# Generative Artificial Intelligence in Medicine: A Mixed Methods Survey of UK General Practitioners

Charlotte Blease, Anna Kharko, Cosima Locher, John Torous, Sophie Anna Rosch, Maria Hägglund, Jens Gaab, Brian McMillan, David Sundemo, Kenneth D. Mandl

Submitted to: Journal of Medical Internet Research
on: August 13, 2024

# *Table of Contents*

# Generative Artificial Intelligence in Medicine: A Mixed Methods Survey of UK General Practitioners

Charlotte Blease[1] PhD; Anna Kharko[1] PhD; Cosima Locher[2] PhD; John Torous[3] MD, MBI; Sophie Anna Rosch[4]; Maria Hägglund[1] PhD; Jens Gaab[2] PhD; Brian McMillan[5] PhD; David Sundemo[6] MD, PhD; Kenneth D. Mandl[3] MD

[1]Uppsala University Uppsala SE
[2]University of Basel Basel CH
[3]Harvard Medical School Boston US
[4]University of Zurich Zurich CH
[5]University of Manchester Manchester GB
[6]University of Gothenburg Gothenburg SE

**Corresponding Author:**
Charlotte Blease PhD
Uppsala University
Department of Women's and Children's Health
Uppsala
SE



## *Abstract*

**Background:** Since November 2022, with the debut of OpenAI's ChatGPT, there has been growing interest in the use of generative artificial intelligence (AI), including in healthcare. However, there is only limited research into doctors' adoption of these tools and their opinions about their application in clinical practice.

**Objective:** This study aimed to explore the opinions of general practitioners (GPs) in the United Kingdom (UK) about the use of generative AI tools (ChatGPT/Bard/Bing AI) in primary care.

**Methods:** Between February 2nd-24th 2024, using a convenience sample, we administered a web-based mixed methods survey of 1000 GPs in the UK to explore their experiences and opinions about the impact of generative AI on clinical practice. Participants were recruited from registered GPs currently working in the UK using the clinician marketing service Doctors.net.uk. Quantitative data were analyzed using descriptive statistics and nonparametric tests. We used thematic content analysis to investigate free-text responses and conducted a qualitative descriptive analysis of written responses ("comments") to 2 open-ended questions embedded in the web-based questionnaire.

**Results:** A total of 1006 GPs responded, with 53% being male and 54% aged 46 years and older. Most GPs (80%) expressed a need for more support and training in understanding these tools. GPs at least somewhat agreed AI would improve documentation (59%), patient information gathering (56%), treatment plans (41%), diagnostic accuracy (40%), and prognostic accuracy (38%). Additionally, 62% believed patients might rely more on AI, 55% felt it could increase inequities, and 54% saw potential for patient harm, but 47% believed it would enhance healthcare efficiency. GPs who used these tools were significantly more optimistic about the scope for generative AI in improving clinical tasks compared with those who did not report using them. Elaborating on the quantitative component of the survey, 31% (307/1006) left comments that were classified into 4 major themes in relation to generative AI in medicine: (1) lack of familiarity and understanding, (2) a role in clinical practice, (3) concerns, and (4) thoughts on future of healthcare.

**Conclusions:** This study highlights UK GPs' perspectives on generative AI in clinical practice, emphasizing the need for more training. Many GPs reported a lack of knowledge and experience with this technology and a significant proportion used non-medical grade technology for clinical tasks, with the risks that this entails. Medical organizations must urgently invest in educating and guiding physicians on AI use and limitations.

(JMIR Preprints 13/08/2024:65318)
DOI: https://doi.org/10.2196/preprints.65318

### Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

# Original Manuscript

# Generative Artificial Intelligence in Medicine:

# A Mixed Methods Survey of UK General Practitioners


## Abstract

**Background: Since November 2022, with the debut of OpenAI's ChatGPT, there has been growing interest in the use of generative artificial intelligence (AI), including in healthcare. However, there is only limited research into doctors' adoption of these tools and their opinions about their application in clinical practice.**

**Objective: This study aimed to explore the opinions of general practitioners (GPs) in the United Kingdom (UK) about the use of generative AI tools (ChatGPT/Bard/Bing AI) in primary care.**

**Methods: Between February 2nd-24th 2024, using a convenience sample, we administered a web-based mixed methods survey of 1000 GPs in the UK to explore their experiences and opinions about the impact of generative AI on clinical practice. Participants were recruited from registered GPs currently working in the UK using the clinician marketing service Doctors.net.uk. Quantitative data were analyzed using descriptive statistics and nonparametric tests. We used thematic content analysis to investigate free-text responses and conducted a qualitative descriptive analysis of written responses ("comments") to 2 open-ended questions embedded in the web-based questionnaire.**

**Results: A total of 1006 GPs responded, with 53% being male and 54% aged 46 years and older. Most GPs (80%) expressed a need for more support and training in understanding these tools. GPs at least somewhat agreed AI would improve documentation (59%), patient information gathering (56%), treatment plans (41%), diagnostic accuracy (40%), and prognostic accuracy (38%). Additionally, 62% believed patients might rely more on AI, 55% felt it could increase inequities, and 54% saw potential for patient harm, but 47% believed it would enhance healthcare efficiency. GPs who used these tools were significantly more optimistic about the scope for generative AI in improving clinical tasks compared with those who did not report using them. Elaborating on the quantitative component of the survey, 31% (307/1006) left comments that were classified into 4 major themes in relation to generative AI in medicine: (1) lack of familiarity and understanding, (2) a role in clinical practice, (3) concerns, and (4) thoughts on future of healthcare.**

***Conclusions:*** This study highlights UK GPs' perspectives on generative AI in clinical practice, emphasizing the need for more training. Many GPs reported a lack of knowledge and experience with this technology and a significant proportion used non-medical grade technology for clinical tasks, with the risks that this entails. Medical organizations must urgently invest in educating and guiding physicians on AI use and limitations.

# Introduction

## Background

Following the launch of ChatGPT in November 2022, there has been increased interest in the role of large language model (LLM)-powered chatbots in medicine. These models are particularly adroit at summarizing information and creating content. They can also engage users in ways that appear to mimic dialogue with their capacity to ‘remember’ previous prompts, and ability to rapidly respond to queries.

A growing body of research shows that documentation and administrative tasks are a leading source of burnout among all clinicians including primary care physicians, many of whom anticipate artificial intelligence (AI) will soon provide assistance [1,2]. LLM-powered chatbots may be especially well-equipped to help with these tasks [3,4]. As noted, these tools have particular strengths in summarizing outputs, and they can also do so in a requested style, literacy level or socio-emotional tone, suggesting scope for assistance with clinical documentation, including notes that patients may read [5,6]. Already, some versions of AI such as Nuance's Dragon Ambient eXperience (DAX) can “listen” into the appointment and translate the dialogue between doctor and patient into clinical notes [7]. So far it is unclear whether these tool improve practice efficiencies, but preliminary findings are promising [8,9]. A growing body of research indicates that LLM-tools could have considerable scope in assisting with patient history-taking [10], and in helping physicians with differential diagnosis especially in more complicated clinical presentations [10–12], and in rare disease diagnoses [13,14].

