Chapter 2: Healthcare Use Cases

Educational Objectives

- Evaluate the role of generative AI in medical imaging, predictive diagnostics, and clinical decision support, assessing its impact on diagnostic accuracy and workflow efficiency.
- Analyze the advantages and limitations of AI-driven diagnostic tools in Radiology, Dermatology, and Gastroenterology, considering real-world case studies and implementation challenges.
- Compare the predictive capabilities of AI with traditional diagnostic methods, particularly in detecting hidden biomarkers from non-traditional sources such as retinal scans and EKGs.
- Assess the effectiveness of AI in clinical decision-making by interpreting benchmark comparisons between AI models, medical students, and board-certified physicians.
- Investigate the ethical and clinical implications of AI integration in electronic health records (EHRs), particularly in decision support systems like the "Green Button" functionality.
- Critique the impact of AI-driven patient self-diagnosis, identifying both the empowerment potential and risks associated with non-clinician use of AI-generated medical insights.
- Examine the strategies healthcare institutions use to address AI-related challenges, such as over-sensitivity in imaging, AI confidence scoring, and workflow adaptation.
- Synthesize best practices for responsible AI adoption in clinical settings, incorporating lessons from real-world hospital implementations and pilot programs.
- Formulate strategies for integrating AI-driven tools into clinical workflows while mitigating the risks of over-reliance, clinician resistance, and patient misinterpretation.
- Apply knowledge of AI advancements to propose innovative ways to enhance diagnostic accuracy, patient care, and interdisciplinary collaboration in healthcare settings.

Introduction

Generative AI is rapidly transforming the field of medicine, unlocking new capabilities for diagnosis, predictive analytics, and clinical decision-making. Unlike traditional AI models that rely on predefined rules and structured data, generative AI uses advanced neural networks to interpret complex medical data, identify patterns, and even generate new insights. These models are already proving valuable in imaging diagnostics, predictive risk modeling, and clinical decision support, offering both opportunities and challenges in their implementation.

This chapter explores the core applications of generative AI in healthcare, including its role in medical imaging, predictive diagnostics from non-traditional biomarkers, and its evolving role in clinical decision support. We will also examine the implications of AI adoption in electronic health records (EHRs) and how patients are now using AI tools to analyze their own medical data. Understanding these advancements will equip clinicians with the knowledge to integrate AI-driven insights into their practice responsibly and effectively.

AI in Medical Imaging

Deep Learning in Radiology, Gastroenterology, and Dermatology

One of the most well-established applications of AI in medicine is in medical imaging. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated impressive accuracy in identifying pathological findings in Radiology, Dermatology, and Gastroenterology. These AI-driven models are trained on millions of labeled images, allowing them to recognize intricate patterns that may be imperceptible to the human eye.

For instance, Google Health developed an AI system capable of detecting breast cancer in mammograms with greater accuracy than human Radiologists, significantly reducing false negatives and false positives. Similarly, AI-powered endoscopy tools are assisting Gastroenterologists in detecting precancerous polyps during colonoscopies, improving the early detection of colorectal cancer. Dermatologists are also leveraging AI-driven skin lesion classifiers to differentiate between benign and malignant growths, improving early melanoma detection rates.

AI-Assisted Colonoscopy for Colorectal Cancer Screening

Colorectal cancer (CRC) is the second-leading cause of cancer-related deaths worldwide, and early detection through colonoscopy is critical for improving survival rates. Computer-aided detection (CAD) systems, powered by deep learning, have been integrated into colonoscopy procedures to help gastroenterologists identify polyps that could develop into cancer.

However, recent studies, including the largest-ever clinical trial on AI-assisted colonoscopy, suggest that while CAD improves the overall detection of polyps, it does not necessarily increase the detection of advanced adenomas—the growths most likely to develop into colorectal cancer. Key findings from the research include:

- CAD increased the number of total polyps detected, particularly small polyps (<5mm in diameter), but it did not improve detection rates of advanced adenomas when used by experienced endoscopists.
- The clinical value of removing smaller polyps remains debated, as unnecessary polyp removal introduces risks such as colon perforation and bleeding.
- The effectiveness of AI-assisted colonoscopy may be greater for less experienced endoscopists, helping standardize detection rates across different skill levels.

Despite these limitations, AI-assisted colonoscopy holds promise for improving screening efficiency, particularly in regions where Gastroenterologists are in short supply. Future advancements in more sophisticated AI models could enhance detection of high-risk polyps while reducing unnecessary interventions. Moreover, real-time AI-assisted tissue classification, which would allow physicians to immediately determine whether a polyp is cancerous, could further streamline diagnosis and treatment, reducing patient anxiety and time to intervention.

