-092820-114757