However, LLM tools have limitations and invite novel concerns. They are unable to discriminate

the quality of the information upon which they are trained, and outputs can be plain wrong, and more worryingly, subtly inaccurate [15]. Such incorrect information, conveyed in an authoritative tone, can potentially lead to patient harm [8,16]. These models may also risk "algorithmic discrimination" via the creation of unfair biases in their recommendations [17,18]. It is unclear whether these risks are worse than the ones that already exist with human mediated care [19]. Findings indicate that these models could compound inequity problems [20]; however, other research suggests that there is promising scope for this AI to be employed, constructively, as a tool to reduce clinical inequities [21,22].

Patient privacy can be put at risk with the adoption of some LLM-powered chatbots in clinical contexts [23]. Owing to their ease of use, conversational fluency and perhaps the tendency of users to anthropomorphize these tools, many patients and clinicians may be tempted to input sensitive, potentially personally-identifiable clinical information [24]. While medical-grade chatbots that comply with healthcare privacy standards exist, with many more under development [25], readily available commercial chatbots such as GPT-4 are not designed for medical use and some medical organizations have issued position statements cautioning that physicians should not be using these tools to undertake clinical work [26,27].

The American Medical Association released a statement in April 2024 outlining principles for the development and deployment of AI in healthcare but did not offer concrete guidance [28]. In the UK, there have been calls to develop professional guidance to support clinicians in using AI [29]. While a multi-agency partnership has been established with a view to regulate the use of AI in clinical settings [30], guidance from NHS England currently advises that the use of AI in clinical settings should be a matter of professional judgement [31]. In the UK and EU, countries must

abide by the Medical Device Regulation (MDR) stating that all medical technical products used in healthcare must be CE-marked and MDR-approved [32]. As generative AI-models are general purpose and not validated as medical grade, it is of particular interest to know if these models are being employed for medical purposes.

AI is rapidly evolving. Knowing how clinicians are currently using and thinking about these tools is critical to better understanding professional uptake and training needs. Greater awareness about how these tools are being used or perceived, for example, could also help guide the development of LLMs to ensure safe adoption. Against the impressive advances and research into generative AI in healthcare, there is limited research into clinicians' use of LLM-powered chatbots.

In an earlier publication emerging from the present study we found that a fifth of GPs (20%; 205/1006) had used generative AI to assist with some aspect of clinical practice [42]. Updating our exploration of clinician uptake of LLMs, we performed a brief scoping review of the literature using the terms "generative artificial intelligence", "large language models," "surveys," and "doctors" in the search engines of PubMed and Google Scholar and explored the grey literature. Findings were limited although preliminary evidence suggests some physicians may be adopting LLM-powered chatbots for a variety of tasks including assisting with documentation [33,34]. Another study of 420 US medical trainees (response rate 50%), found that 40% had used ChatGPT [35]. In July 2023, five hospitals in the South Metropolitan Health Service in Perth, Australia, were directed to cease using ChatGPT after it was found that some staff members had used the chatbot to write clinical notes [36]. In June 2023, New York Times reported that doctors are asking ChatGPT how to improve their communication with patients [37].

There is a dearth of research into the perspectives of practicing clinicians on these tools. Given that primary care doctors offer the fundamental gateway to healthcare for many patients, and because these physicians work under considerable frontline pressure, the current study aimed to gauge the views of general practitioners about these tools. We sought to explore the opinions of general practitioners' (GPs) use of this new generation of chatbots in the UK.

# Methods

## Main Survey

We conducted an anonymous nationwide, mixed methods, web-based survey of GPs in the UK (n=1006; see Appendix 1). The study employed a convenience sample to solicit the opinions of GPs using the membership of the clinician marketing service Doctors.net.uk [38]. This is the largest professional network and online information service for doctors in the UK; it is free to join and requires doctors to give their General Medical Council (GMC) registration details [38]. Currently, Doctors.net.uk has 254,741 members out of a total of approximately 379,208 registered doctors in the UK (67%) [39]. The study was selected to be part of an "omnibus" monthly survey administered by Doctors.net.uk which focus on topical medical issues and have predetermined sample sizes of 1000 participants as this is in line with other sample sizes of healthcare providers [40]. The study team has obtained samples from Doctors.net.uk in previous studies using similar methods [1,2,41].

At Doctors.net.uk, a percentage of GPs active within the community consent to being sent survey invitations via email, which differs according to those who are active on the website at any given period. Depending on how GPs consented to receive survey invitations, between 2 February to

22nd February 2024, our study was advertised via email and displayed on the Doctors.net.uk home pages or only displayed on doctors' homepages, among a sample of GPs. During the survey administration period, approximately 21,000 GPs were active in the community.

We asked Doctors.net.uk to invite a random sample of GPs from across the UK. This sampling was stratified by regional location using demographic information about currently registered GPs working in the UK provided by the General Medical Council (GMC) in the GMC Data Explorer [39]. Using this stratification, Doctors.net.uk embedded invitations within the homepages of 13,967 GPs who were active on the website in the last 90 days; of these, 2,775 who consented to research communications were also sent an email to participate.

We used the CHERRIES checklist for survey administration and reporting (see Appendix 2). Ethical approval for this survey was obtained from the Faculty of Psychology, University of Basel, Switzerland (Protocol # EKFP-034-23-1). The invited participants were informed that no personally identifying data would be collected. The survey platform, www.doctors.net.uk, operates on a secure platform that ensures that personal data is numerically stored and fully anonymous – that is, not linked to the participants' responses. All personal data, such as email addresses, were removed from respondents' IDs before the transfer of the data to the research team. Www.doctors.net.uk meets the requirements of UK and EU laws on the General Data Protection Regulation. All respondents provided informed consent before participating. On completion, respondents were recompensed for their time with £7.50 (US $8.80) worth of "1000 eSR" points via Doctors.net.uk, which are exchangeable for web-based shopping vouchers.

## Survey

The study team developed a survey instrument to examine GPs' use and opinions about generative AI in clinical practice (see Appendix 1). The survey, which encompassed 4 items, was timed to take around 5 minutes to complete and pre-tested and piloted with 6 UK GPs to assess face-validity. The survey was divided into three main sections and specifically asked about GPs' experiences and opinions with respect to Open AI's ChatGPT, Google's Bard or Microsoft's Bing AI. We emphasize that Chat GPT/Bard/Bing AI are not specifically trained on medical data. We recognize that there is currently a plethora of LLMs, including but not limited to Anthropics Claude-series, Metas Llama-models, Snapchats MyAI and Mistrals AI-models. However, given the commercial availability of Chat GPT/Bard/BingAI at the time of data collection, we focused on them in this study.

