

VOLUMEN 17 • NÚMERO 32
NÚMERO ESPECIAL: FILOSOFÍA DE LA MEDICINA

*euph*Ψία

REVISTA DE FILOSOFÍA

PRESENTACIÓN

Introducción de los editores, Mario Gensollen y Alger Sans Pinillos

VARIA

La abducción en el razonamiento médico, Cristina Barés Gómez y Matthieu Fontaine

Jeopardizing Biomedical Epistemic Niches, Lorenzo Magnani

Revolución en los modelos sanitarios: diseño, complejidad e instituciones, Anna Estany

Challenges and Controversies of Generative AI in Medical Diagnosis, Jordi Vallverdú

El lugar de la cirugía en la filosofía de la medicina, Cecilia M. Calderón Aguilar

DISCUSIÓN

Ética de la innovación médica, Antonio Sitges-Serra

Análisis y definición de los conceptos de *salud y enfermedad*, Adreu Segura



ISSN 2683-2518

Challenges and Controversies of Generative AI in Medical Diagnosis

Jordi Vallverdú
ICREA Academia
Universitat Autònoma de Barcelona
jordi.vallverdu@uab.cat

Resumen

Este artículo proporciona una exploración exhaustiva del papel transformador de los modelos de inteligencia artificial generativa, específicamente Redes Generativas Antagónicas (GAN) y Autoencoders Variacionales (VAE), en el ámbito del diagnóstico médico. Basándose en la filosofía de la medicina y la epidemiología, el artículo examina las dimensiones técnicas, éticas y filosóficas de la integración de modelos generativos en la atención médica. Un estudio de caso con Emily resalta el crucial apoyo que la inteligencia artificial generativa puede ofrecer en diagnósticos médicos complejos. La discusión se extiende a la aplicación de GAN y VAE en la imagen médica, enfatizando su potencial para mejorar diagnósticos, planificación de tratamientos e investigación médica. El artículo profundiza además en desafíos y controversias, abordando problemas de precisión anatómica, sesgos en datos de entrenamiento, interpretabilidad de imágenes médicas generadas por

inteligencia artificial y consideraciones éticas, como el fenómeno de “Dr. Google” y sus implicaciones para el autodiagnóstico, especialmente en el contexto del creciente papel de los modelos de inteligencia artificial generativa en la atención médica. La sección final enfatiza la necesidad de alfabetización en salud, el uso responsable de la información en línea y la toma de decisiones colaborativa entre pacientes y proveedores de atención médica. Abogamos por colaboraciones interdisciplinarias para establecer pautas éticas y garantizar un uso responsable de la inteligencia artificial en la atención médica.

Palabras clave: inteligencia artificial generativa; ChatGPT; diagnóstico médico; Dr. Google; autodiagnóstico; ética en atención médica; McDonalización de la atención médica.

Abstract

This paper provides a comprehensive exploration of the transformative role of generative AI models, specifically Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), in the realm of medical diagnosis. Drawing from the philosophy of medicine and epidemiology, the paper examines the technical, ethical, and philosophical dimensions of integrating generative models into healthcare. A case study featuring Emily underscores the pivotal support generative AI can offer in complex medical diagnoses. The discussion extends to the application of GANs and VAEs in medical imaging, emphasizing their potential in improving diagnostics, treatment planning, and medical research. The paper further delves into challenges and controversies, addressing issues of anatomical accuracy, biases in training data, interpretability of AI-generated medical images, and ethical considerations, for example, the phenomenon of "Dr. Google" and its implications for self-diagnosis, particularly in the context of the increasing role of generative AI models in healthcare. The concluding section emphasizes the need for health literacy, responsible use of online

information, and collaborative decision-making between patients and healthcare providers. We advocate for interdisciplinary collaborations to establish ethical guidelines and ensure responsible AI use in healthcare.

Key words: Generative AI; ChatGPT; Medical Diagnosis; Dr. Google; Self-Diagnosis; Healthcare Ethics; Healthcare's McDonalization.

1. Introduction

In the evolving landscape of healthcare and public health, medical diagnosis is poised for a revolution. Artificial intelligence, particularly generative models, presents an unprecedented opportunity to transform the way we approach diagnostics. This introduction outlines the significance of this convergence within the broader contexts of philosophy of medicine and epidemiology.

Meet Emily, a 38-year-old woman who recently experienced a medical conundrum that showcases the pivotal role of generative AI in healthcare. Emily's case serves as a compelling backdrop to delve into the profound implications of generative AI in medical diagnosis. Emily began noticing unusual skin rashes and intermittent fevers, which she found concerning. Her first step was to consult a healthcare chatbot —an AI-driven virtual assistant accessible through her smartphone. She described her symptoms to the chatbot, and it provided initial recommendations and potential diagnoses based on its programmed knowledge. However, the chatbot was not able to definitively identify the underlying cause of Emily's symptoms due to their complexity and atypical presentation. Recognizing the limitations of AI, Emily promptly scheduled an appointment with her primary care physician. This is where generative AI enters the equation. In the clinic, the doctor used a generative AI system to assist in Emily's

diagnosis. By inputting Emily's medical history, symptoms, and available test results, the generative AI model could analyze her case comprehensively. Generative AI models, such as GANs and VAEs, excel at data synthesis and pattern recognition. In this case, they can assist in identifying complex patterns and correlations within Emily's symptoms and medical history that might not be immediately apparent to a human clinician. By simulating various potential diagnoses and outcomes, generative AI can help guide the healthcare provider in their diagnostic decision-making. The generative AI system generated a range of possible diagnoses and ranked them based on their likelihood, taking into account both common and rare conditions. This data-driven approach allowed Emily's physician to consider a broader spectrum of potential diagnoses, including some that may not have been on their radar. In Emily's case, generative AI played a pivotal role in supporting her diagnosis by augmenting the clinical expertise of her healthcare provider. The integration of generative AI helped in the evaluation of her complex symptoms and contributed to a more accurate diagnosis.