As AI technology advances, its role in colorectal cancer screening is expected to expand and evolve, ultimately supporting clinicians in delivering more accurate, efficient, and safer diagnostic procedures.

Challenges of AI in Medical Imaging

Despite its successes, AI in medical imaging is not without challenges. One common issue is over-sensitivity, where AI models may flag benign lesions as malignant, leading to unnecessary biopsies and patient anxiety. In real-world applications, AI-powered imaging systems have been known to flag a high number of benign lung nodules as potentially cancerous, resulting in an increase in invasive procedures.

To address this, a more refined approach involves integrating multi-modal AI systems that combine imaging data with patient history and biomarker analysis. This enhancement reduces false positives while maintaining high sensitivity, ensuring that critical diagnoses are not overlooked while minimizing unnecessary interventions.

As AI technology advances, its role in **colorectal cancer screening and medical imaging** is expected to **expand and evolve**, ultimately supporting clinicians in delivering **more accurate, efficient, and safer diagnostic procedures**.

Predictive Diagnostics Beyond Imaging

AI’s Ability to Detect Hidden Biomarkers

Beyond imaging, AI is revealing previously unknown diagnostic capabilities by analyzing routine medical tests. Deep learning models have demonstrated an unprecedented ability to extract hidden physiological markers from EKGs and retinal scans, predicting patient characteristics such as gender, potassium levels, and even future disease risks. Researchers have developed AI systems capable of detecting early Parkinson's disease signs based on retinal scan data—an advancement that could revolutionize early intervention strategies.

One study found that AI models analyzing electrocardiograms (ECGs) were able to predict a patient's serum potassium levels with high accuracy, an insight previously thought to be undetectable without direct blood testing. Similarly, AI-powered ECG analysis has been shown to predict mortality risk, heart failure, and future cardiovascular disease events, with accuracy surpassing traditional risk models.

A particularly notable example is the AI-ECG Risk Estimator (AIRE) platform, a deep learning model trained on over a million ECGs. AIRE goes beyond traditional ECG interpretation by providing patient-specific survival curves, predicting both all-cause mortality and time-to-death from a single ECG. It demonstrated strong predictive performance across multiple international cohorts, with **C-indices ranging from 0.638 to 0.787** for various cardiovascular risks, including ventricular arrhythmia, atherosclerotic cardiovascular disease, and heart failure.

The **C-index (concordance index)** is a statistical measure used to evaluate the discriminative ability of predictive models, particularly in survival analysis. It represents the probability that, for a randomly selected pair of individuals, the model correctly predicts which patient will experience an event first. A **C-index of 0.5** indicates no predictive ability (equivalent to random chance), while a **value of 1.0** signifies perfect prediction. The observed range of **0.638 to 0.787** suggests that AIRE offers **moderate to strong predictive performance**, significantly improving risk stratification beyond traditional methods. These findings suggest that AI-driven ECG analysis could become a cornerstone of personalized medicine, offering both early warnings and actionable clinical insights.

Importantly, AI-ECG models have shown utility even in individuals with normal ECGs as assessed by cardiologists. In a validation study, patients with seemingly normal ECGs but flagged as high-risk by AIRE had significantly higher mortality rates than those classified as low-risk. This suggests that AI may detect subtle, subclinical patterns invisible to human interpretation—potentially redefining the role of routine ECGs in risk assessment.

Beyond cardiac risk, AI-derived ECG biomarkers have demonstrated associations with systemic conditions, including biological aging, metabolic disorders, and even brain health. Genome-wide association studies (GWAS) have linked AI-predicted ECG survival rates with genetic loci associated with cardiovascular structure, aging, and metabolic syndrome. Additionally, phenome-wide analyses (PheWAS) have identified correlations between AI-predicted survival and echocardiographic parameters such as left ventricular ejection fraction, atrial size, and pulmonary pressures—further validating the biological plausibility of AI-ECG predictions.

These findings raise a profound question: how much more diagnostic potential exists within existing medical tests that AI can uncover? If a standard ECG—an inexpensive and widely available test—can predict mortality, heart failure, and systemic health risks, what other latent biomarkers are hidden within routine clinical data? The ongoing integration of AI into clinical workflows promises not just improved diagnostic precision but a fundamental shift in how we assess and manage disease risk across populations.