Part A examined GPs' experiences with using generative AI to assist with any aspect of clinical practice and, as noted, responses to this section are published separately as a brief report [42]. Part B requested participants to offer their opinions about these generative AI tools in the delivery of care. This section used a mixture of closed ended Likert scale survey items and included two optional open-ended questions. Section C requested demographic information, including participant gender, age, their GP role, the regional location of their practice, and the number of patients in their general practice. Further screening questions were embedded within the survey to determine whether respondents were currently registered and practicing as GPs in the UK. All participants were requested to answer all the closed-ended questions and participants were advised they could leave the survey at any time.

## Quantitative Analysis

After survey collection, quantitative survey responses were entered into Excel (Microsoft Corp), and descriptive statistical analysis was carried out using JASP (v 0.9.2; University of Amsterdam), RStudio (v 2024.04.2+764), and SPSS (v 27; IBM Corp). We summarized respondents' characteristics through descriptive statistics (count, percentage). We compared our sample of GPs to those on the GMC registry based on age, gender and practice location through a chi-square of independence test.

To explore the impact of professional experience of AI on opinions about generative AI we divided respondents into those who have used AI to assist in clinical practice and those who have not (Item Q1, see Appendix 1). We analyzed the impact of AI use on AI potential for improvement (Item Q2) as well as impact on patients and clinicians (Item Q3). Due to these items comprising of several statements which are rated on a 7-point agreement Likert scale, we used chi-square of independence test to compare the two GP groups.

Lastly, we analyzed what GP characteristics affected use of generative AI. We conducted a bidirectional stepwise regression analysis to investigate the predictors of AI usage in clinical practice. The response variable was whether respondents *used AI tools in their clinical practice* (0 = Not used AI, 1 = Used AI). Predictor variables included *gender* (Male; Female; Other; Prefer not to say), *age* (35 or under; 36 to 45; 46 to 55; 56 or over), *practice size* (up to 5,000 patients; 5,001-7,500 patients; 7,501-10,000 patients; 10,001-12,500 patients; 12,501 patients or more), *role* (GP Partner/Principal; Salaried GP; Locum GP; GP Registrar; Secondary Care doctor; Other), and *practice location* (Major conurbation; Large town/city; Medium town/city; Small town/city; Village/hamlet; Other). For this analysis, participants' data were excluded if they did not indicate their gender (Prefer not to say; $n$ = 8). Predictors were added or removed

stepwise based on Akaike's Information Criterion (AIC). The full model included interactions between gender and the other predictors, while the null model included no predictors. The most parsimonious model with the lowest AIC value was selected.

## Qualitative Analysis

The two optional open-ended questions formed the basis of the qualitative analysis. These questions allowed participants to respond in detail to the topic of the questionnaire: *"Please add any additional comments you might have about the use of Chat GPT/Bard/Bing AI in primary care"* (Item Q5, see Appendix 1), and *"Please add any additional comments you might have about the use of the Chat GPT/Bard/Bing AI in medicine"* (Item D3).

We carried out inductive thematic coding of the fully anonymized data [43]. Inductive thematic coding is a robust method for identifying themes that emerge directly from the data, ensuring that the analysis is grounded in the participants' actual experienced and opinions. This approach promotes a deep understanding of context-specific nuances in participants' comments. Responses were analyzed by two members of the research team (CB and JT). CB is a philosopher of medicine, ethicist, and health informatician from the United Kingdom, and JT is a psychiatrist and health informatician from the United States. The transcripts were initially read several times to achieve familiarization with participants' responses. Next, an inductive coding process was used, in which brief descriptive labels ("codes") were applied to each comment. Multiple codes were applied to the comments with multiple meanings. Comments and codes were reviewed and compared to investigate similarities and differences. As a result of this iterative process, first-

order codes were grouped into second-order themes. CB and JT met to discuss coding decisions, and subsequently, minor revisions were made.

# Results

## Quantitative Analysis

## Study population

Of the 2,775 who received the email and home page invitations, 2,683 opened the email invite, and 604 clicked on the survey link with 532 completing the survey (532/2775, 19%); the remainder (n=474) accessed and completed the survey via their homepage (474/11,192, 4%). Of the total number of 1006 GPs who responded, 53% (531/1006) were male, and 54% (544/1006) were aged 46 years or older (see Table 1).

[INSERT TABLE 1 HERE]

Requested to describe their role, almost half reported ‘GP Partner/Principal’ (45%; 455/1006), followed by ‘Salaried GP’ (34%; 341/1006), ‘Locum GP’ (18%; 181/1006), and lastly ‘GP registrar (3%; 29/1006). As seen in Table 2, most were based in a ‘small town/city’ (30%; 305/1006) and had ‘12,501 or more’ patients listed in the practice (34%; 344/1006).

[INSERT TABLE 2 HERE]

Our participant pool exhibited some distinctions from the GPs registered with the GMC in the UK as of March 2024. In the GMC registry, the majority of GPs were female (40,871/72,254, 56.57%), contrasting with our sample where male respondents prevailed (531/1,006, 52.78%).

Age distribution in our sample ranged from 35 and under (76/1,006, 7.5%) to 56 years or older (196/1,006, 19.48%), with the largest part aged between 36 and 45 years (386/1,006, 38.37%). Similarly, 34.59% (348/1,006) were aged 46 to 55. In comparison, the GMC registry showed a different age profile, with the majority aged between 41 and 60 years (38,775/72,254, 53.66%), while 32.13% were between 21 and 40 years old (23,218/72,254). A smaller proportion was aged 61 to 80 years (10,030/72,254, 12.88%), or 81 or above (231/72,254, 0.32%). Although proportional regional representation in our sample closely resembled that of the GMC registry, smaller variations were observed (see Table 3). Using descriptive statistical analysis, our sample exhibited slight underrepresentation in London (10,293/71,900, 14.32% vs 113/1006, 11.23%) and Scotland (6,920/71,900, 9.62% vs 84/1006, 8.35%), and slight overrepresentation in the North West (7,750/71,900, 10.78% vs 121/1006, 12.03%), North East and Yorkshire (8,683/71,900, 12.08% vs. 131/1,006, 13.02%), and the Midlands (10,328/71,900, 14.36% vs 162/1006, 16.1%).

[INSERT TABLE 3 HERE]

## GP characteristics associated with use of generative AI

Differences in participant characteristics were observed between those who used AI and those who did not (see Table 4).