In a world where healthcare systems grapple with escalating demands, expanding populations, and emerging health threats, the need for precise, efficient diagnostic tools has never been more pressing (Hirosawa et al., 2023). Timely and accurate diagnoses underpin effective patient care, treatment strategies, and disease prevention. Generative AI models, exemplified by Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), are at the forefront of this transformation. These models, renowned for their data synthesis capabilities, offer a myriad of applications in healthcare. They hold the power to revolutionize medical imaging by enhancing radiological and pathological

assessments. They facilitate disease classification and aid in epidemiological research, all while reshaping the landscape of medical diagnosis. Generative models, as a class of AI algorithms, are designed to generate data that closely mimics existing datasets. They accomplish this by learning the underlying structure of the data, enabling them to create new data points that adhere to this structure. In healthcare, generative models are particularly valuable for their role in generating synthetic medical images and improving the quality of real-world datasets.

The philosophical perspective, embedded in the philosophy of medicine, delves into the epistemological and ethical dimensions of medical practice (Vallverdú, 2016). It prompts questions about how knowledge is acquired, validated, and applied in healthcare. With the integration of generative models, the very nature of medical knowledge is being redefined, raising inquiries into the essence of disease, the dynamics of patient-doctor relationships, and the ethical underpinnings of healthcare. Epidemiology, the science of understanding disease patterns, causality, and prevention, plays a pivotal role in the context of generative AI. These models augment disease surveillance, refine modeling, mitigate bias, and contribute to the assessment of causality, fundamentally altering the landscape of epidemiological research.

This paper embarks on a comprehensive exploration of the multifaceted dimensions of generative AI in medical diagnosis, informed by the critical considerations of philosophy of medicine and epidemiology. Through the examination of technical and ethical aspects, we aim to provide an encompassing understanding of the transformative potential and challenges that

arise with the integration of generative models into healthcare and public health.

2. Generative Models in Medical Imaging

Medical imaging plays a pivotal role in modern healthcare for diagnostic and research purposes. The development of generative models, particularly Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), has introduced innovative approaches to generate realistic and high-fidelity medical images. This section explores the utilization of these generative models in the context of medical image generation, highlighting their contributions, challenges, and potential applications.

Medical image generation is a critical component in medical research, education, and diagnostics. Traditional methods often rely on simulation techniques or manual creation of synthetic images, which may lack the complexity and diversity found in real-world medical data. Generative models, such as GANs and VAEs, have emerged as powerful tools for addressing these limitations by learning and replicating the underlying data distribution. Generative Adversarial Networks (GANs), introduced by Goodfellow et al. (2014), have gained popularity for their ability to generate realistic images through an adversarial training process. In the medical imaging domain, GANs have been employed to synthesize images of various modalities, including X-rays, MRIs, and CT scans. The generator and discriminator components of GANs collaborate to produce images that closely resemble authentic medical data. Despite their success, GANs face challenges in generating medically relevant images with sufficient

detail and anatomical accuracy. Addressing these challenges often involves fine-tuning model architectures, incorporating domain-specific knowledge, and optimizing training parameters. Additionally, the potential for the generation of artifacts and unrealistic features requires careful consideration.

Variational Autoencoders (VAEs), AEs, proposed by Kingma and Welling (2013), are another class of generative models that utilize probabilistic frameworks for image generation. VAEs are particularly well-suited for capturing latent structures in data, making them valuable in the synthesis of medical images. One key advantage of VAEs is their ability to learn meaningful latent representations of medical images. This facilitates the generation of diverse images by manipulating latent variables, allowing for controlled exploration of the data space. This property is particularly beneficial in applications such as anomaly detection and data augmentation.

The COVID-19 pandemic has brought attention to challenges and pitfalls associated with the utilization of artificial intelligence (AI) in medical applications, particularly in the context of diagnosing and prognosticating COVID-19 through medical imaging. The paper titled “Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans” addresses key issues in this domain (Roberts et al., 2021; Heaven, 2021). These include concerns related to data quality and bias, emphasizing the necessity for diverse and representative datasets to ensure robust generalization across different populations. The study also delves into challenges associated with the translation of AI models from research settings to real-world clinical environments, emphasizing the importance of rigorous

testing and validation in diverse clinical contexts. Furthermore, the paper underscores the interpretability and explainability of AI models, advocating for transparent models to establish trust among healthcare professionals and patients. Ethical considerations, encompassing patient privacy, consent, and fairness in algorithmic decision-making, are highlighted. The study also explores practical challenges in the clinical implementation of AI tools, stressing the need for seamless integration with existing healthcare systems and protocols. Continuous monitoring and improvement of AI models are emphasized to adapt to changes in the disease, medical practices, or emerging research findings, underlining the iterative refinement required for sustained effectiveness. The paper underscores the importance of a cautious, evidence-based approach, considering AI as a complementary tool rather than a definitive solution in the dynamic and rapidly evolving landscape of global health crises. Collaborative efforts between medical professionals, researchers, and AI developers are paramount for navigating these challenges and ensuring the responsible and effective utilization of AI in healthcare.

The use of generative models in medical image generation extends beyond research and education, influencing various applications in clinical practice. From generating synthetic datasets for training robust machine learning algorithms to aiding in the creation of patient-specific simulations, the impact of GANs and VAEs is profound. As these models continue to evolve, future directions may involve enhancing interpretability, addressing ethical considerations, and expanding their utility in personalized medicine. Generative models, particularly GANs and VAEs, have demonstrated significant potential in generating realistic and

diverse medical images. While challenges persist, ongoing research and advancements in model architectures are likely to further enhance their applicability in medical imaging, ultimately contributing to improved diagnostics, treatment planning, and medical research.