AI in Risk Prediction

Healthcare organizations have begun to implement AI-driven retinal scan analysis tools to predict cardiovascular disease risk among patients with diabetes. These tools can help to identify microvascular changes in the retina that correlated with high cardiovascular risk. By integrating this AI assessment into routine screenings, clinicians can intervene earlier, improving patient outcomes and reducing emergency hospitalizations.

Building on these advancements, a novel deep learning framework, Eye-AD, has demonstrated that AI can detect early signs of systemic diseases, including cardiovascular conditions and neurodegenerative disorders, through retinal imaging. Using Optical Coherence Tomography Angiography (OCTA), AI has been trained to detect microvascular abnormalities that may indicate heightened disease risk before traditional symptoms appear.

The Eye-AD model, trained on 5,751 OCTA images from 1,671 participants, achieved an AUC of 0.9355 for detecting early-onset Alzheimer's disease (EOAD) and 0.8630 for mild cognitive impairment (MCI). This suggests that retinal imaging could serve as a noninvasive and cost-effective screening tool not only for Alzheimer's but also for cardiovascular risk stratification.

Retinal microvascular changes have been strongly linked to systemic conditions such as hypertension, diabetes, and atherosclerosis—all of which contribute to cardiovascular disease risk. Histopathological evidence has confirmed that Alzheimer's patients exhibit retinal microvascular deterioration, similar to cerebral small vessel disease seen in stroke patients. AI models now leverage these vascular changes to detect subclinical atherosclerosis and predict future cardiovascular events.

Additionally, a multi-center study found that AI models analyzing retinal images could detect arterial stiffening, a precursor to cardiovascular disease, with higher accuracy than traditional risk scores. The ability to integrate AI-driven retinal scans into clinical practice holds promise for transforming preventive medicine, allowing for early interventions in high-risk patients before irreversible damage occurs.

These AI tools could be refined to assess cognitive decline risks, metabolic disorders, and early-stage cardiovascular disease using retinal imaging. This convergence of AI, ophthalmology, and cardiovascular medicine represents a significant leap toward proactive, personalized healthcare.

AI vs. Human Clinicians in Decision-Making

Comparing AI to Medical Students and Physicians

One of the most controversial aspects of AI in healthcare is its ability to perform at or above the level of trained medical professionals in clinical decision-making. Large language models (LLMs) such as GPT-4, Med-PaLM, and Anthropic's Claude have been benchmarked against medical students and practicing physicians, showing high accuracy in answering complex clinical questions and providing differential diagnoses.

In 2023, a study at Stanford University tested an AI model against board-certified physicians on clinical reasoning tasks. The AI matched or exceeded the physicians in several diagnostic cases, raising ethical and practical concerns about how these models should be incorporated into patient care. Should they be used as second opinions, or could they eventually replace certain aspects of clinical diagnosis?

More recently, a 2025 randomized controlled trial published in *Nature Medicine* further explored AI's role beyond diagnostic reasoning, specifically assessing its impact on management reasoning—the ability to balance treatment decisions, testing strategies, and risk management. In the study, 92 practicing physicians were randomized into two groups: one using GPT-4 alongside conventional resources and the other relying solely on conventional resources to answer expert-developed clinical vignettes based on real patient cases.

Physicians who used AI assistance scored significantly higher on expert scoring rubrics than those who relied on traditional resources alone (mean difference = 6.5%, 95% CI = 2.7 to 10.2, $P < 0.001$). However, they also spent more time per case (mean difference = 119.3 seconds, 95% CI = 17.4 to 221.2, $P = 0.02$), suggesting that while AI improved performance, it required additional cognitive effort to integrate AI-derived insights into decision-making. Notably, there was no significant difference between physicians using AI and AI alone

(−0.9%, 95% CI = −9.0 to 7.2, P = 0.8), raising further questions about whether human clinicians meaningfully enhance AI-driven decisions or merely validate them.

These findings suggest that LLMs can enhance physician decision-making in complex cases, particularly in treatment and management reasoning, not just diagnostics. As AI models continue to evolve, the medical community must determine how to integrate them effectively—should they serve as decision-support tools to complement human expertise, or will they take on more autonomous roles in clinical practice? Validating these findings in real-world clinical settings will be critical in shaping AI's future role in medicine.

The Role of AI in Electronic Health Records (EHRs)

The "Green Button" for AI-Driven Decision Support

A major advancement in AI-assisted healthcare is the integration of AI into Electronic Health Records (EHRs). Several health systems are now deploying AI-powered decision support tools embedded directly into their EHR interfaces, aiming to enhance clinical workflows and improve patient outcomes. One such tool is the "Green Button" functionality, which allows clinicians to query aggregate patient data in real time to inform treatment decisions.

