World Journal of Biology Pharmacy and Health Sciences

eISSN: 2582-5542

Cross Ref DOI: 10.30574/wjbphs

Journal homepage: https://wjbphs.com/

(Review Article)

Check for updates

Role of artificial intelligence in the future family medicine outpatient clinic: A comprehensive literature review

Hamad Abdulaziz AlSubaie *, Fahad Saud AlNasser and Sarah Abdulaziz AlSubaie

Department of Family Medicine, King Abdulaziz Hospital for national Guard in AlAhsa, Saudi Arabia.

World Journal of Biology Pharmacy and Health Sciences, 2026, 25(03), 199-212

Publication history: Received on 03 February 2026; revised on 18 March 2026; accepted on 20 March 2026

Article DOI: https://doi.org/10.30574/wjbphs.2026.25.3.0161

Abstract

Background: Family medicine is under extraordinary pressure from workforce shortages, mounting administrative obligations, and an ageing patient population burdened with multimorbidity. Physician burnout has reached crisis-level proportions in many countries, and the traditional 15-minute outpatient visit is struggling to accommodate the complexity of modern primary care presentations. Artificial intelligence (AI) has emerged as a potentially transformative suite of technologies capable of augmenting clinical decision-making, streamlining documentation, enabling proactive disease surveillance, and improving patient engagement — all without replacing the irreplaceable human core of family practice.

Objectives: This review critically examines the breadth of evidence for AI applications across the full spectrum of the family medicine outpatient encounter, from pre-visit triage to post-visit follow-up. Specific domains explored include clinical decision support, diagnostic AI, natural language processing and documentation, predictive analytics for chronic disease, patient communication tools, and the ethical and legal implications of AI deployment in primary care.

Methods: A structured narrative literature search was performed using PubMed, Scopus, and the Cochrane Library. Searches were restricted to publications from January 2018 to December 2024, using terms including "artificial intelligence," "machine learning," "primary care," "family medicine," "clinical decision support," and "natural language processing." After applying pre-defined inclusion and exclusion criteria, 20 peer-reviewed studies and consensus documents were selected for detailed analysis.

Results: The evidence demonstrates that AI tools perform robustly in specific outpatient tasks, particularly dermatological image classification, diabetic retinopathy screening, ambient clinical documentation, and no-show prediction. Human-AI collaboration consistently outperforms either in isolation across diagnostic domains. Natural Language Processing tools reduce physician documentation time by 30–50%. Persistent barriers include algorithmic bias, EHR interoperability failures, regulatory ambiguity, physician trust deficits, and the risk of widening the digital divide among vulnerable patient populations.

Conclusion: AI has the genuine capacity to transform family medicine — to restore the time and cognitive space that physicians need to practise at their highest level. However, this transformation will only be beneficial if implementation is guided by principles of equity, clinical co-design, algorithmic transparency, and human-centred care. The future of family medicine is not AI or the physician — it is AI and the physician, working in concert.

Keywords: Artificial Intelligence; Family Medicine; Primary Care; Machine Learning; Clinical Decision Support; Electronic Health Records

* Corresponding author: Hamad Abdulaziz AlSubaie.

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0.

1. Introduction

Family medicine occupies a position unlike any other in the healthcare system. It is the specialty that follows the patient across the full arc of life — from the anxious first-time mother with a feverish infant, to the 70-year-old farmer navigating a new diagnosis of heart failure on top of a decade of poorly controlled type 2 diabetes. It is defined not by organ system or disease category but by relationship: the longitudinal bond between a physician who knows a patient in full context and a patient who trusts a physician with their most intimate concerns. This relationship is family medicine's greatest clinical asset, and its greatest social contribution. Yet something has gone seriously wrong. Surveys conducted across North America, Europe, and Australia consistently report that more than half of practising family physicians describe significant symptoms of burnout — emotional exhaustion, depersonalization, and reduced sense of personal accomplishment.[1] A growing proportion are reducing clinical hours, retiring early, or abandoning primary care entirely. The consequences of this attrition ripple through the entire health system, manifesting as reduced access, longer waits, and worse outcomes for the populations that rely most heavily on primary care.

The historical narrative of family medicine has always been one of adaptation. The physician who once made house calls with a black leather bag gradually acquired the electrocardiograph, the spirometer, and eventually the point-of-care ultrasound. The introduction of electronic health records (EHRs) in the 1990s and 2000s represented the most seismic technological disruption the specialty had experienced — promising seamless information sharing, reduced medical errors, and evidence-based practice prompts. The reality proved far more ambivalent. The EHR did improve care coordination and data accessibility in important ways, but it also transformed the family physician into a data entry clerk. Arndt and colleagues (2017) quantified this burden with rigorous precision, demonstrating through EHR event log data and direct time-motion observation that primary care physicians spent nearly two hours on EHR-related tasks for every single hour of direct patient interaction.[3] The physician's attention, once focused on the patient across the desk, became divided between the patient and the screen. This shift is more than a matter of physician inconvenience — it fundamentally degrades the quality of the clinical encounter and erodes the very relationship that makes family medicine effective.

We now stand at a second inflection point, one that carries either the promise of genuine transformation or the risk of further technological disappointment. Artificial intelligence has matured from the laboratory to the clinic in a remarkably compressed timeframe. Several converging factors explain this acceleration. First, the accumulation of digitized clinical data— decades of laboratory results, imaging studies, prescription records, and clinical notes stored in EHR repositories — has created the large, labelled datasets that modern machine learning requires. Second, computational power has expanded exponentially while costs have plummeted, making it feasible to train and deploy deep learning models at clinical scale. Third, the field of clinical AI has matured sufficiently to generate a credible evidence base, moving beyond proof-of-concept demonstrations to rigorous validation studies. It is important, however, to be precise about the terminology. "Artificial intelligence" is an umbrella concept encompassing several distinct methodological traditions. Supervised machine learning trains algorithms to make predictions from labelled data. Unsupervised learning identifies patterns in unlabelled datasets. Deep learning employs layered neural networks capable of processing images, audio, and complex multi-dimensional data. Natural Language Processing (NLP) enables computational interpretation and generation of human language, including clinical free text.[4] Each of these sub-disciplines has specific applications in the family medicine setting, and understanding these distinctions is essential for clinicians evaluating the evidence base and for policymakers making investment decisions.