On the other hand, generative models offer a potent solution for data augmentation in medical imaging datasets, a crucial aspect in training robust machine learning models. By generating synthetic images that closely mimic the characteristics of real medical data, GANs and VAEs contribute to the enrichment and diversification of training sets. This aids in overcoming challenges related to limited and imbalanced datasets, ultimately enhancing the generalization capabilities of machine learning algorithms applied in medical image analysis. The inherent capacity of generative models to discern patterns within data makes them effective tools for image denoising in medical imaging. The ability to reconstruct clean and high-fidelity images from noisy input data is invaluable for improving the quality of diagnostic images. GANs, in particular, excel in removing noise while preserving relevant anatomical details, thereby aiding clinicians in accurate interpretation and diagnosis. In medical imaging, where details are paramount, the application of generative models for super-resolution has garnered attention. GANs and VAEs can enhance the resolution of medical images, providing finer details that may be critical for diagnosis and treatment planning. This capability opens avenues for improved visualization and analysis, contributing to the overall efficacy of medical imaging techniques.

The integration of generative models into medical imaging challenges conventional notions of medical knowledge acquisition. As these models operate as complex, data-driven

systems, their ability to learn and generate representations may surpass the explicit understanding of the underlying biological and physiological processes. This challenges the traditional paradigm of medical knowledge rooted in human-expertise-driven interpretations, raising questions about the interpretability and reliability of AI-generated medical images.

The reliance on generative models for medical image generation and analysis introduces ethical considerations, particularly in the context of AI-based diagnoses. The trustworthiness and reliability of diagnoses made by these models become central concerns. The interpretability of model decisions, potential biases in training data, and the accountability of AI systems in healthcare demand careful ethical scrutiny. Balancing the benefits of improved diagnostics with the need for transparency and ethical standards becomes pivotal as generative models continue to permeate medical imaging practices. As the field of generative models in medical imaging evolves, ethical implications must be addressed proactively. Striking a balance between innovation and patient safety, ensuring transparency in model decision-making, and establishing clear ethical guidelines are imperative. Engaging in interdisciplinary collaborations involving clinicians, ethicists, and technologists can contribute to the development of responsible and ethically sound practices in the application of generative models in medical imaging.

Generative models, with their multifaceted applications in data augmentation, denoising, and super-resolution, bring transformative potential to medical imaging. However, their integration prompts us to critically examine the philosophical and ethical dimensions of these advancements. By navigating the controversies and addressing ethical considerations, the medical

community can harness the benefits of generative models while ensuring the responsible and ethical use of AI in the pursuit of improved healthcare outcomes.

3. Diagnostic Applications

The integration of generative artificial intelligence (AI) tools, with a particular emphasis on the widely recognized ChatGPT, into medical diagnostic practices has ushered in a paradigm shift that presents both challenges and opportunities within the field of healthcare (Garg et al., 2023). The opportunities lie in the potential enhancement of diagnostic capabilities, efficiency, and accessibility (Kulkarni, 2023).

Generative AI, when harnessed by medical professionals, provides an avenue for rapid analysis of vast datasets, aiding in the identification of complex patterns associated with various medical conditions. Moreover, the user-friendly interfaces of platforms like ChatGPT open up opportunities for lay individuals to engage in self-diagnosis, fostering a sense of empowerment and proactive health management. However, these opportunities are accompanied by formidable challenges. One significant challenge revolves around the need for robust validation and rigorous testing to ensure the reliability and accuracy of generative AI models in diverse clinical scenarios. The interpretability of AI-generated diagnoses poses another hurdle, necessitating transparent frameworks to instill confidence among healthcare professionals and mitigate the risk of misinterpretation. Additionally, issues of data privacy, security, and ethical considerations loom large, especially when considering the

sensitive nature of medical information. The standardization of diagnostic processes through generative AI may inadvertently contribute to oversimplification and the overlooking of nuanced aspects inherent in individual patient cases. Furthermore, concerns about biases in training data and potential disparities in health outcomes must be diligently addressed to prevent unintended consequences. While the use of generative AI, exemplified by ChatGPT, holds promise for revolutionizing medical diagnostics, it is imperative to approach its implementation with a balanced consideration of the associated challenges. A collaborative effort between healthcare professionals, technologists, and ethicists is vital to harness the full potential of generative AI in medical diagnosis while ensuring the delivery of accurate, ethical, and patient-centered healthcare.

In (Kuroiwa et al, 2023), the objective was to evaluate ChatGPT's performance in self-diagnosing prevalent orthopedic conditions, encompassing carpal tunnel syndrome (CTS), cervical myelopathy (CM), lumbar spinal stenosis (LSS), knee osteoarthritis (KOA), and hip osteoarthritis (HOA), while also examining the degree to which it recommends medical consultations. Over a 5-day period, the study authors consistently posed identical questions to ChatGPT, categorizing responses based on correctness, and assessing reproducibility between days and raters. Notably, the ratios of correct answers varied across conditions, with CTS achieving a perfect score, CM demonstrating a notable disparity, and LSS, KOA, and HOA exhibiting differing levels of accuracy. Importantly, responses recommending medical attention were categorized by the strength of the recommendation, revealing variations in the utilization of phrases such as "essential," "recommended," "best," and "important." While

ChatGPT exhibited potential as an initial step in healthcare access, the study underscored inconsistencies in accurate self-diagnosis, emphasizing the necessity of refining symptom identification and incorporating clear language in natural language processing (NLP) systems to prompt users to seek expert medical opinions. The findings advocate for further investigation into the application of AI, such as ChatGPT, in clinical studies to comprehensively elucidate its role and potential enhancements in healthcare practices. (Menen et al., 2023) explored the diagnostic accuracy of ChatGPT as a medical information resource for common and rare diseases in the context of laypeople engaging in self-diagnosis, reflecting the increasing trend of seeking online medical advice. Assessing 50 clinical case vignettes, including 10 rare presentations, the findings demonstrated that ChatGPT 4 effectively diagnosed common cases within two suggested diagnoses when utilized by lay individuals for self-diagnosis. However, for rare diseases, the model required eight or more suggestions to solve 90% of cases. Notably, ChatGPT 4 consistently outperformed ChatGPT 3.5 in terms of diagnostic accuracy within this user group. The study also included a comparison between ChatGPT and human medical doctors. The authors suggested that while ChatGPT showed promise as a diagnostic tool to assist lay individuals in self-diagnosing challenging cases, caution should be exercised in non-professional use, despite its commendable diagnostic accuracy.