For instance, if a physician encounters a patient with an unusual set of symptoms, they can use AI-driven decision support systems (AI-CDSS) to analyze thousands of similar cases within the hospital's database, identifying trends and optimal treatment pathways. Research has shown that AI-assisted EHR decision support can improve adherence to clinical guidelines and reduce diagnostic errors by 20%, ultimately leading to more efficient and evidence-based care delivery.

Recent advancements in AI integration within EHRs extend beyond decision support. AI-powered natural language processing (NLP) algorithms are now being used to extract meaningful insights from unstructured clinical data, such as physician notes and patient history. By converting free-text entries into structured data, NLP enhances the usability of EHRs, reducing administrative burden and improving the accuracy of documentation. Additionally, deep learning models are being integrated into EHR systems to predict patient deterioration before symptoms become critical, allowing for earlier interventions.

Challenges of AI in EHRs

Despite its benefits, AI integration in EHRs faces significant challenges. One major concern is clinician trust and workflow integration. Many physicians report frustration with AI recommendations that appear to contradict their clinical judgment, leading to hesitation in relying on AI-assisted decision-making. This challenge highlights the broader issue of AI interpretability, often referred to as the “black box” problem.

To address this, some health systems have introduced AI confidence scoring systems, providing clinicians with transparency into how certain the AI is about its recommendations. This allows physicians to weigh AI insights alongside their own expertise, leading to a more collaborative approach to clinical decision-making. Additionally, ongoing research suggests that human-centered design principles should be applied to AI-CDSS interfaces to ensure they align with physician workflows rather than disrupt them.

Another challenge is the risk of AI bias, which arises when algorithms are trained on non-representative datasets. Studies have found that AI models trained on historically biased healthcare data may reinforce disparities, disproportionately affecting certain patient populations. To mitigate this risk, healthcare institutions must implement rigorous validation and auditing frameworks to ensure AI models are fair, unbiased, and generalizable across diverse demographics.

The Rise of Patient-Driven AI Use

Patients Using AI to Review Their Own Data

Perhaps one of the most unexpected developments in generative AI is its increasing adoption by patients. Traditionally, AI in healthcare has been clinician-centric, but new trends suggest that patients are actively using AI tools to analyze their own medical records.

For example, some parents are now leveraging AI to review their children’s medical histories, often cross-referencing symptoms with published case studies to challenge initial physician assessments. A particularly compelling case in 2023 involved Courtney, a mother who used AI to correctly diagnose her 4-year-old son, Alex, with tethered cord syndrome after 17 physicians over three years failed to identify the root cause of his symptoms.

Alex’s symptoms first appeared after his parents purchased a bounce house during the COVID-19 lockdown. He began experiencing chronic pain, which initially seemed to resolve with Motrin but later escalated into a complex array of symptoms. His dental issues—chewing on objects and suspected teeth grinding—led them to consult a dentist, who ruled out dental pathology but referred him to an orthodontist due to concerns about airway obstruction and a narrow palate. A palatal expander was placed, and for a brief period, his symptoms improved.

However, his condition worsened, with growth stalling, headaches intensifying, and gait abnormalities emerging. His right leg became dominant, and his left leg was dragged behind when walking. Despite consulting multiple specialists—including a neurologist, ENT, pediatrician, internist, and musculoskeletal specialist—no unifying diagnosis was reached. Instead, each specialist focused on their area of expertise, addressing isolated symptoms but failing to connect the clinical picture holistically.

The breakthrough came when Courtney, frustrated by the lack of answers, turned to ChatGPT. She painstakingly inputted his symptoms, MRI findings, and clinical history into the AI model, highlighting subtle but crucial details—such as his inability to sit crisscross applesauce—that suggested an underlying structural abnormality. ChatGPT proposed tethered cord syndrome, a condition where the spinal cord is abnormally attached to surrounding tissues, leading to progressive neurological dysfunction.

With this AI-suggested diagnosis in hand, Courtney joined support groups for families affected by tethered cord syndrome, where she found striking similarities to Alex's case. She then sought out a new neurosurgeon, who immediately recognized the condition upon reviewing his MRI. The specialist identified occulta spina bifida and confirmed that Alex's spinal cord was tethered, a diagnosis that had previously been overlooked.