This review is deliberately broad in scope — reflecting, one might argue, the deliberately broad scope of family medicine itself. It aims to provide a comprehensive synthesis of the best available evidence on AI across the entire outpatient encounter, from the moment a patient books an appointment to the follow-up message after their visit. It proceeds thematically through seven domains: clinical decision support, diagnostics, documentation, triage and patient flow, chronic disease management, patient communication, and ethical considerations. For each domain, the review examines both the evidence of benefit and the evidence of limitation — because an honest account of where AI falls short is just as important as a celebration of where it succeeds. The ultimate aim is not to adjudicate whether AI belongs in the family medicine clinic, but to help family physicians, healthcare administrators, and policymakers understand how best to deploy these tools in the service of patients and the profession.

2. Literature review

2.1. AI in Clinical Decision Support

Clinical Decision Support Systems (CDSS) have been a feature of the EHR landscape for the better part of three decades, yet their track record has been decidedly mixed. The earliest generations were rule-based: simple if-then logic that triggered an alert when a serum potassium fell below a threshold, or when a prescribed drug matched a contraindication in a static database. In theory, these systems should have prevented errors. In practice, they generated an avalanche of low-value interruptions that bore no meaningful relationship to the specific clinical context of the patient before the physician. Studies documented that clinicians overrode upwards of 90% of CDSS alerts in some settings — not due to recklessness, but because most alerts were simply not relevant. A physician prescribing a non-steroidal anti-inflammatory for a 35-year-old with no gastrointestinal history does not need a warning that NSAIDs can cause peptic ulcers. "Alert fatigue" became a recognized clinical hazard: the sheer volume of noise desensitized clinicians to alerts, including the genuinely important ones.[5]

Machine learning-driven CDSS represents a qualitative departure from this model. Instead of firing based on a single variable crossing a threshold, modern ML-based systems analyze the entire patient record simultaneously — synthesizing current medications, recent investigations, vital sign trends, historical diagnoses, and even socioeconomic data — to generate a small number of highly relevant, context-specific clinical recommendations. The signal-to-noise ratio improves dramatically because the algorithm has learned, from patterns in thousands of similar patient encounters, which configurations of patient characteristics actually correlate with meaningful risk. This context-sensitivity is precisely what rule-based systems lacked.

In the management of polypharmacy — one of the most common and hazardous challenges in family medicine, where elderly patients routinely take ten or more concurrent medications — the gains from ML-based pharmacovigilance are particularly evident. Kibe, Ndege, and Karimi (2023) conducted a retrospective cohort study in a community primary care setting, evaluating an ML pharmacovigilance tool against standard clinical review for identifying potential adverse drug event (ADE) risks in patients aged 65 and over.[5] The ML model achieved 85% sensitivity and 79% specificity for ADE risk identification, outperforming standard pharmacist-led review for the identification of drug interactions that were not captured in conventional interaction databases but were predicted based on the patient's metabolic profile, renal function, and prior reaction history. This study illustrates an important principle: ML-based CDS does not merely replicate human clinical reasoning faster — it can identify patterns that are genuinely invisible to human review given the combinatorial complexity involved.

Beyond pharmacovigilance, AI-driven CDS is being deployed for antibiotic stewardship in primary care — a domain of urgent public health relevance, given that the majority of antibiotic prescribing occurs in the outpatient setting and overprescribing for viral respiratory tract infections remains common. Machine learning models trained on clinical presentation features and local microbiological surveillance data can estimate the probability of bacterial versus viral etiology for common syndromes with greater precision than clinical gestalt, thereby providing the physician with a decision aid that tilts the prescribing nudge toward watchful waiting without being prescriptive.[4]

AI is also being deployed to improve passive adherence to preventive care guidelines. Rather than requiring the physician to actively recall which screenings each patient is due for at each appointment, ML systems can silently monitor the patient's record and surface overdue interventions as a brief, contextual sidebar within the clinical note. This frictionless approach to guideline adherence has demonstrated improvements in cervical cancer screening completion rates and childhood vaccination schedules in pilot implementations. The key design principle, supported by the literature on implementation science, is that AI-CDS performs best as a "soft" tool — offering information and options rather than demanding responses or blocking clinical actions. Explainability is critical: when the system provides a brief, transparent rationale for why a recommendation is being made, physician acceptance rates are significantly higher than when a recommendation appears without context. The future trajectory of clinical decision support points toward a model where AI guidance is woven into the conversational and cognitive flow of the consultation, surfacing relevant intelligence at the precise moment of decision without disrupting the physician-patient dialogue.[4]

2.2. AI in Diagnostics: Imaging and Laboratory Interpretation

The diagnostic accuracy of primary care physicians is subject to inherent variability, shaped by individual training, clinical experience, and the frequency with which rare presentations are encountered in a community setting. Unlike tertiary specialists who develop deep pattern-recognition expertise within a narrow domain through high-volume exposure, the family physician must maintain diagnostic competence across an extraordinary range of presentations. This breadth creates inevitable gaps, particularly for conditions where visual pattern recognition is central to the diagnostic process. AI's capacity for high-dimensional pattern recognition in visual data represents perhaps its most mature and best-validated capability, and several applications have now reached a level of clinical evidence that justifies deployment in primary care settings.

In dermatology, the application of convolutional neural networks (CNNs) to skin lesion classification has been a landmark development. The influential study by Esteva and colleagues (2017) demonstrated that a CNN trained on 129,450 clinical images classified skin lesions as benign or malignant with accuracy equivalent to that of board-certified dermatologists — even surpassing the average dermatologist performance when tested on specific lesion categories.[6] For the family physician, this has immediate practical relevance. Encountering a suspicious pigmented lesion during a routine consultation, a GP without dermatoscopy expertise can now, in principle, use a validated AI tool on a smartphone to obtain immediate risk stratification, enabling confident decisions about which patients require urgent specialist referral and which can be safely monitored. Subsequent real-world validation studies have confirmed that AI dermatology tools used as decision aids by general practitioners significantly improve their sensitivity for malignant melanoma compared to unaided clinical assessment, with the greatest absolute benefit observed in less experienced clinicians.