Moreover, it becomes evident that, notwithstanding the inherent capabilities of natural language frameworks, there is a critical need to master the art of prompting effectively to yield precise results. Learning to formulate queries in a manner that optimally leverages the language models is imperative for

achieving accurate diagnostic outcomes (Caruccio et al., 2024). Additionally, the authors provided several interesting ideas and tools: the introduction of a novel processing pipeline designed for seamless interaction with language models, transforming them into diagnostic decision support systems; the development of an innovative prompt engineering methodology tailored specifically for engaging with language models, ensuring precision in diagnostic results; a comparative analysis evaluating the efficacy of traditional machine learning models in providing intelligent diagnoses for low-risk diseases; an assessment comparing the performance metrics, including accuracy, precision, recall, and F1-score, between traditional machine learning approaches and ChatGPT in intelligent diagnosis; a comparative evaluation of the diagnostic capabilities of ChatGPT against Google BARD; a performance analysis of ChatGPT in comparison with domain-specific NLP models; and the creation of a new bot employing the most effective predictive models, capable of interacting with individuals and offering preliminary diagnoses based on their presented symptoms. These contributions collectively advance our understanding of the nuanced interplay between language models, prompt formulation, and diagnostic precision, paving the way for improved utilization of artificial intelligence in medical contexts.

(Walker et al., 2023) investigated ChatGPT-4, the latest version of an artificial intelligence (AI) chatbot designed to respond to freely formulated and intricate questions. With the potential to emerge as a future standard for accessing medical information, particularly for health care professionals and patients, the research aimed to assess the reliability of the medical information provided by ChatGPT-4. Focusing on five hepato-

pancreatico-biliary (HPB) conditions with high global disease burdens, the study employed the Ensuring Quality Information for Patients (EQIP) tool, comprising 36 items divided into content, identification, and structural data subsections. Additionally, five guideline recommendations per condition were reformulated as questions and posed to ChatGPT, with agreement between guidelines and AI responses independently evaluated. The results revealed that ChatGPT-4 provided medical information comparable to static internet sources, as reflected in the median EQIP score and the agreement between guideline recommendations and ChatGPT responses. Despite the current limitations, the findings suggested that large language models, such as ChatGPT, had the potential to become a standard source for patients and health care professionals seeking medical information in the future.

In conclusion, the incorporation of generative artificial intelligence, particularly exemplified by ChatGPT, into the landscape of medical diagnostics signifies a transformative paradigm shift in healthcare. This shift is marked by promising opportunities for enhanced diagnostic capabilities, efficiency, and accessibility, as demonstrated by the rapid analysis of extensive datasets and the empowering potential for self-diagnosis among individuals. However, these opportunities are met with formidable challenges that necessitate careful consideration and strategic navigation.

The challenges encompass the imperative for rigorous validation and testing to ensure the reliability and accuracy of generative AI models across diverse clinical scenarios. The interpretability of AI-generated diagnoses emerges as a critical hurdle, requiring transparent frameworks to instill confidence

among healthcare professionals and mitigate the risk of misinterpretation. Issues of data privacy, security, and ethical considerations loom large, given the sensitive nature of medical information. Moreover, the standardization of diagnostic processes through generative AI introduces the risk of oversimplification, potentially overlooking nuanced aspects inherent in individual patient cases. Concerns about biases in training data and the potential for disparities in health outcomes must be vigilantly addressed to avert unintended consequences.

While the integration of generative AI, exemplified by ChatGPT, holds immense promise for revolutionizing medical diagnostics, its implementation necessitates a balanced and collaborative approach. Healthcare professionals, technologists, and ethicists must work in concert to harness the full potential of generative AI in medical diagnosis. This collaborative effort is crucial to ensure the delivery of accurate, ethical, and patient-centered healthcare in the dynamic landscape shaped by the advancements in artificial intelligence.

4. Other Controversial Aspects of Generative AI in Medical Scenarios

The rapidly advancing landscape of generative artificial intelligence (AI), exemplified by the ChatGPT chatbot, has prompted considerable interest in its potential application within the healthcare domain. Notably, ChatGPT's remarkable performance on the United States Medical Licensing Exam (USMLE) has ignited discussions about its possible integration into healthcare practices (Sallam, 2023). However, a critical issue

has emerged regarding ChatGPT's role in authorship, as it has been credited as a co-author on scientific papers. The study conducted by Ide, Hawke, and Nakayama (2023) systematically evaluates whether ChatGPT aligns with the authorship criteria set forth by the International Committee of Medical Journal Editors (ICMJE). The comprehensive analysis undertaken in this study unveils that while ChatGPT demonstrates the capability to fulfill certain authorship criteria, it falls short in meeting others, particularly those pertaining to final approval and accountability. The researchers, therefore, assert that, based on the current criteria set by the ICMJE, it is inappropriate for ChatGPT to be designated as an author. The study underscores the paramount importance of transparency, advocating for explicit acknowledgment of ChatGPT's involvement in research studies. Discussions within the commentary delve into potential shifts in authorship criteria, aligning with the perspective of the World Association of Medical Editors, which posits that chatbots such as ChatGPT are ineligible for authorship roles. Additionally, the study cites the Science Family of Journals' stringent policy, explicitly forbidding the use of AI-generated text or figures without explicit permission and prohibiting AI programs from assuming authorship roles, with a violation considered scientific misconduct.