[INSERT TABLE 4 HERE]

Analysing the characteristics correlated with GPs reported use of generative AI in clinical practice, we carried out a regression (see Table 4 in Appendix 3). Based on AIC statistic, the model that best explained the variance in AI use consisted of Gender and Practice location, with Female and Medium town/city as

reference categories (see Equation 1), $F = 5.499$, $df = 6991$, Adj. $R^2 = 0.26$, $p < .001$.

$$[\text{Used AI} = \beta_0 + \beta_1 \text{ x Gender} + \beta_2 \text{ x Practice Location} + \epsilon]$$

$$\text{Used\_AI} = \beta_0 + \beta_1 \text{ x Location}_{\text{Major conurbation}} + \beta_2 \text{ x Location}_{\text{Large town/city}} + \beta_3 \text{ Location}_{\text{Small town/city}} + \beta_4 \text{ x Location}_{\text{Village/hamlet}} + \beta_5 \text{ x Location}_{\text{Other}} + \beta_6 \text{ x Gender}_{\text{Male}} + \epsilon$$

Male GPs were significantly more like than female GPs to have used AI, as well as those practising in bigger cities (see Table 4 in Appendix 3).

## GPs' opinions about the impact of Generative AI

GPs reported a variety of opinions about the impact of generative AI on clinical tasks with at least one in five respondents reporting "don't know". In total, however, GPs were more positive than negative about the impact of generative AI across all tasks except for empathy (see Figure 1).

[INSERT FIGURE 1 HERE]

Requested to offer opinions about the impact of these tools on care delivery, GPs reported more mixed opinions (see Figure 2).

[INSERT FIGURE 2 HERE]

Finally, asked for their opinions on possible litigation risks associated with the use of generative AI, across the whole sample, 2% (n=22) reported these tools would "decrease my risk of having legal action taken against me", 34% (n=340) believed that use would "neither decrease nor increase my risk", 20% (n=201) believed that use would "increase my risk of having legal action

taken against me", and 44% (n=443) said that they "didn't know".

## GPs' opinions based on their reported use of Generative AI

When comparing between respondents who reported using or not using AI for assistance in clinical practice, significant differences emerged for all survey items describing AI's potential for improvement with clinical tasks (see Table 1 in Appendix 3). GPs who reported using generative AI tended to anticipate greater potential to improve clinical tasks than GPs who did not report using it, with the latter also providing more 'Don't know' answer (see Figure 3).

[INSERT FIGURE 3 HERE]

Significant differences also emerged for all survey items describing generative AI's potential to impact patients and clinicians (see Table 2 in Appendix 3). GPs who reported using generative AI in clinical practice tended to anticipate both more positive and fewer harmful effects and provided fewer 'Don't know' answers than those who did not report using it (see Figure 4).

[INSERT FIGURE 4 HERE]

Lastly, of the GPs who reported experience with using generative AI in clinical practice, GPs offered a range of opinions about the impact of these tools on litigation risk (see Table 5).

[INSERT TABLE 5 HERE]

Chi-square of independence indicated that the differences in answer distribution were significant (see

Table 3 in Appendix 3).

## Results of the Qualitative Survey

### Overview

A total of 307 (31% out of the 1006 respondents) left comments in response to at least 1 question (6,468 words in total). Comments were brief (typically, sentence fragments, or 1-3 sentences). GP respondents who submitted comments to each question were not significantly different in terms of sex and age from those who did not submit comments; however, they did differ depending on whether they reported using generative AI in clinical practice, or the number of reported patients in their practice, with GPs who used AI leaving more comments.

Owing to the iterative thematic analysis process, four major categories of GPs' views were identified in relation to generative AI in medicine: (1) lack of familiarity and understanding, (2) its role in clinical practice, (3) concerns, and (4) the future (see Figure 5). These categories were further subdivided into 12 themes, which are described in subsequent sections with illustrative comments. Numbers in parentheses are identifiers that ascribe comments to individual participants; in addition, gender is marked as (Male or Female), GP role is marked as ('Partner/Principal', 'Salaried,' 'Locum', or 'Registrar), and whether GPs reported using any generative AI to assist in any aspect of clinical practice (Y or N).

[INSERT FIGURE 5 HERE]

### Lack of awareness

### No knowledge

**A dominant theme was lack of knowledge about generative AI tools. Multiple short comments stated, “not aware of these,” “no idea,” “no knowledge.” Similarly, many GPs responded that they didn’t know enough to comment; for example, “I don’t know anything about how it works so have been unable to offer an opinion” (#20, Female, Partner/ Principal, N). Some were openly candid about their lack of knowledge in relation to general practice: “I honestly have never used any AI software and therefore find it hard to have insight” (#228, Male, Locum, N), “I just don’t know how this will impact us” (#995, Male, Partner/Principal, N,). Some expressed anxieties but were open about their lack of awareness: for example, “I am fearful, but ignorant!” (#529, Female, Partner/Principal, N). Relatedly, some GPs also identified a lack of training, or need for more education on generative AI. For example: “Need more training and information about them” (#94, Female, Salaried, N), “No idea what it’s about and we haven’t been given any help/support/info regarding this” (#45, Female, Salaried, N), and “I would love to learn more about this” (#298, Female, Locum, N).**

## No experience

Another related dominant theme was a lack of experience with generative AI tools. Multiple short comments reflected that many respondents had “never used,” or had “not tried” these technologies with some comments combining lack of experience with lack of knowledge; for example: “not familiar with chat GPT or other AI, never used them” (#3, Female, Salaried, N). A few respondents signaled some awareness but lack of experience; for example, “I have never used them. If I was having to write lectures or essays I might?” (#986, Male, Partner/Principal, N), “I’m not too familiar with these but think for information gathering they could be useful…” (#165, Female, Salaried, N), and “I’d be interested in giving it a go.” (#265, Female, Salaried, N).

## A role in clinical practice

## Optimism

Another major theme was optimism about the role of current AI in medicine. Many comments were brief, but remarks straddled a broad range of opinions. Most common was the sentiment of guarded optimism suggesting “benefits but some concerns” (#479, Male, Locum, N). For example: “correctly used will improve healthcare” (#59, Male, Partner/Principal, Y) and “potential for good, but also big risk of harm” (#329, Male, Salaried, N). A few GPs were more strident in their enthusiasm; for example: “Exciting possibilities,” (#205, Female, Salaried, N), “ChatGPT would be very good in primary care” (#311, Male, Partner/Principal, Y), “Bring it on!” (#399, Male, Partner/Principal, Y)

## Documentation and administrative tasks

Some GPs expressed buoyancy about the potential for generative AI in documentation and administrative tasks; for example: “Could potentially be immensely useful for documentation, and for collating

information from notes," (#47, Male, Salaried, Y), and "I think it can be excellent in administrative contexts" (#290, Male, Salaried, N). Some comments described the usefulness of these tools in, "letter writing", for example, "improving text used for template letters e.g., letters in support of housing/benefits/ education" (#156, Female, Salaried, Y).