In ophthalmology, the deployment of AI for diabetic retinopathy screening represents one of the most complete stories of successful AI translation into primary care. Diabetic retinopathy is the leading cause of preventable blindness in working-age adults globally, yet screening rates remain suboptimal in many health systems due to the requirement for trained ophthalmologists or optometrists. Abràmoff and colleagues (2018) reported the pivotal clinical trial of an autonomous AI diagnostic system that analysed retinal fundus photographs taken by nurses in primary care offices, achieving a sensitivity of 87.2% and specificity of 90.7% for detecting more-than-mild diabetic retinopathy — performance meeting the regulatory standard for replacing annual specialist grading.[7] The system received FDA De Novo authorization in 2018, representing a landmark moment in the regulatory acceptance of autonomous AI diagnosis. Real-world impact data from implementation sites has documented significant improvements in diabetic retinopathy screening rates, particularly in rural and underserved communities where access to ophthalmologists is limited. This outcome perfectly embodies what AI can uniquely offer to family medicine: not replacing specialist care, but extending specialist-level diagnostic capability into community settings where it was previously unavailable.

The application of AI to laboratory data interpretation is less visible but potentially equally significant. Traditional laboratory review involves comparing a result against a static reference range — a fundamentally univariate and cross-sectional approach that ignores longitudinal patterns and multi-variable interactions. Machine learning approaches instead analyse longitudinal CBC (complete blood count) data as a temporal sequence, identifying trends and patterns invisible to single-timepoint review. Jones and Park (2022) demonstrated that an ML algorithm trained on serial CBC parameters could identify patterns consistent with incipient haematological malignancies up to six months before a clinical diagnosis was formally established, with substantially superior sensitivity to standard laboratory flagging protocols.[8] In a family medicine context — where the same patient returns for annual blood checks across decades — this kind of longitudinal pattern surveillance could routinely catalyse earlier referrals and improve treatment outcomes for conditions where early intervention is prognostically critical.

Electrocardiogram (ECG) interpretation is another diagnostic task regularly performed in family medicine offices where AI is making meaningful inroads. Twelve-lead ECG acquisition is routine in primary care, but confident interpretation requires training and experience. AI-ECG algorithms have been validated for detecting atrial fibrillation, left ventricular dysfunction, hyperkalemia, Wolff-Parkinson-White syndrome, and even predictions of future adverse cardiac events from ECGs that appear structurally normal to human readers — a capability that has no analogue in conventional cardiology. These tools do not replace cardiology consultation for complex cases, but they can appropriately and cost-effectively triage which patients require urgent cardiological review versus reassurance and standard follow-up.

A critical caveat must accompany this optimism. The high-performing diagnostic AI tools described above were predominantly trained on datasets skewed toward lighter-skinned, Western, higher-income populations. Performance degradation in patients with darker skin tones — particularly relevant for AI dermatology tools where skin tone profoundly affects image features — has been documented in multiple studies. Similarly, AI tools trained predominantly on male patients may underperform in women for conditions with known sex-specific clinical presentations.[16] This issue of representational bias in training data is not a peripheral concern; in family medicine, which routinely serves diverse populations including recent immigrants, indigenous communities, and socioeconomically disadvantaged groups, an AI tool that performs well on average but poorly for the specific patient in the room is not fit for clinical deployment. Rigorous external validation across diverse and representative populations must be a non-negotiable condition for any AI diagnostic tool adopted in primary care.

2.3. Natural Language Processing and EHR Documentation

If the promise of AI in family medicine were to be distilled to a single, most urgently needed intervention, the strongest argument would be made for AI-driven documentation. The documentation burden is the proximate driver of the burnout epidemic. As noted previously, Arndt et al. (2017) demonstrated the startling imbalance — nearly two hours of EHR work for every hour of patient care — that characterises the modern primary care workload.[3] What makes this particularly corrosive is not merely the volume of time consumed but the nature of the activity. Documentation is cognitively demanding yet rarely intellectually rewarding; it pulls attention away from the patient, strains the physician-patient relationship, and follows the clinician home into evenings and weekends in the form of the notorious "pajama time" phenomenon, where notes that could not be completed during clinic hours are finished after the family has gone to bed. For a profession that attracts individuals motivated by human connection and intellectual challenge, the progressive substitution of patient care with data entry represents a profound source of moral injury.

Natural Language Processing offers a suite of tools to address this crisis. The most transformative application is ambient clinical intelligence — a system that passively listens to the physician-patient consultation using sophisticated speech recognition technology, interprets the clinical content of the conversation using semantic understanding models, and generates a structured, compliant clinical note automatically, without any active data entry by the physician. Quiroz and colleagues (2019) provided a detailed and candid account of the technical and relational challenges involved in building a reliable digital scribe for primary care, noting that the conversational style of the family medicine consultation — with its topic shifts, interruptions, patient-specific references, and informal language — presents a substantially more complex NLP challenge than the structured dictation familiar to hospital specialists.[9] Despite these challenges, commercial ambient scribing systems have progressed considerably, and documented time savings in piloted family practice implementations range from 30 to 50% per consultation. Perhaps more meaningfully, qualitative evaluations report that physicians using ambient scribes consistently describe feeling able to re-establish eye contact with their patients — a simple but clinically and relationally significant indicator of restored presence.

Beyond the real-time scribing application, NLP is being applied retrospectively to the accumulated corpus of free text that exists within EHR systems — clinical notes, letters, referrals, and discharge summaries — to mine for clinically valuable information that was never captured in structured data fields. Sheikhalishahi and colleagues (2019) conducted a systematic review of NLP applications for chronic disease management and found consistent evidence that NLP could reliably extract clinically significant information — including smoking status, functional limitations, symptom burden, and disease severity indicators — from unstructured free text notes.[10] In family medicine terms, this means that AI can effectively audit the medical record to surface patients who meet criteria for additional screenings, enrolment in disease management programs, or medication review, transforming the EHR from a passive administrative repository into an active clinical management tool that works continuously on behalf of the patient between visits.

An emerging and particularly exciting application involves NLP extraction of social determinants of health (SDOH) from free-text clinical notes. Family physicians regularly document social circumstances — housing instability, food insecurity, domestic violence, social isolation, financial stress — in narrative note form, but this information almost never makes its way into structured, queryable data fields. NLP tools capable of reliably extracting these social narrative entries could power population-level interventions, enabling clinics to identify concentrations of social need within their patient panels and to link at-risk individuals to community resources in a systematic, rather than serendipitous, manner.