(Cox, 2023) delved into the intricate terrain of causality in epidemiology by engaging in a Socratic dialogue with ChatGPT, a prominent large language model (LLM). The focus was on discussing the interpretation of epidemiological associations between fine particulate matter (PM2.5) and human mortality risks. Reflecting common patterns of human reasoning and argumentation, ChatGPT initially asserted that "It is well-established that exposure to ambient levels of PM2.5 does increase

mortality risk” and emphasized the importance of reducing PM2.5 exposure for public health. Fine particulate matter (PM2.5) refers to tiny particles or droplets in the air that are 2.5 micrometers or smaller in size. These particles can originate from various sources such as vehicle emissions, industrial processes, and natural sources like wildfires. However, through systematic questioning, the dialogue revealed a nuanced shift as ChatGPT acknowledged that “It is not known with certainty that current ambient levels of PM2.5 increase mortality risk.” Despite recognizing strong evidence of an association, ChatGPT underscored the uncertainty about causation due to potential omitted confounders. This dialogue exemplified the complex and controversial nature of employing generative AI in medical scenarios, raising questions about the reliability and interpretation of its conclusions within the hotly debated field of causality in epidemiology (related directly to legal aspects of these practices; Vallverdú, 2022). The utilization of generative AI in medical diagnosis introduces several epistemic challenges for understanding causality, for example in observational studies (Golinelli et al., 2023; Sanmarchi et al., 2023). Interpretability and explainability are compromised due to the inherent complexity of large language models, making it difficult to discern the model’s decision-making process. Omitted confounders, variables not considered in the model, pose a risk of inaccurate causal inferences. The quality and representativeness of training data, along with biases inherent in the data, can skew causal relationships. Generative AI models may struggle to express and handle uncertainty associated with medical diagnoses, where probabilistic reasoning is common. The dynamic nature of medical knowledge, continually evolving with new discoveries, presents a challenge for AI systems to stay current.

Additionally, the interaction between humans and AI, as well as the level of trust users place in AI-generated information, can impact the accurate understanding of causal relationships in medical contexts.

ChatGPT, a leading force in AI technology, holds immense promise for transforming clinical medicine, impacting data analysis, clinical trial recruitment, and patient education (Eysenbach, 2023). Its potential spans assisting in basic research, from drug discovery to disease prediction and therapeutic target assessment. Challenges arise from ChatGPT's limitations, specifically in updating real-time training data and providing only general answers. Concerns include its incapacity to offer detailed and comprehensive diagnoses, particularly for common symptoms, as well as its inability to pass life-support exams without specific medical training, casting doubt on its readiness for critical clinical applications.

In clinical applications, reliance on ChatGPT for diagnoses poses risks of inaccuracy or delayed treatment. Its application in human-computer interaction, particularly in mental health, holds potential for improving usability. Yet, controversies emerge, such as its use in medical education and writing (Boscardin et al., 2023), raising ethical concerns and journal policies. Some scientific journals have stipulated the need for clear attribution of ChatGPT-generated content, while others, like *Nature*, have refused to accept it as an author due to concerns about responsibility for generated content. Despite its potential to enhance decision-making, reduce errors, and offer personalized treatment recommendations, AI, including ChatGPT, demands cautious consideration due to privacy issues and biases. While it has the power to revolutionize clinical and translational medicine, a balanced approach is

necessary, leveraging its capabilities while addressing potential risks and negative impacts (Xue et al., 2023).

5. Dr. Google and Self-Diagnosis

“Dr. Google” is a colloquial term used to describe the phenomenon where individuals turn to internet search engines, particularly Google, to seek information about medical conditions, symptoms, and potential diagnoses. This practice has become increasingly common in recent years, reflecting the growing accessibility of health-related information online. While the term is not an official medical or scientific designation, it captures the widespread tendency for people to use online resources for self-diagnosis and health information. “Dr. Google” encapsulates the modern phenomenon where individuals, armed with readily available information on the internet, engage in self-diagnosis (Lee et al., 2014; 2015). This section examines the motivations behind this behavior and its impact on the traditional doctor-patient relationship. In doing so, we delve into the philosophical aspects of patient autonomy and the implications for informed decision-making. The advent of generative models in medical diagnosis has ushered in a new era of precision and efficiency in healthcare. Simultaneously, the rise of “Dr. Google” as a ubiquitous source for self-diagnosis has prompted a reevaluation of the dynamics between patients, healthcare providers, and information accessibility. This section seeks to bridge the gap between these two phenomena and uncover the underlying philosophical and medical implications.

Patient autonomy, a cornerstone of medical ethics, underscores the right of individuals to make decisions about their own health. However, the advent of “Dr. Google” introduces a nuanced dynamic, where patients, armed with information from various sources, navigate the delicate balance between empowerment and potential misinformation. The philosophical discourse on patient autonomy becomes a critical lens through which we evaluate the implications of self-diagnosis. Patient autonomy, rooted in the principle of informed consent, emphasizes an individual’s right to be involved in decisions about their healthcare. The advent of the internet and search engines has exponentially increased the accessibility of health-related information. Patients can now actively seek information about symptoms, conditions, and treatment options, empowering them to engage more actively in discussions with healthcare providers. This aligns with the idea that informed patients are better equipped to make decisions that align with their values and preferences. The empowerment facilitated by information access, however, is accompanied by the challenge of navigating a vast sea of information. “Dr. Google” introduces a nuanced dynamic where patients must discern the reliability of sources, interpret complex medical information, and differentiate between credible information and potential misinformation. This delicate balance requires patients to critically evaluate the information they encounter, emphasizing the need for health literacy and the ability to distinguish between evidence-based knowledge and anecdotal or unreliable content.