Tethered cord syndrome, particularly in cases of spina bifida occulta, can be challenging to diagnose because external signs—such as birthmarks, skin dimples, or tufts of hair over the spine—are often subtle and easily missed. In young children, symptoms like chronic pain, gait abnormalities, and bladder dysfunction can be misattributed to behavioral or developmental delays rather than a neurosurgical condition. The diagnosis is typically made through a combination of MRI imaging and clinical presentation, but given the variability in symptoms, many cases go undetected until significant neurological impairment occurs.

Alex underwent surgical intervention to detach the tethered spinal cord, preventing further neurological deterioration. While his postoperative recovery is ongoing, his family is optimistic about his long-term prognosis.

This case underscores both the potential and the limitations of AI in clinical diagnosis. While AI tools like ChatGPT may help identify rare or overlooked conditions by synthesizing vast amounts of data, they also lack the ability to independently validate findings or assess clinical nuance. This raises ethical and practical concerns about how AI should be integrated into patient care—not as a replacement for clinicians, but as a tool to enhance differential diagnosis, reduce diagnostic delays, and empower patients to advocate for thorough evaluations.

AI-driven health chatbots and personalized risk prediction tools are also becoming more common. Many patients now rely on AI-powered applications that scan EHR data, genetic markers, and lifestyle habits to provide individualized health assessments. This shift has profound implications for shared decision-making, as AI-literate patients may come to clinical encounters with AI-generated insights, expecting physicians to validate or refute them.

Ethical Considerations and Risks

While patient-driven AI use can empower individuals to advocate for better care, it also introduces several risks. One major concern is the potential for misinterpretation of AI-generated findings. Patients without medical training may misunderstand AI recommendations, leading to unnecessary anxiety, self-diagnosis errors, or even delays in seeking proper medical care.

Another issue is the potential strain on physician-patient relationships. If AI findings contradict a doctor's professional assessment, patients may lose trust in their healthcare providers. This growing trend raises an important question: How should physicians address AI-generated patient inquiries in clinical practice?

To mitigate these challenges, hospitals and health organizations must develop clear guidelines for AI-literate communication. Clinicians should be trained to contextualize AI-driven insights, helping patients differentiate between AI-generated hypotheses and clinically validated diagnoses. Furthermore, AI developers must prioritize transparency, ensuring that AI tools designed for patient use clearly explain their limitations and do not provide misleading medical advice.

Conclusion

Generative AI is no longer a future concept—it is actively reshaping healthcare delivery today. From radiology and dermatology to predictive diagnostics and EHR integration, AI is augmenting clinical workflows, improving diagnostic accuracy, and even empowering patients to take an active role in their healthcare. However, these advancements come with challenges, including over-reliance on AI, physician trust concerns, and ethical considerations regarding patient use. By understanding these real-world applications and their implications, healthcare professionals can better integrate AI into their practice in a way that enhances care quality while maintaining professional oversight.

References

Ardila, D., Kiraly, A. P., Bharadwaj, S., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest CT. *Nature Medicine, 25*(6), 954-961.

Goh, E., Gallo, R.J., Strong, E. et al. GPT-4 assistance for improvement of physician performance on patient care tasks: a randomized controlled trial. Nat Med (2025). https://doi.org/10.1038/s41591-024-03456-y

Google Health. (2020). AI system outperforms radiologists in breast cancer detection. Retrieved from https://health.google.com

Hao, J., Kwapong, W.R., Shen, T. *et al.* Early detection of dementia through retinal imaging and trustworthy AI. *npj Digit. Med.* **7**, 294 (2024). https://doi.org/10.1038/s41746-024-01292-5

Holohan, M. (2023). A boy saw 17 doctors over 3 years for chronic pain. ChatGPT found the diagnosis. Retrieved from https://www.today.com/health/mom-chatgpt-diagnosis-pain-rcna101843

McKinney, S. M., Sieniek, M., Godbole, V., et al. (2020). International evaluation of an AI system for breast cancer screening. *Nature, 577*(7788), 89-94.

Reynolds, S. (2023). Is AI Ready to Play a Leading Role in Colorectal Cancer Screening? National Cancer Institute. Retrieved from https://www.cancer.gov/news-events/cancer-currents-blog/2023/colonoscopy-cad-artificial-intelligence

Sau, Arunashis et al (2024). Artificial intelligence-enabled electrocardiogram for mortality and cardiovascular risk estimation: a model development and validation study. The Lancet Digital Health, Volume 6, Issue 11, e791 - e802