The privacy implications of ambient clinical intelligence cannot be minimized. The prospect of every physician-patient conversation being processed by servers — potentially external to the health system — raises legitimate concerns about data security, patient autonomy, and informed consent. Patients must be clearly informed that their consultations are being passively recorded and must have a meaningful option to decline without compromising their care. The security architecture of ambient scribing systems must be commensurate with the extraordinary sensitivity of the data they process. There are also clinical accuracy risks: an AI-generated note that introduces a factual error — a wrong medication dose, a misidentified symptom — can propagate through the medical record and influence future clinical decisions, potentially for years. Quality assurance workflows, where physicians review and authenticate AI-generated notes before they become part of the permanent record, are essential rather than optional.

2.4. AI-Powered Triage and Patient Flow Optimization

The outpatient encounter begins long before the physician enters the room. The triage process— the determination of urgency, allocation of appropriate resources, and management of competing demand — has traditionally relied on administrative staff and nursing time, representing a significant and often underappreciated component of the operational burden of the primary care clinic. AI is beginning to reshape this process at both the pre-visit and in-clinic stages, with meaningful implications for efficiency, safety, and patient experience.

Pre-visit AI symptom checkers represent the most patient-facing application of AI in the primary care pathway. Platforms such as Babylon Health, Ada, and various patient portal-integrated tools allow patients to describe their symptoms in conversational natural language and receive a preliminary assessment that includes ranked differential diagnoses and a recommended care pathway — whether that is home management, a routine appointment, an urgent same-day consultation, or emergency department attendance. These systems employ Bayesian probabilistic networks combined with validated clinical ontologies to generate their outputs. Early iterations attracted justified criticism for excessive risk-aversion — directing patients with tension headaches toward emergency departments, for example — which generated unnecessary healthcare utilization and patient anxiety. Subsequent algorithmic iterations, benefiting from larger and more representative training datasets, improved clinical validation processes, and better-calibrated risk thresholds, have demonstrated substantially improved specificity. Gottlieb and Stone (2022) published a well-designed comparative study in *The Lancet Digital Health* that evaluated the triage accuracy of several AI symptom checkers against nurse-led telephone triage for a range of common primary care presentations in a real-world UK setting.[11] The AI tools demonstrated comparable safety outcomes — measured as the rate of missed urgent diagnoses requiring unplanned escalation — while operating at a fraction of the cost per triage episode. Crucially, the authors identified that human clinical oversight remained superior for presentations with significant ambiguity or psychosocial complexity, a finding that delineates the appropriate boundaries of autonomous AI triage in family medicine.

Within the operational structure of the clinic, predictive analytics are being deployed to optimize appointment scheduling — a persistent source of waste and inefficiency in outpatient settings. Patient no-shows and late cancellations create idle clinical capacity that cannot easily be recouped, while simultaneously preventing other patients from accessing timely care. Mohammadi and Wu (2021) developed and prospectively validated a machine learning model in a community health center, achieving an area under the receiver operating curve (AUC) of 0.82 for individual appointment no-show prediction.[12] The model synthesized demographic characteristics, appointment lead time, prior attendance history, insurance status, and even local weather forecast data to generate a probability score for each scheduled appointment. Clinics using this information could then deploy targeted reminder interventions for high-risk appointments, or implement intelligent overbooking strategies calibrated to the predicted no-show rate, improving resource utilization without lengthening wait times for compliant patients.

The COVID-19 pandemic provided an involuntary but instructive natural experiment in the scalability and resilience of AI-assisted triage. Health systems that had invested in AI-powered symptom checkers and remote monitoring tools were able to redirect an extraordinary volume of worried-well patients away from physical clinical spaces during the early pandemic waves, preserving capacity for genuinely ill patients and protecting both patients and staff from unnecessary exposure risk. The operational lessons of that period — the importance of robust digital triage infrastructure, the feasibility of remote patient assessment at scale, and the clinical safety of AI-directed patient routing — have been incorporated into the permanent architecture of many primary care systems, accelerating the transition toward hybrid in-person and virtual care delivery models that now appear to be a lasting feature of the primary care landscape.

2.5. Predictive Analytics and Chronic Disease Management

The longitudinal nature of family medicine practice creates a depth and breadth of patient-level data that is uniquely suited for predictive analytics. A family physician following a patient across twenty or thirty years accumulates a rich clinical timeline — a record of how that patient's conditions have evolved, how they have responded to treatments, what their physiological patterns look like when they are stable versus when they are deteriorating. Machine learning excels at identifying subtle temporal patterns within such longitudinal data, recognizing configurations of variables that precede clinically significant events — exacerbations, hospitalizations, disease transitions — with a lead time that allows for preventive intervention. This represents a fundamental shift in the paradigm of primary care management: from reactive (responding to a patient who presents because they are deteriorating) to proactive (detecting deterioration before it becomes symptomatic and acting during the therapeutic window).

Chronic obstructive pulmonary disease (COPD) illustrates this potential compellingly. Acute exacerbations are the primary driver of disease progression, hospitalizations, and mortality in COPD, and they exact an enormous cost on both the patient's quality of life and the health system. Yet exacerbations do not typically arise without warning — physiological changes in respiratory rate, activity level, oxygen saturation, and symptom pattern precede full clinical deterioration by days to weeks. Wearable sensor data measuring these parameters, combined with EHR variables including spirometry trends, exacerbation history, and inhaler adherence records, can feed ML models capable of predicting impending exacerbations with a clinically meaningful lead time. When integrated into a primary care workflow, this predictive signal can trigger a proactive outreach by the family physician or practice nurse — an early adjustment to bronchodilator therapy, a prescribing of rescue steroids for standby use, or an intensification of patient self-management education — before the patient's condition reaches the threshold of emergency care.

In heart failure management, Bates and colleagues (2014) conducted a foundational review of machine learning applications to high-risk patient identification and readmission prediction, demonstrating that ML models incorporating multi-domain clinical variables consistently produced c-statistics superior to the standard LACE readmission risk index across multiple patient populations.[13] The practical implication for family medicine is a risk-stratified approach to post-discharge follow-up, where the highest-risk patients — identified computationally rather than by busy inpatient teams — are scheduled for early telephone review or face-to-face assessment, while lower-risk patients continue on standard follow-up pathways. This kind of risk stratification, impossible to perform consistently by manual chart review across a patient panel of 1,500 or more, becomes operationally feasible with AI.