Patient autonomy, as a cornerstone of medical ethics, engages in a philosophical discourse that questions the nature and limits of individual decision-making in healthcare. The

empowerment derived from “Dr. Google” reflects a shift toward a more participatory model of healthcare, where patients actively engage in their own well-being. However, this philosophical discourse also highlights the responsibility of healthcare providers to guide patients through the complexities of information, ensuring that autonomy is exercised within a framework of accurate, evidence-based knowledge. The implications of self-diagnosis within the context of patient autonomy extend beyond individual decision-making. The challenge lies in striking a balance between patient empowerment and the potential risks associated with inaccurate self-diagnoses. The philosophical discourse on patient autonomy prompts a critical examination of how healthcare systems can support patients in utilizing information responsibly and making informed decisions that contribute to their overall well-being. Recognizing the influence of “Dr. Google” on patient autonomy underscores the evolving role of healthcare providers. Rather than viewing patient-initiated information seeking as a challenge, healthcare professionals can embrace it as an opportunity for collaborative decision-making. This involves fostering open communication, addressing patient concerns, and providing guidance to ensure that patient autonomy is exercised in a manner that enhances, rather than compromises, overall health outcomes.

Generative models, with their capacity to analyze vast datasets and provide accurate medical diagnoses, contribute to the information landscape available to patients (and also to the medical experts!; Horak et al., 2023). The practice of self-diagnosis by lay people using internet information and generative AI models introduces several challenges and potential problems that need

careful consideration. Here are some key challenges associated with self-diagnosis in this context:

- a) **Misinterpretation of Information:** Lay individuals may lack the medical expertise to accurately interpret complex information provided by generative AI models. The output of these models might be misinterpreted, leading to incorrect conclusions about one's health condition. Misunderstanding medical terminology or misinterpreting probabilities could contribute to unnecessary anxiety or incorrect self-assessment.
- b) **Overemphasis on Rare Conditions:** Internet information and generative AI models often provide a wide range of potential diagnoses based on symptoms. Lay individuals may be prone to overemphasizing rare or severe conditions without considering more common and benign explanations. This can lead to unnecessary stress, anxiety, and potentially inappropriate actions or treatment decisions.
- c) **Confirmation Bias:** Individuals seeking health information online may unintentionally exhibit confirmation bias, where they selectively focus on information that confirms their preconceived beliefs or fears. This can reinforce inaccurate self-diagnoses and hinder the consideration of alternative, more accurate information.
- d) **Incomplete or Inaccurate Information:** The quality of health information available on the internet varies widely. Lay people may encounter incomplete or inaccurate information that has not been vetted by medical

professionals. Generative AI models, while powerful, may also be limited by the quality and representativeness of the data on which they were trained, leading to potential biases or inaccuracies in their predictions.

- e) **Delay in Seeking Professional Help:** Relying solely on self-diagnosis through internet information and generative AI models may lead to delays in seeking professional medical advice. Conditions that require timely intervention or treatment may be overlooked or underestimated, resulting in adverse health outcomes.
- f) **Psychological Impacts:** Engaging in self-diagnosis can have psychological impacts, including heightened anxiety, stress, or unnecessary worry. The fear of serious illnesses, even when unfounded, may adversely affect mental well-being and quality of life.
- g) **Inadequate Understanding of Diagnostic Uncertainty:** Lay individuals may not fully grasp the inherent uncertainties associated with medical diagnoses. Generative AI models provide probabilities rather than definitive answers, and users may struggle to comprehend the nuanced nature of these probabilities, potentially leading to unwarranted confidence or excessive concern.
- h) **Privacy and Security Concerns:** The use of online platforms for health information raises privacy and security concerns. Individuals may unknowingly expose sensitive health data, and the trustworthiness of platforms in protecting this information becomes crucial.

- i) **Medicalization of Normal Variations:** Normal variations in health can be misconstrued as pathological conditions through self-diagnosis, leading to unnecessary medicalization of natural bodily changes. This may result in unnecessary medical visits, tests, and interventions.

In addressing these challenges, it is essential to promote health literacy, guide individuals in critically evaluating online information, and emphasize the complementary role of professional medical advice in the diagnostic process. Healthcare professionals play a vital role in educating the public about the limitations and potential risks associated with relying solely on self-diagnosis through internet information and generative AI models.

6. The McDonaldization of Healthcare: Generative AI in the Hands of Private Companies¹

The concept of the “McDonaldization of Society,” introduced by sociologist George Ritzer, explores the standardization, efficiency, and predictability that characterize modern organizational structures, drawing parallels with the fast-food industry (Ritzer, 1993). Extending this metaphor to the realm of healthcare, particularly with the omnipresence of generative AI in the hands of private companies, provides a lens through which we can examine the potential implications for users. This essay explores

¹ The contents of this section were firstly explained in an invited talk at the University of Girona (Càtedra Ferrater Mora) on ‘Subrogate Cognition,’ in October 2023.