## Assisting with medical tasks

Some comments suggested generative AI could "improve triage"; for example, "They could be a great tool for initial patient screening and care navigation" (#328, Male, Salaried, Y). The concept of generative AI as a co-pilot for clinical tasks was also identified: comments described these tools as assisting with, or augmenting, medical tasks that doctors would ultimately oversee; for example: "I see ChatGPT or other AI as an aide for example spotting patterns of presentations and helping work out optimum. monitoring as well as helping with recall" (#1124, Female, Partner/Principal, Y), "list differentials generated by AI e.g., I may not have considered z diagnosis, but I did consider x and y" (#419, Male, Salaried, N), "May be helpful in aiding differential diagnosis, treatment plans, and self-care" (#814, Male, Partner/Principal, N).

## Concerns

## Deep skepticism

Another prominent theme was skepticism about the use of AI in medicine. These comments were brief but emphatic. For example, "I have deep concerns" (#172, Male, Partner/Principal, N), "I am very skeptical and nervous about it" (#163, Female, Salaried, N), "Absolutely would not use in this context" (#836, Female, Locum, N), "At this stage represents huge risks" (#713, Male, Partner/Principal, N), "It's crap" (#536, Male, Partner/Principal, Y).

## Errors and inaccuracies

Some GPs also offered justifications for this skepticism concerns by alluding to the risks of errors and inaccuracies with generative AI technologies; again, however, remarks were brief. For example: "All the

time they make stuff up" (#252, Female, Salaried, N), "It's not accurate," (#321, Female, Locum, N). Only a few GPs offered more developed reflections on these limitations "AI cannot understand context" (#376, Male, Partner/Principal, N), "Looks so authoritative that may be misleading" (#87, Male, Salaried, Y), "AI Language models hallucinate therefore cannot be relied on for evidence-based decision making" (#972, Male, Partner/Principal, Y).

## Legal concerns

Another concern described by GPs was accountability for mistakes and legal implications for generative AI tools; for example: "who will take the blame????" (#898, Male, Salaried, N), "unclear at this stage where liability would lie" (#835, Male, Registrar, N), "Who will take the litigation risk if a physician uses AI tools? If the litigation risk can be reduced or removed, then I think AI would be more welcome" (#419, Male, Salaried, N). The risks of defensive medicine were also raised: "Increased risk of using i.e. a patient will sue me for not doing a million tests that ChatGPT suggested" (#560, Male, Partner/Principal, Y), "I can see LLM output being uses in legal action against doctors which will encourage defensive medicine" (#627, Male, Registrar, N).

## The future

## Too early to call

A prominent theme was the idea that AI was "a "fledgling technology" (#1022, Male, Salaried, N) and at "in an infant stage" (#1066, Male, Partner/Principal, Y). Some GPs believed it was too premature to make a judgment call on the impact on medicine: for example, "It is too early to say how AI will change clinical practice" (#456, Male, Salaried, N), and "AI is still quite green. Time will tell what impact will have" (#746, Male, Locum, Y).

*Safety research and training*

Anticipating a potential future role for AI in medicine, some comments expressed the need for

more research into the safety of these technologies. For example: “needs more trials done before used in clinical practice” (#263, Female, Salaried, N), “Good in theory, needs more improvements to be fully reliable” (#542, Male, Salaried, N); “exciting developments but my understanding is they need a lot more improvement to reduce hallucinations before could be put into practice” (#394, Female, Partner/Principal, N); “More research is needed to assess safety before implementation.” (#62, Prefer not to say, Salaried, N)

To future-proof the safe adoption of these tools, some GPs also suggested that more training and education would be needed; for example: “It will be coming, but I hope we are trained in it by then,” (#839, Female, Salaried, N), and “More training needed if to be used regularly. How can we use it to our advantage safely” (#142, Female, Salaried, Y), “Needs training & guidance & time” (#485, Male, Locum, Y).

## Doctor-patient interactions

Many comments expressed the opinion that face-to-face interactions between doctors and patients would always be necessary. Some of these comments reflected the view that medical expertise required human interactions; for example, “I think the use of AI will remove from general practitioner their ‘gut feeling,’ ‘physical examination,’ ‘history of patient,’ and the benefit of the face-to-face assessment” (#698, Male, Partner/Principal, N), “However good these things get, there is still no substitute for that gut-feeing something isn’t right with a patient” (#722, Female, Salaried, Y).

Other comments expressed the view that patients will always desire human interactions: for example, “It devalues the significant benefit from someone telling their doctor face to face what they are worried about” (#396, Female, Salaried, N), “a lot of my job is about genuine interpersonal interaction and empathy and I cannot see how AI will achieve this” (#229, Female, Partner/Principal, N).

# Discussion

## Main findings

Few studies have explored the opinions of physicians about large language model chatbots on clinical practice. This mixed methods study of 1006 practicing physicians explored the experiences and opinions of UK GPs about Chat GPT, Bard, and Bing AI on their practice. Building on previously published findings from this survey which found 20% (n=205) of UK GPs in the sample reported using generative AI tools in clinical practice [42]; in the present study we further explored the experiences and opinions of British primary care doctors about these tools. We found that male GPs working in the UK were significantly more like than female GPs to have used AI, as well as those practising in bigger cities.

Notably, 80% of surveyed GPs at least somewhat agreed that they needed more support/training in understanding these tools. 62% of GPs at least somewhat agreed that more patients would rely on these tools instead of seeking medical attention. While more than a quarter responded that they “didn’t know” what the impact of generative AI would be on care delivery, more than half of surveyed GPs at least somewhat agreed that these tools would increase risks of inequities (55%), and patient harm (54%). In contrast, however, nearly half (47%) believed these tools would increase efficiencies in healthcare.

We found significant differences in respondents’ opinions depending on whether GPs reported using generative AI tools in any aspect of clinical practice. Among those who had used these tools compared with those who had not, GPs were more likely to at least somewhat agree they needed more support and understanding in these tools (87% vs 78%), that patients would rely on

these tools before seeking medical attention (70% vs 60%), and these tools would increase efficiencies in care (71% vs 41%). In contrast, those who reported using Chat GPT, Bard or Bing AI to assist with clinical practice, were also more likely to at least somewhat agree that these tools could increase inequities in care delivery (57% vs 54%), and patient harm (57% vs 51 %), with more who had used AI at least somewhat disagreeing with these statements (27% vs 15% and 27% vs 13%, respectively).