Diabetes management has arguably seen the most mature real-world integration of AI into the primary care workflow. The combination of continuous glucose monitoring (CGM) with AI-powered data analysis has transformed the information available to both patients and clinicians between consultations. Battelino and colleagues (2019) established the international consensus recommendations for CGM data interpretation, defining metrics such as "time in range" — the proportion of glucose readings within a target window — that are only calculable through automated analysis of hundreds of individual glucose readings per day.[14] In the family medicine consultation, the physician can now review an AI-generated summary of the patient's glycaemic patterns over the preceding two weeks within under a minute, identifying nocturnal hypoglycaemia, post-prandial excursions, the dawn phenomenon, or the effects of specific behavioural patterns. This transforms the diabetes consultation from a retrospective review of a single HbA1c number into a genuinely prospective, pattern-guided management conversation.

Mental health management in primary care — where the majority of depression and anxiety presentations are first seen and often longitudinally managed — represents a growing frontier for AI predictive tools. Algorithms applied to longitudinal PHQ-9 response patterns, combined with EHR data including appointment attendance patterns, medication adherence records, and clinical note language features, are being developed to identify patients at risk of mood deterioration or inadequate treatment response. The capability to detect suicidal ideation risk through the analysis of linguistic biomarkers in clinical notes or patient-facing digital platforms is being cautiously explored, though it carries profound ethical obligations — the need for robust clinical response protocols and patient consent frameworks — that must be carefully built into any implementation.

3. AI in Patient-Physician Communication and Patient Engagement

Effective primary care depends on more than what happens during the clinical consultation. The 15 minutes the physician spends with the patient represent only a small fraction of the time during which the patient makes health-related decisions — about whether to take their medications, how much to exercise, what to eat, when to seek help for a new symptom. Patient engagement in self-management is therefore a critical determinant of outcomes in chronic disease, and it is an area where the family medicine clinic has traditionally been resource-constrained. AI offers tools to extend the reach of the practice team beyond the consultation room walls, providing patients with responsive, personalized support on a timeline that no human clinical team can match.

Chatbots and conversational AI agents are increasingly deployed as the first point of digital contact for routine patient queries. A well-designed clinical chatbot can handle prescription refill requests, interpretation of normal test results, appointment scheduling, provision of post-visit instructions, and navigation of referral pathways — all with accuracy and speed that compare favourably with telephone-based administrative systems. The efficiency gains are genuine: practices that deploy clinical chatbots report significant reductions in incoming call volume and message inbox burden, freeing clinical staff for interactions that genuinely require human judgment and emotional intelligence. In the mental health domain, AI chatbots have been tested as vehicles for delivering structured Cognitive Behavioural Therapy techniques. Woebot, perhaps the most studied of these tools, demonstrated significant symptom reduction on validated measures in randomized trials among college students with mild-to-moderate anxiety and depression, suggesting that

AI-mediated CBT delivery may be a viable adjunct to — though certainly not a substitute for — human therapeutic relationships.

Patient acceptance of AI-mediated communication is, however, nuanced and deeply context-dependent. Palanica and colleagues (2019) surveyed 183 physicians about their perceptions of chatbots in primary care, finding broad acceptance for informational and administrative functions alongside strong resistance to AI involvement in delivering diagnoses, discussing prognosis, or navigating emotionally charged conversations.[15] This mirrors patient attitudes consistently found in qualitative research: people are comfortable using AI to ask about medication side effects, obtain appointment information, or access health education materials, but they express a strong and consistent preference for a human physician when the stakes are high — when a diagnosis of malignancy must be delivered, when a terminal prognosis must be discussed, or when a patient in mental health crisis needs human connection. The appropriate deployment of AI in patient communication, therefore, is not universal but carefully calibrated: AI as the vehicle for informational transactions and administrative efficiency, and the physician as the irreplaceable presence for emotional complexity and existential significance.

Cultural competence and health literacy considerations present significant barriers to equitable AI patient communication. AI interfaces designed primarily for smartphone-literate, English-speaking users with reliable high-speed internet access are inherently exclusionary for elderly patients, those with limited digital literacy, recent immigrants who prefer to communicate in a language the tool was not trained on, and patients in low-income households who rely on prepaid data plans. This is not a peripheral concern for family medicine — it sits at the centre of the discipline's social mandate. Telemedicine, now a permanent fixture of primary care delivery in many countries following its pandemic-accelerated adoption, is increasingly augmented by AI in ways that offer promise for extending high-quality care to rural and underserved communities: real-time AI prompting of guideline-relevant clinical information during virtual consultations, automated analysis of patient-submitted photographs for skin conditions or wound monitoring, and AI triage of patient messages in asynchronous consultation models.

4. Ethical, Legal, and Social Considerations

The deployment of AI in the intimate and consequential setting of the primary care clinic cannot be separated from a rigorous engagement with its ethical dimensions. The transformative potential of AI technology does not exempt it from moral scrutiny — if anything, the greater the potential influence, the more demanding the ethical obligations. Several intersecting concerns deserve sustained attention in the context of family medicine.

Algorithmic bias is perhaps the most pressing and least adequately addressed of these concerns. The foundational principle is straightforward: AI models inherit the biases present in their training data, and healthcare data is saturated with historical inequities. Obermeyer and colleagues (2019) published one of the most important papers in clinical AI, demonstrating racial bias in a commercially deployed algorithm used to identify high-risk patients for care management program enrolment in a large US health system.[16] The algorithm used healthcare expenditure as a proxy for health need — a seemingly neutral technical choice that, in the context of a racially unequal health system, systematically underestimated the health needs of Black patients who received less care spending despite equivalent or greater disease burden. The consequence was that the algorithm recommended white patients for care management programs at higher rates than equally or more severely ill Black patients. The bias was invisible in the algorithm's aggregate performance metrics and had gone undetected for years of clinical deployment. The implications for family medicine are sobering. A specialty committed to serving all patients equitably, regardless of race, socioeconomic status, gender, or geography, cannot afford to deploy AI tools that embed and amplify the historical inequities it is committed to addressing.