how the user experience in healthcare, facilitated by user-friendly and powerful chatbot systems, may be subject to a form of “McDonaldization.” Generative AI in healthcare, often proprietary to private companies, tends to streamline and standardize diagnostic processes. Chatbot systems, designed for efficiency, guide users through structured interactions for symptom assessment and preliminary diagnosis. While this can enhance accessibility and speed, it may also contribute to a standardized approach that oversimplifies the complexity of individual health cases. User-friendly chatbot systems, akin to the predictability of fast-food ordering processes, offer a standardized and guided experience. Users are prompted through predefined pathways, ensuring a predictable interaction. While this facilitates ease of use, it may inadvertently overlook the uniqueness of individual health conditions and symptoms. Also, the proliferation of generative AI in healthcare often results in a homogenization of health information. Commonly used chatbot systems may draw from similar databases and algorithms, providing users with standardized information (Nova, 2023). This raises concerns about potential biases and limited diversity in the information presented, potentially overlooking nuanced healthcare needs. Private companies that develop and deploy generative AI in healthcare systems wield significant decision-making power. Users, in relying on these systems, cede a degree of autonomy as the algorithms and pathways are determined by corporate entities. This centralization raises questions about the influence of profit motives and corporate interests on healthcare decision-making. Much like the emphasis on speed and accessibility in fast-food services, generative AI in healthcare prioritizes quick access to information and diagnoses. While this

meets the demand for immediate information, it may overshadow the importance of thorough and nuanced medical assessments that healthcare professionals provide.

The McDonaldization of Healthcare through generative AI potentially limits the human connection in the user experience. While chatbot systems offer rapid responses, they lack the empathy and contextual understanding that human healthcare providers bring to patient interactions. This depersonalization may impact the quality of the user experience. The metaphor of the “McDonaldization of Healthcare” through the lens of generative AI, predominantly controlled by private companies, highlights both advantages and potential drawbacks. While user-friendly chatbot systems offer unprecedented accessibility and efficiency, there are concerns about the standardization, homogenization, and centralization of healthcare processes. Striking a balance between the benefits of technology and the nuances of individual health needs is crucial to prevent the unintended consequences of a McDonaldized healthcare experience. As we navigate this evolving landscape, it is imperative to critically assess the impact of generative AI on user perspectives and advocate for a healthcare system that prioritizes both efficiency and the holistic well-being of individuals.

7. Conclusions

The integration of generative artificial intelligence, exemplified by ChatGPT, signifies a transformative paradigm shift in healthcare, offering promising opportunities for enhanced diagnostic capabilities and accessibility. Rapid analysis of extensive datasets

and the potential for lay individuals to engage in self-diagnosis present avenues for empowerment. However, formidable challenges include the imperative for robust validation, interpretability of AI-generated diagnoses, and ethical considerations surrounding data privacy and biases. Specific studies evaluating ChatGPT's diagnostic accuracy reveal variations across conditions, emphasizing the need for refinement in symptom identification. The critical importance of effective prompting techniques is underscored, with proposed tools and methodologies to improve diagnostic outcomes. Research on ChatGPT-4 suggests its potential as a standard source for medical information, despite existing limitations. The paper concludes that while generative AI holds promise for revolutionizing medical diagnostics, a balanced and collaborative approach involving healthcare professionals, technologists, and ethicists is essential to ensure accurate, ethical, and patient-centered healthcare. Controversial aspects, including issues of authorship and challenges in epidemiology, are explored. The philosophical implications of "Dr. Google" highlight the empowerment of patients through information access, tempered by the need for health literacy. Challenges in self-diagnosis using AI models, such as misinterpretation and privacy concerns, are acknowledged. The metaphor of the "McDonaldization of Healthcare" draws attention to the standardization and efficiency introduced by generative AI, posing potential drawbacks of depersonalization and centralized decision-making. We also suggest the need for caution in implementing AI for critical clinical applications and emphasize the importance of balancing technology's benefits with potential risks. While methodological creativity serves as a catalyst for advancement in scientific disciplines, its implementation

concurrently becomes a source of confusion (Vallverdú, 2017). The scientific controversies spanning the 20th and early 21st centuries surrounding epidemiology are intricately intertwined with parallel debates on the employed statistical techniques. Hence, it becomes imperative to acknowledge the epistemological challenges that accompany the implementation of methodological creativity's outcomes. Considering the impending integration of generative AI into our epistemic and operational endeavors, fostering an open yet critical mindset towards their possibilities and challenges is paramount. Embracing the revolutionary potential of generative AI demands a vigilant approach, addressing concerns such as interpretability, biases, and ethical considerations. Future exploration involves addressing challenges through advancements in prompt formulation, refining diagnostic precision, and fostering collaboration among diverse stakeholders, paving the way for a more nuanced and ethically grounded integration of generative AI in healthcare.

Viewing the revolution brought about by generative artificial intelligence (AI) in healthcare, classic epidemiologists such as Fischer, Bradford Hill, and Snow would likely appreciate the opportunities for efficient data analysis and pattern identification in identifying population-level health trends. The potential of generative AI to rapidly analyze extensive datasets aligns with the foundational principles they laid for understanding and controlling the spread of diseases. However, concerns may arise regarding challenges in causal inference, particularly in meeting criteria for causation, given the complexity of large language models. Ethical considerations, in line with their historical emphasis on ethical research, would likely lead these epidemiologists to scrutinize issues such as biases in training data,

data privacy, security, and the unintended consequences of AI-generated diagnoses. While they might see potential benefits in patient empowerment and a patient-centered approach, they would likely caution against overreliance on AI, emphasizing the complementary role of healthcare professionals. Furthermore, classic epidemiologists might appreciate the broader public health implications of generative AI, particularly in disease prediction and data analysis, but would advocate for rigorous validation and continuous exploration to address challenges and ensure the responsible integration of AI in epidemiology.