Regarding the potential for Chat GPT/Bard/Bing AI to improve tasks in care, GPs at least somewhat agreed that these tools would improve documentation (59%), patient information gathering (56%), the creation of personalized treatment plans (41%), diagnostic accuracy (40%), and prognostic accuracy (38%). Most GPs (65%) at least somewhat disagreed that these tools could improve the delivery of empathy. Again, opinions differed depending on whether GPs reported using generative AI: GPs who used these tools were more likely to be positive than those who had not, at least somewhat agreeing that these tools would improve documentation (78% vs 54%), patient information gathering (67% vs 53%), the creation of personalized treatment plans (61% vs 35%), diagnostic accuracy (59% vs 35%), prognostic accuracy (51% vs 34%), and the delivery of empathy (29% vs 9%).

Across the whole sample, around a third (34%) believed that use of these tools would “neither decrease nor increase” the risk of litigation being taken against them, with 20% believing that use would increase the risk of an action being taken, with 44% saying they “didn’t know”. Here again, use of these tools influenced attitudes: 35% of those GPs who reported using generative AI compared with 16% who had not believed these tools would “neither decrease nor increase their risk”; and 6% who used it believed it would decrease their legal risk compared with 1%

who had not used it.

Results from the qualitative section of the survey supported and elaborated on these experiences and opinions. A dominant theme was a lack of deep knowledge, awareness, or experience with generative AI tools. Another leading theme was the role of these tools in clinical practice with many GPs expressing optimism for the potential for benefits in primary care – especially when it comes to documentation and administrative tasks – even if this buoyancy was often tempered with some concerns. Indeed, concerns about including about risks of clinical errors and inaccuracies, as well as legal anxieties about accountability and blame, comprised another overarching theme. Reflecting on the future of generative AI in primary care, many GPs expressed uncertainties about the scope for impact, while the need for safety research and physician training were identified as growing needs. Finally, and again supporting the quantitative component of the study, some GPs also expressed the view that doctor-patient interactions would always be necessary to preserve empathy and humanistic aspects of care.

## Comparison with prior work

The results of this study mirror other recent exploratory studies which suggest clinicians may be adopting these tools in greater numbers than anticipated. For example, in June 2023, in a Medical Economics study, more than 1 in 10 surveyed healthcare professionals reported using LLM-model powered chatbots such as ChatGPT, and nearly 50% expressed an intent to use these technologies in the future for tasks such as data entry, medical scheduling or research.[33] Similarly, in October 2023 a small survey (n=138) conducted with psychiatrists affiliated to the American Psychiatric Association found that 44% of respondents had used ChatGPT 3.5 and 33% had used 4.0 *"to assist with answering clinical questions"* with 70% of psychiatrists

believing that "*documentation will be/is more efficient*" as a result of these tools [34]. In that study, 8 in 10 agreed "clinicians need more support/training in understanding these tools" with respondents expressing divergent opinions about the promise and harms associated with these chatbots in healthcare. An August 2023 survey by the American Medical Association found that US healthcare workers also shared mixed views on AI with the most encouragement around documentation and the most concern around privacy [40]. Our study echoes these findings.

We also observed some complementary predictions with medical informaticians and other experts working in AI/ML and related fields. Similar to the views of researchers and experts investigating the role of AI healthcare [44], many GPs optimistically advanced the view that these tools might function best as "co-pilots" assisting them in a variety of tasks. Like expert informaticians, GPs in our survey expressed the need for more AI training. This education concern echoes the forecasts of a Delphi poll of international health informatics experts who reported consensus that in 10 years (by 2029), advances in AI/ML would prompt workplace changes including incursions into the disintermediation of physician expertise, and increased AI/ML training requirements [45].

GPs in our study, especially those who had used generative AI, believed patients would increasingly use these tools. Notably, experts in the aforementioned international Delphi poll predicted that by 2029, AI and machine learning will significantly enhance diagnostic accuracy, particularly benefiting individuals with limited access to specialists, minority groups, and those with rare diseases; these panelists also predicted that AI/ML-tools would improve access to expert doctor knowledge [45]. Similarly, in the recent Medical Economics survey, 8 in 10 Americans surveyed believed tools such as ChatGPT had the potential to improve the quality of

healthcare, reduce costs and increase the accessibility of care [33].

## Strengths and limitations

The main strength of this survey is that it is the largest conducted anywhere exploring clinicians' opinions about the impact of generative AI on medicine. The use of a web-based survey may have facilitated honest feedback about UK GPs' views on generative AI, as reflected in the candour of our participants' responses.

The study has several limitations. The use of a convenience sample means that it is not possible to infer that our sample was representative of the opinions of GPs in the UK. Although we strove to stratify the sample according to geographic location, our respondents differed from those registered with the GMC. While our sample closely resembled the geographical representation of UK GPs, our sample was younger and included more male GPs. Our respondents were restricted to those GPs who used Doctors.net.uk, including during the administration of the survey and response biases may have influenced the findings. For example, it is possible that GPs with stronger (negative or positive) views about generative AI may have been more inclined to participate. Selection bias can also be a contributory factor explaining the majority of male GPs observed in the sample, in contrast to the majority of females in the underlying GMP population. In addition, comments in the qualitative component of the survey were often brief owing to the constraints of the web-based survey design. As noted in the Methods, we specifically restricted the survey to opinions about commercially available LLM-powered chatbots, but respondents may have had different opinions about medical grade models such as Google's PALMMed2 and this should be explored further too.

In the current study, we chose to evaluate three specific models; Chat GPT/Bard/Bing AI. We believe it is a reasonable choice, owing to the public renown and availability of these models. Naming the models was also deemed necessary, not to confuse other types of AI with generative AI. However, it is possible that more advanced AI-users, aware of the privacy concerns, instead use locally hosted open-source models such as iterations of Llama or Mistral. The answers from such users may have been missed in the current survey.

More robust, nationally representative surveys are required to obtain a clearer picture of doctors' views and experiences with generative AI. In addition, future research could usefully engage the expert opinions of health informaticists, medical AI researchers, and ethicists in relation to the use of medical-grade generative AI in clinical practice. Finally, more extensive, stand-alone survey research is required to understand the perspectives of patients with using generative AI, including their opinions about doctors using these tools.

## Conclusions

This descriptive analysis provides exploratory insights into the views of 1006 GPs in the UK regarding their experienced and opinions about generative AI in on patient care and in their practice. There was an overriding shared consensus among GPs that they needed more support/training in understanding generative AI tools. Notably, however, GPs who used these tools were significantly more optimistic about the scope for generative AI in improving clinical tasks compared with those who did not report using them. GPs who adopted these tools were more likely to agree that generative AI could improve documentation, patient information gathering, the creation of personalized treatment plans, diagnostic accuracy, prognostic accuracy, and empathic care. These GPs were also more likely to believe generative AI would improve

efficiencies in care.

Conspicuously, however, GPs held disparate opinions about the litigation risks of using these tools, including when it came to accountability if things go wrong. Many GPs expressed lack of knowledge and experience with these tools. Lack of awareness was most notable in the absence of discussion about the patient privacy concerns with the adoption of these tools.