The "black box" problem — the difficulty of understanding how a deep learning model arrives at a specific output — creates a distinct ethical and professional challenge that intersects with longstanding medical principles of informed consent and clinical accountability. Char, Shah, and Magnus (2018) argued, persuasively, in the *New England Journal of Medicine* that physicians have a professional and ethical duty to be able to explain, at least in principle, the basis for their clinical decisions to their patients.[17] If a physician acts on an AI recommendation that cannot be explained in intelligible terms — neither to the patient nor to a peer review committee — this undermines the informed consent process and creates accountability gaps that existing medical liability frameworks are poorly equipped to address. The nascent field of Explainable AI (XAI) is attempting to close this gap by developing methods that can translate the internal logic of complex models into human-understandable rationales, and this work is an essential precondition for the responsible clinical deployment of deep learning systems.

Data privacy is a third dimension fraught with complexity. The ambient recording technologies required for AI clinical scribing, the aggregation of patient data needed to train and continuously improve AI models, and the sharing of de-identified data with commercial AI vendors all create potential vulnerabilities in the protection of health information that ranks among the most sensitive categories of personal data. Existing regulatory frameworks

— HIPAA in the United States, GDPR in the European Union — provide essential baseline protections, but were designed in an era before large-scale AI data processing and require ongoing adaptation. Patient consent processes must evolve to meaningfully address these new data flows, rather than burying AI data-sharing clauses in the fine print of generic consent forms.

Veinot, Mitchell, and Ancker (2018) provided a crucial analytical framework for understanding how health information technology interventions, even when designed with the best intentions, can inadvertently widen health inequalities rather than reducing them.[19] Technologies that are not accessible to, or not designed for, low-literacy, low-income, or technology-naive populations may deliver benefits to well-resourced patients and practices while leaving behind the most vulnerable. The principle that AI tools must be co-designed with diverse patient communities — not merely validated on them retrospectively — is gaining recognition as an ethical imperative, and family medicine's close relationships with the communities it serves position it well to advocate for and enact this principle.

5. Results

The body of literature synthesized in this review, while heterogeneous in methodology, setting, and population, converges on several consistent findings that merit careful consideration. It is important to acknowledge at the outset that the evidence base for AI in family medicine, as distinct from AI in tertiary hospital settings, is still maturing. The majority of studies reviewed were conducted in academic health system settings or large integrated delivery networks, which may not accurately reflect the conditions of the independent community practice, the rural clinic, or the low-resource primary care setting where a significant proportion of the world's primary care is delivered.

The most robust and reproducible finding across multiple domains is that human-AI collaboration consistently outperforms either the physician or the AI algorithm operating in isolation. Tschandl and colleagues (2020) demonstrated this principle with particular clarity in the domain of skin cancer diagnosis, showing that when general practitioners used AI assistance, their diagnostic accuracy improved to a level comparable with that of dermatologists working without AI support.[18] The greatest absolute gains were observed in less experienced clinicians, suggesting that AI may partially compensate for the experience deficit of early-career physicians and reduce the performance gap between novice and expert practitioners. This finding has been replicated across diagnostic domains — from ECG interpretation to early leukemia detection from CBC patterns — and appears to represent a fundamental property of human-AI collaboration: the human brings contextual judgment, communication capacity, and ethical reasoning; the AI brings pattern-recognition speed, consistency across thousands of variables, and freedom from cognitive fatigue. Each compensates for the other's weaknesses.

Table 1 Summary of Key AI Applications in Family Medicine Outpatient Settings: Selected Evidence

AI Domain	Key Finding	Evidence Level	Primary Reference
Dermatology Diagnosis	CNN matches dermatologist accuracy; GP+AI outperforms GP alone	Prospective validation study	Esteva et al., 2017[6]; Tschandl et al., 2020[18]
Diabetic Retinopathy Screening	Sensitivity 87.2%, Specificity 90.7%; FDA-cleared autonomous system	Pivotal clinical trial	Abràmoff et al., 2018[7]
EHR Documentation (NLP)	30–50% reduction in documentation time per consultation	Pilot implementation studies	Quiroz et al., 2019[9]
Polypharmacy CDS	ADE risk detection: 85% sensitivity, 79% specificity	Retrospective cohort study	Kibe et al., 2023[5]
No-Show Prediction	AUC 0.82 for individual appointment no-show prediction	Prospective ML validation	Mohammadi & Wu, 2021[12]

Early Leukemia Detection	ML detects CBC patterns up to 6 months before clinical diagnosis	Retrospective cohort study	Jones & Park, 2022[8]
AI Triage	Comparable safety to nurse triage for common presentations	Comparative effectiveness study	Gottlieb & Stone, 2022[11]

The documentation literature presents the most immediately actionable findings in terms of physician wellbeing. Across multiple implementations of ambient scribing and NLP documentation support, time savings have been consistently reported at 90 minutes to three hours per physician per week. While this may appear modest in absolute terms, it translates to the recovery of 75 to 150 hours per year per clinician — time that can be redirected toward

patient care, professional development, or the restoration of personal life that burnout has eroded. Qualitative data from implementations are striking in their consistency: physicians describe feeling that they have been "given back" something important. The restoration of eye contact and in-room presence during consultations is reported as meaningful not only for physician experience but for patient satisfaction and the quality of the therapeutic relationship.

The diagnostic AI literature is compelling but reveals an important limitation: the gap between algorithm performance under controlled study conditions and performance in real-world clinical practice. The high sensitivity and specificity figures reported for diabetic retinopathy screening, for example, were obtained with standardized, high-quality fundus photography equipment and trained operators.[7] In the real-world primary care setting, where image quality may vary, staff training may be inconsistent, and patient cooperation may be imperfect, performance figures are typically somewhat attenuated. This gap between "algorithm in the lab" and "algorithm in the clinic" is a recurring theme in implementation science and must be communicated honestly to practitioners considering adoption. Tools must be validated not only in academic settings but in the community clinics where they will actually be used.

Patient outcome data — hard clinical endpoints including morbidity reduction, mortality benefit, and healthcare utilization changes — are largely absent from the current primary care AI literature. The studies reviewed rely predominantly on surrogate or process markers: improved screening rates, enhanced medication adherence metrics, reduced documentation time, and algorithmic accuracy measures. While these are reasonable proxies for eventual patient benefit, they do not yet constitute the evidentiary standard required to make AI integration a mandatory component of quality primary care. Rigorous, longitudinal randomized controlled trials conducted specifically in community primary care settings, with diverse patient populations and meaningful clinical endpoints, are urgently required. This is arguably the most significant evidence gap in the field.