References

- Boscardin, C.K., Gin, B., Golde, P.B., and Hauer, K.E. (2023). "ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity". *Academic Medicine*. doi: 10.1097/ACM.0000000000005439. Online ahead of print.
- Caruccio, L., Cirillo, S., Polese, G., Solimando, G., Sundaramurthy, S., and Tortora, G. (2024). "Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot". *Expert Systems with Applications* 235: 121186.
- Garg, R.K., Urs, V.L., Agarwal, A.A., Chaudhary, S.K., Paliwal, V., and Kar, S.K. (2023). "Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review". *Health Promotion Perspectives* 13(3): 183.
- Golinelli, D., Sanmarchi, F., Nuzzolese, A., Toscano, F., Bucci, A., and Nante, N. (2023). "A Guide to AI in epidemiology:

- ChatGPT and the STROBE checklist for observational studies". *European Journal of Public Health* 33(Supplement 2): ckad160.1213.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2020). "Generative adversarial networks". *Communications of the ACM* 63(11): 139-144.
- Cox Jr, L.A. (2023). "Pushing Back on AI: A Dialogue with ChatGPT on Causal Inference in Epidemiology". In *AI-ML for Decision and Risk Analysis: Challenges and Opportunities for Normative Decision Theory* (pp. 407-423). Cham: Springer.
- Eysenbach, G. (2023). "The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers". *JMIR Medical Education* 9(1): e46885.
- Heaven, W.D. (2021). "Hundreds of AI tools have been built to catch covid. None of them helped". *MIT Technology Review*. <<https://www.technologyreview.com/2021/07/30/1030329/machine-learning-ai-failed-covid-hospital-diagnosis-pandemic>>
- Hirosawa, T., Harada, Y., Yokose, M., Sakamoto, T., Kawamura, R., and Shimizu, T. (2023). "Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: A pilot study". *International Journal of Environmental Research and Public Health* 20(4): 3378.
- Horak, J., Novak, A., and Voumik, L.C. (2023). "Healthcare generative artificial intelligence tools in medical diagnosis, treatment, and prognosis". *Contemporary Readings in Law and Social Justice* 15(1): 81-98.

- Ide, K., Hawke, P., and Nakayama, T. (2023). "Can ChatGPT be considered an author of a medical article?". *Journal of Epidemiology* 33(7): 381-382.
- Kingma, D.P., and Welling, M. (2013). "Auto-encoding variational bayes". *arXiv* preprint: 1312.6114. <<https://arxiv.org/abs/1312.6114>>
- Kulkarni, P.A., and Singh, H. (2023). "Artificial intelligence in clinical diagnosis: opportunities, challenges, and hype". *JAMA* 330(4): 317-318.
- Kuroiwa, T., Sarcon, A., Ibara, T., Yamada, E., Yamamoto, A., Tsukamoto, K., and Fujita, K. (2023). "The potential of ChatGPT as a self-diagnostic tool in common orthopedic diseases: exploratory study". *Journal of Medical Internet Research* 25: e47621.
- Lee, K., Hoti, K., Hughes, J.D., and Emmerton, L. (2014). "Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions". *Journal of Medical Internet Research* 16(12): e262.
- Lee, K., Hoti, K., Hughes, J.D., and Emmerton, L.M. (2015). "Consumer use of "Dr Google": a survey on health information-seeking behaviors and navigational needs". *Journal of Medical Internet Research* 17(12): e288.
- Mehnen, L., Gruarin, S., Vasileva, M., and Knapp, B. (2023). "ChatGPT as a medical doctor? A diagnostic accuracy study on common and rare diseases". *medRxiv* Preprint. <<https://www.medrxiv.org/content/10.1101/2023.04.20.23288859v2.full>>
- Nova, K. (2023). "Generative AI in healthcare: advancements in electronic health records, facilitating medical languages, and

- personalized patient care". *Journal of Advanced Analytics in Healthcare Management* 7(1): 115-131.
- Ritzer, G. (2011). *The McDonalidization of Society: An Investigation into the Changing Character of Contemporary Social Life*. Newbury Park: Pine Forge Press.
- Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A.I., Etmann, Ch., McCague, C., Beer, L., Weir-McCall, J.R., Teng, Z., Gkrania-Klotsas, E., AIX-COVNET, Rudd, J.H.F., Sala, E., and Schönlieb, C. B. (2021). "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans". *Nature Machine Intelligence* 3(3): 199-217.
- Sallam, M. (2023). "The utility of ChatGPT as an example of large language models in healthcare education, research and practice: systematic review on the future perspectives and potential limitations". *medRxiv* Preprint. <<https://www.medrxiv.org/content/10.1101/2023.02.19.23286155v1>>
- Sanmarchi, F., Bucci, A., Nuzzolese, A.G., Carullo, G., Toscano, F., Nante, N., and Golinelli, D. (2023). "A step-by-step researcher's guide to the use of an AI-based transformer in epidemiology: an exploratory analysis of ChatGPT using the STROBE checklist for observational studies". *Journal of Public Health*. doi: 10.1007/s10389-023-01936-y
- Vallverdú, J. (2016). "Causalidad y epidemiología". En Estany, A., y Puyol, A. (eds.), *Filosofía de la epidemiología social* (pp. 99-118). Madrid: Plaza y Valdés-CSIC.

- Vallverdú, J. (2017). "Creatividad epistemológica como factor de cambio y/o confusión: epidemiología y estadística". *Revista Iberoamericana de Argumentación* 15: 92-98.
- Vallverdú, J. (2022). "Ethics & causality in judicial algorithms". En Castro Toledo, F.J. (coord.), *La transformación algorítmica del sistema de justicia penal* (pp. 23-33). Navarra: Thompson Reuters-Aranzadi.
- Walker, H.L., Ghani, S., Kuemmerli, C., Nebiker, C.A., Müller, B.P., Raptis, D.A., and Staubli, S.M. (2023). "Reliability of medical information provided by ChatGPT: assessment against clinical guidelines and patient information quality instrument". *Journal of Medical Internet Research* 25: e47479.
- Xue, V.W., Lei, P., and Cho, W.C. (2023). "The potential impact of ChatGPT in clinical and translational medicine". *Clinical and Translational Medicine* 13(3): e1216.