We conclude that many physicians may already be using generative AI at scale and finding use for these tools in clinical practice. We caution that medical organizations must urgently invest more sustained efforts in educating and guiding physicians on the safe adoption and limitations of generative AI in patient care. We suggest that relevant federal and national medical organizations also make it clear to physicians what types of generative AI can be used, and how.

## Additional statements

### Acknowledgments: The authors would like to thank Nicola Miles and Gary Kingsnorth at Doctors.net.uk for assistance with administering the survey.

**Conflicts of interest:** The authors have no conflicts of interest to disclose.

**Contributors: Conceptualization: CB, KM; Data analysis: AK, CL, CB, JT, ASR; Investigation: CB, KM; Project Administration: CB, JG, CL, MH; Supervision: CB, CL, AK; Visualization: AK; Writing–original draft preparation: CB, AK, CL; Writing–review & editing: CB, KM, JT, AK, CL, ASR, BM, DS, JG, MH. All authors read and approved the final manuscript.**

**Funding statement: This study was supported by the project "Beyond Implementation of eHealth" (2020-01229) awarded by the Swedish Research Council for Health, Working Life and Welfare (FORTE), at Uppsala University, Sweden, and by the Department of Clinical Psychology and Psychotherapy, University of Basel, Switzerland.**

**Data availability statement:** Raw data are available as a supplementary file.

## References

1. Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM. Artificial Intelligence and the Future of Primary Care: Exploratory Qualitative Study of UK General Practitioners' Views. Journal of medical Internet research 2019;21(3):e12802.

2. Blease C, Bernstein MH, Gaab J, Kaptchuk TJ, Kossowsky J, Mandl KD, Davis RB, DesRoches CM. Computerization and the future of primary care: A survey of general practitioners in the UK. PloS one 2018;13(12):e0207418.

3. Haupt CE, Marks M. AI-generated medical advice—GPT and beyond. Jama American Medical Association; 2023;329(16):1349–1350.

4. Blease C, Torous J, McMillan B, Hägglund M, Mandl KD. Generative language models and

open notes: exploring the promise and limitations. JMIR Medical Education JMIR Publications Toronto, Canada; 2024;10:e51183.

5. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, Faix DJ, Goodman AM, Longhurst CA, Hogarth M. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA internal medicine American Medical Association; 2023;183(6):589–596.

6. Sharma A, Lin IW, Miner AS, Atkins DC, Althoff T. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. Nature Machine Intelligence Nature Publishing Group UK London; 2023;5(1):46–57.

7. Nuance. Automatically document care with DAX Copilot. 2024. Available from: https://www.nuance.com/healthcare/dragon-ai-clinical-solutions/dax-copilot/see-the-dragon-ambient-experience.html [accessed Aug 12, 2024]

8. Baker HP, Dwyer E, Kalidoss S, Hynes K, Wolf J, Strelzow JA. ChatGPT's Ability to Assist with Clinical Documentation: A Randomized Controlled Trial. JAAOS-Journal of the American Academy of Orthopaedic Surgeons LWW; 2022;10–5435.

9. Rosenberg GS, Magnéli M, Barle N, Kontakis MG, Müller AM, Wittauer M, Gordon M, Brodén C. ChatGPT-4 generates orthopedic discharge documents faster than humans maintaining comparable quality: a pilot study of 6 cases. Acta Orthopaedica Nordic Orthopaedic Federation; 2024;95:152.

10. Tu T, Palepu A, Schaekermann M, Saab K, Freyberg J, Tanno R, Wang A, Li B, Amin M, Tomasev N, Azizi S, Singhal K, Cheng Y, Hou L, Webson A, Kulkarni K, Mahdavi SS, Semturs C, Gottweis J, Barral J, Chou K, Corrado GS, Matias Y, Karthikesalingam A, Natarajan V. Towards Conversational Diagnostic AI. arXiv; 2024. Available from: http://arxiv.org/abs/2401.05654 [accessed Jan 18, 2024]

11. Kanjee Z, Crowe B, Rodman A. Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge. JAMA 2023;

12. McDuff D, Schaekermann M, Tu T, Palepu A, Wang A, Garrison J, Singhal K, Sharma Y, Azizi S, Kulkarni K, Hou L, Cheng Y, Liu Y, Mahdavi SS, Prakash S, Pathak A, Semturs C, Patel S, Webster DR, Dominowska E, Gottweis J, Barral J, Chou K, Corrado GS, Matias Y, Sunshine J, Karthikesalingam A, Natarajan V. Towards Accurate Differential Diagnosis with Large Language Models. arXiv; 2023. Available from: http://arxiv.org/abs/2312.00164 [accessed Apr 5, 2024]

13. Mehnen L, Gruarin S, Vasileva M, Knapp B. ChatGPT as a medical doctor? A diagnostic accuracy study on common and rare diseases. medRxiv Cold Spring Harbor Laboratory Press; 2023;2023–04.

14. Wojtara M, Rana E, Rahman T, Khanna P, Singh H. Artificial intelligence in rare disease diagnosis and treatment. Clinical Translational Sci 2023 Nov;16(11):2106–2111. doi: 10.1111/cts.13619

15. Hutto E. Dr. OpenAI Lied to Me. Medpage Today. 2023. Available from: https://www.medpagetoday.com/opinion/faustfiles/102723 [accessed Nov 11, 2023]

16. Ingram D. A mental health tech company ran an AI experiment on real users. Nothing's stopping apps from conducting more. NBC News. 2023. Available from: https://www.nbcnews.com/tech/internet/chatgpt-ai-experiment-mental-health-tech-app-koko-rcna65110 [accessed Aug 13, 2023]

17. Teno JM. Garbage in, Garbage out—Words of Caution on Big Data and Machine Learning in Medical Practice. JAMA Health Forum American Medical Association; 2023. p. e230397–e230397.

18. Gross N. What ChatGPT Tells Us about Gender: A Cautionary Tale about Performativity and Gender Biases in AI. Social Sciences MDPI; 2023;12(8):435.

19. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. BMC medical ethics 2017;18(1):19.

20. Zack T, Lehman E, Suzgun M, Rodriguez JA, Celi LA, Gichoya J, Jurafsky D, Szolovits P, Bates DW, Abdulnour R-EE. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. The Lancet Digital Health Elsevier; 2024;6(1):e12–e22.