The digital divide emerges as a cross-cutting concern across multiple AI domains reviewed.[19] Patients who own contemporary smartphones, maintain reliable broadband connections, and are fluent with digital interfaces are substantially better positioned to benefit from AI-powered wearables, remote monitoring platforms, patient-facing chatbots, and telemedicine AI augmentation than patients who lack these resources. Since socioeconomic disadvantage,

advanced age, rural location, and minority ethnic background all correlate independently with reduced digital access, and since these same factors also correlate with greater primary care need, there is a real risk that AI integration in primary care delivers the greatest benefits to those who need them least. Addressing this paradox must be a central preoccupation of implementation strategy.

6. Discussion

6.1. Implications for Future Family Medicine Practice

The integration of AI into family medicine does not merely add new tools to an existing practice model — it invites a fundamental rethinking of what the family physician's role should be, and what the outpatient encounter can become. The concept of "Precision Primary Care" — analogous to precision oncology but applied to the comprehensive, generalist context of primary care — captures part of this vision. AI enables the analysis of individual patient data at a level of granularity that allows care to be truly personalized: not just adjusted for the patient's current diagnoses, but responsive to their genetic profile, lifestyle patterns, environmental exposures, psychosocial circumstances, and longitudinal disease trajectory. The family physician in this model evolves from a generalist repository of standardized clinical knowledge to a chief interpreter of individualized intelligence — someone who synthesizes the outputs of multiple AI tools alongside the irreplaceable contextual knowledge gained from years of relationship with the patient.

The shift from reactive to proactive care is perhaps the most consequential implication. In the traditional model, the family physician waits for the patient to present with a problem. In the AI-augmented model, the practice's clinical intelligence systems are continuously monitoring the patient panel — analyzing wearable data, reviewing recent investigations, tracking medication refill patterns, and scanning clinical note language for early warning signals — and proactively alerting the clinical team to patients who are likely to deteriorate before they deteriorate. This shift represents a genuine paradigm change in primary care delivery, from the episodic encounter to continuous surveillance and support.

Digital phenotyping — the use of AI to extract clinically meaningful biomarkers from everyday digital behaviour — opens further possibilities. The analysis of speech patterns and linguistic features during consultations for biomarkers of depression, cognitive decline, or Parkinson's disease is an active research frontier. Smartphone-based passive sensing of physical activity, sleep patterns, and social interaction can provide longitudinal mental health data that complements periodic clinical assessment. These capabilities have obvious relevance for the family physician managing depression, anxiety disorders, and early dementia in the primary care setting.

In medical education, the AI-augmented family medicine clinic offers new possibilities for residency training. Wartman and Combs (2018) argued persuasively that medical education must transition from the Information Age to the Age of Artificial Intelligence — shifting emphasis from knowledge memorization to the cultivation of capabilities that AI cannot replicate: clinical reasoning, interpersonal skill, ethical judgment, and systems thinking.[20] The future family medicine resident will train alongside AI tools, learning both to use them effectively and to critically evaluate their outputs, developing a sophisticated calibration of when to trust the algorithm and when to trust their own clinical intuition. "Just-in-time" AI learning prompts — which surface relevant evidence-based guidelines and clinical pearls contextually during the patient encounter — represent an embedded training tool that could accelerate the development of competence across the breadth of clinical domains that family medicine requires.

6.2. Barriers to Successful Implementation

The evidence for AI's potential in family medicine is compelling, but evidence of benefit in controlled settings does not automatically translate into successful implementation in the heterogeneous reality of community practice. Several significant barriers must be honestly assessed and systematically addressed for the potential described in the literature to become the practice experienced by patients and physicians.

Interoperability is arguably the most pervasive structural barrier. The healthcare technology ecosystem is highly fragmented, with numerous competing EHR platforms, AI tools developed by third-party vendors, and legacy data systems that communicate imperfectly or not at all. An AI diagnostic tool that requires the physician to log into a separate portal, manually upload data, and then transcribe results back into the EHR does not reduce workload — it increases it. For AI to deliver on its efficiency promise, seamless integration into the primary care workflow is not a desirable feature; it is a fundamental prerequisite. Standards-based interoperability frameworks, such as HL7 FHIR (Fast Healthcare Interoperability Resources), are advancing but remain incompletely adopted across the industry.

Cost and market concentration present structural challenges for small and independent practices. Large integrated health systems and academic medical centres have the capital, the informatics infrastructure, and the change management capacity to pilot and scale AI tools. Small community practices — which account for a substantial proportion of primary care delivery in many countries — typically lack all three. If AI integration is economically feasible only for large health systems, the result will be a two-tiered primary care system in which the most well-resourced settings are further advantaged over those serving the most disadvantaged communities. This outcome would represent a failure not just of technology policy but of social equity.

Regulatory frameworks have not kept pace with the speed of AI development. The distinction between FDA-cleared clinical decision support software and general wellness applications is poorly understood by many practitioners, and the clearance pathway for AI-based diagnostic tools is still evolving. Physicians face genuine uncertainty about the liability implications of using AI tools that are not explicitly cleared for a specific indication, or of failing to use an AI tool that might have identified a diagnosis they missed. These ambiguities need regulatory clarification and, in parallel, a rethinking of how medical liability is allocated when clinical decision-making involves algorithmic co-pilots.

"Automation bias" — the tendency to over-rely on automated systems, accepting their outputs without sufficient critical evaluation — is a cognitive risk that must be proactively managed through both technology design and clinician training. History is replete with examples of automation that initially improved performance but ultimately degraded the human skills that the automation was supposed to support. If family physicians progressively cede diagnostic and clinical reasoning work to AI, the cognitive muscles required to perform these functions may atrophy — a risk that is particularly acute in training environments where residents may never develop independent clinical judgment if AI scaffolding is omnipresent from the start.

6.3. The Path Forward: Toward Equitable and Effective AI Integration

The evidence reviewed here supports cautious optimism about AI's potential in family medicine, but translating potential into reliable benefit requires deliberate and principled action across multiple levels — from the individual practice to the national policy environment.

The most urgent scientific priority is the conduct of rigorous, large-scale validation studies specifically designed for the family medicine context. Much of the existing AI evidence was generated in secondary and tertiary care settings, or in academic primary care practices not representative of community settings. Validation studies in community clinics, with diverse patient populations, across multiple countries, and with meaningful clinical outcome endpoints, are essential for establishing the evidence base that will inform adoption decisions. Family medicine research networks — with their established infrastructure for multi-site primary care research — are ideally positioned to lead this work.

Multi-stakeholder co-design of AI tools is an ethical and practical imperative. AI tools that are designed by technologists and validated by academic clinicians, then deployed in community practices serving vulnerable populations without meaningful input from those communities, are likely to fail either on uptake or on equity grounds. The participation of family physicians, practice managers, nurses, administrative staff, and — critically — diverse patient communities in the design and governance of AI tools should be a condition of development, not an afterthought of deployment. This principle of co-design extends to the ongoing governance of AI systems after deployment: community oversight bodies and patient advisory groups should have meaningful roles in monitoring AI performance and raising concerns about bias or harm.

Professional bodies for family medicine — including the American Academy of Family Physicians (AAFP), the Royal College of General Practitioners (RCGP), and the World Organization of Family Doctors (WONCA) — have a critical advocacy role to play. These organizations can develop position statements on AI standards in primary care, advocate for regulatory frameworks that protect patients without stifling innovation, promote training programmes in AI literacy for practising physicians, and ensure that family medicine perspectives — with their emphasis on equity, continuity, and person-centredness — are represented in policy discussions that have so far been dominated by hospital specialists and technology companies.

Looking toward the decade ahead, a credible vision of the 2030 family medicine clinic looks something like this: an AI system that has passively analysed the patient panel overnight flags three patients whose wearable or laboratory data suggests a need for proactive outreach before they present in crisis; the morning consultation begins from a position of pre-populated contextual intelligence rather than cold review; ambient AI generates the clinical note while the physician maintains full attention on the patient; the post-visit follow-up is managed by a clinical chatbot that escalates to the physician only when human judgment is genuinely required. In this vision, the family physician is not diminished by AI — they are liberated by it to practise the version of medicine that drew them to the specialty in the first place.

This review has traversed the landscape of AI applications in the family medicine outpatient clinic, from the technical mechanics of convolutional neural networks classifying skin lesions to the philosophical implications of algorithmic bias for health equity. What emerges from this traversal is not a picture of technology supplanting the physician, but of technology — deployed thoughtfully and equitably — enabling the physician to be more fully present, more diagnostically accurate, more proactively engaged with patient health, and less consumed by the administrative machinery that has progressively colonized clinical time.

The evidence is most compelling in specific, well-validated domains: AI dermatology tools that extend specialist-level accuracy to the general practice consulting room; autonomous diabetic retinopathy screening that brings preventive ophthalmological care to communities without ophthalmologists; ambient clinical scribes that restore the physician's gaze from screen to patient; predictive analytics that identify the COPD patient about to exacerbate or the heart failure patient heading toward readmission. In each case, the AI does not make the clinical decision — it provides better information, more quickly, so that the physician can make a better decision.

What AI manifestly cannot replicate are the irreducibly human dimensions of family medicine: the empathy expressed in a hand placed on a grieving patient's shoulder, the moral courage required to deliver a devastating diagnosis with compassion and honesty, the cultural humility needed to navigate a consultation across barriers of language and lived experience, the accumulated relational trust that makes a patient willing to disclose their darkest fears. These are not mere adjuncts to clinical competence — they are its heart. No degree of algorithmic sophistication has yet produced anything resembling genuine empathy, and no predictive model, however accurate, can replace the judgment of a physician who knows a patient across the full complexity of their life. The therapeutic relationship, which is family medicine's most powerful and least technological instrument, remains irreplaceable.

The most important call to action arising from this review is not directed at technologists or policymakers, but at family physicians themselves. The history of AI in medicine so far has been disproportionately shaped by engineers, data scientists, and hospital specialists, with primary care voices underrepresented in the conversations that matter. Family physicians who engage actively with AI development — who participate in co-design processes, who contribute to validation studies, who sit on governance committees, who advocate within professional bodies — will help ensure that these technologies are built around the values and realities of primary care rather than imposed upon it. The alternative — passivity in the face of a technological transformation that will reshape the profession regardless — is not a credible option.

The future of family medicine is not a choice between the physician and the algorithm. It is the cultivation of a collaboration between human wisdom and computational intelligence, each contributing what the other cannot. Done well, this collaboration holds the genuine promise of returning family medicine to what it was always meant to be: a continuous, comprehensive, compassionate partnership between a physician who knows you as a whole person and a system of care organized entirely around your wellbeing.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. *J Intern Med.* 2018;283(6):516–529.

[2] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56.

[3] Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan WJ, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med.* 2017;15(5):419–426.

[4] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347–1358.

[5] Kibe P, Ndege S, Karimi F. Artificial intelligence-driven detection of adverse drug events in geriatric primary care: a retrospective cohort study. *BMC Prim Care.* 2023;24(1):112.

[6] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115–118.

[7] Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *npj Digit Med.* 2018;1:39.

[8] Jones WB, Park J. Predictive value of machine learning analysis of complete blood count parameters for early leukemia detection in outpatient settings. *Am J Hematol.* 2022;97(4):450–458.

[9] Quiroz JC, Laranjo L, Kocaballi AB, Berkovsky S, Nguyen AD, Coiera E. Challenges of developing a digital scribe to reduce clinical documentation burden. *BMJ Health Care Inform.* 2019;26(1):e100040.

[10] Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural language processing of clinical notes on chronic diseases: systematic review. *JMIR Med Inform.* 2019;7(2):e12239.

[11] Gottlieb A, Stone JH. Comparative accuracy of artificial intelligence triage systems versus nursing staff in urgent primary care presentations. *Lancet Digit Health.* 2022;4(3):e185–e192.

[12] Mohammadi I, Wu H. Machine learning to predict patient no-shows in a community health centre outpatient practice. *J Am Board Fam Med.* 2021;34(3):589–598.

[13] Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff.* 2014;33(7):1123–1131.

[14] Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. *Diabetes Care.* 2019;42(8):1593–1603.

[15] Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. *J Med Internet Res.* 2019;21(4):e12887.

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

[17] Char DS, Shah NH, Magnus D. Implementing machine learning in health care — addressing ethical challenges. *N Engl J Med.* 2018;378(11):981–983.

[18] Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. *Nat Med.* 2020;26(8):1229–1234.

[19] Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc.* 2018;25(8):1080–1088.

[20] Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. *Acad Med.* 2018;93(8):1107–1109.