21. Ferryman K, Mackintosh M, Ghassemi M. Considering Biased Data as Informative Artifacts in AI-Assisted Health Care. Drazen JM, editor. N Engl J Med 2023 Aug 31;389(9):833–838. doi: 10.1056/NEJMra2214964

22. Pierson E, Shanmugam D, Movva R, Kleinberg J, Agrawal M, Dredze M, Ferryman K, Gichoya JW, Jurafsky D, Koh PW, Levy K, Mullainathan S, Obermeyer Z, Suresh H, Vafa K. Use large language models to promote equity. arXiv; 2023. Available from: http://arxiv.org/abs/2312.14804 [accessed Feb 8, 2024]

23. Marks M, Haupt CE. AI Chatbots, Health Privacy, and Challenges to HIPAA Compliance. JAMA 2023;

24. Blease C. Open AI meets open notes: surveillance capitalism, patient privacy and online record access. Journal of Medical Ethics Institute of Medical Ethics; 2024;50(2):84–89.

25. Adams K. Epic to Integrate GPT-4 into Its EHR Through Expanded Microsoft Partnership. MedCity News. 2023. Available from: https://medcitynews.com/2023/04/epic-to-integrate-gpt-4-into-its-ehr-through-expanded-microsoft-partnership/ [accessed Jul 31, 2023]

26. American Psychiatric Association. The Basics of Augmented Intelligence: Some Factors Psychiatrists Need to Know Now. American Psychiatric Association. 2023. Available from: https://www.psychiatry.org/News-room/APA-Blogs/The-Basics-of-Augmented-Intelligence [accessed Aug 13, 2023]

27. AMA. ChatGPT and generative AI: What physicians should consider. American Medical

Association. 2023. Available from: https://www.ama-assn.org/system/files/chatgpt-what-physicians-should-consider.pdf [accessed Sep 11, 2023]

28. American Medical Association. Augmented intelligence in medicine. 2024. Available from: https://www.ama-assn.org/practice-management/digital/augmented-intelligence-medicine [accessed Aug 6, 2024]

29. Smith H, Downer J, Ives J. Clinicians and AI use: where is the professional guidance? Journal of Medical Ethics Institute of Medical Ethics; 2023; Available from: https://jme.bmj.com/content/early/2023/08/22/jme-2022-108831.abstract [accessed Apr 15, 2024]

30. National Institute for Health and Care Excellence. Artificial intelligence (AI) and digital regulations service. 2024. Available from: https://www.nice.org.uk/about/what-we-do/digital-health/multi-agency-advisory-service-for-ai-and-data-driven-technologies [accessed Aug 6, 2024]

31. NHS England. Artificial Intelligence. 2023. Available from: https://transform.england.nhs.uk/information-governance/guidance/artificial-intelligence/#:~:text=Although%20AI%2Dbased%20technology%20is,user%2C%20using%20your%20professional%20judgement. [accessed Aug 6, 2024]

32. European Commission. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance. ). 2017 Apr. Available from: https://op.europa.eu/en/publication-detail/-/publication/83bdc18f-315d-11e7-9412-01aa75ed71a1/language-en [accessed Aug 6, 2024]

33. Shryock T. AI Special Report: What patients and doctors really think about AI in health care. Medical Economics. 2023. Available from: https://www.medicaleconomics.com/view/ai-special-report-what-patients-and-doctors-really-think-about-ai-in-health-care [accessed Aug 22, 2023]

34. Blease C, Worthen A, Torous J. Psychiatrists' Experiences and Opinions of Generative Artificial Intelligence in Mental Healthcare: An Online Mixed Methods Survey. Psychiatry Research Elsevier; 2024;115724.

35. Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, MacDonald N, Holmes KL, Kho A. An exploratory survey about using ChatGPT in education, healthcare, and research. Plos one Public Library of Science San Francisco, CA USA; 2023;18(10):e0292216.

36. Taylor J. AMA calls for stronger AI regulations after doctors use ChatGPT to write medical notes. The Guardian 2023 Jul 27; Available from: https://www.theguardian.com/technology/2023/jul/27/chatgpt-health-industry-hospitals-ai-regulations-ama [accessed Aug 22, 2023]

37. Kolata G. When Doctors Use a Chatbot to Improve Their Bedside Manner. New York Times 2023 Jun 12; Available from: https://www.nytimes.com/2023/06/12/health/doctors-chatgpt-artificial-intelligence.html#:~:text=Most%20surprising%20to%20Dr.%20Lee,than%20waiting%20time%20or%20cost. [accessed Sep 20, 2023]

38. Doctors.net.uk. Doctors.net.uk. 2024. Available from: https://www.doctors.net.uk/ [accessed Feb 27, 2024]

39. General Medical Council. Register data summary. 2024. Available from: https://gde.gmc-uk.org/the-register/register-summary/register-data-summary [accessed Feb 28, 2024]

40. American Medical Association. AMA Augmented Intelligence Research Physician sentiments around the use of AI in heath care: motivations, opportunities, risks, and use cases. 2023 Nov. Available from: https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf [accessed Aug 6, 2024]

41. Blease CR, Kharko A, Dong Z, Jones RB, Davidge G, Hagglund M, Turner A, DesRoches C, McMillan B. Experiences and opinions of general practitioners with patient online record access: an online survey in England. BMJ open British Medical Journal Publishing Group; 2024;14(1):e078158.

42. Blease C, Locher C, Gaab J, Hagglund M, Mandl KD. Generative Artificial Intelligence in Primary Care: An online survey of UK General Practitioners. BMJ Health & Care Informatics In press;

43. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology 2006;3(2):77–101.

44. Topol E. Deep medicine: how artificial intelligence can make healthcare human again. Hachette UK; 2019.

45. Blease C, Kharko A, Locher C, DesRoches CM, Mandl KD. US primary care in 2029: A Delphi survey on the impact of machine learning. PloS one Public Library of Science San Francisco, CA USA; 2020;15(10):e0239947.

## Tables and Figures

**Table 1**. Respondents' characteristics.

| Characteristic | **Total** (*n* = 1,006) |
|---|---|
| **Gender**, *n* (%) | |
| Female | 467 (46.42%) |
| Male | 531 (52.78%) |

| | |
|---|---|
| Prefer not to say | 8 (0.8%) |
| **Age**, *n* (%) | |
| 35 or under | 76 (7.55%) |
| 36 to 45 | 386 (38.37%) |
| 46 to 55 | 348 (34.59%) |
| 56 or over | 196 (19.48%) |

**Table 2.** Practice characteristics.

| Characteristic | **Total** (*n* = 1,006) |
|---|---|
| **Practice location** | |
| Large town / city (e.g. Nottingham, Cardiff) | 182 (18.09%) |
| Major conurbation (e.g. London, Glasgow) | 172 (17.10%) |
| Medium town / city (e.g. Worcester, Dundee) | 224 (22.27%) |
| Small town / city (e.g. Thetford, Omagh) | 305 (30.32%) |
| Village / hamlet | 115 (11.43%) |
| Out of hours area | 3 (0.3%) |
| Rural / remote | 1 (0.1%) |
| Other * | 4 (0.4%) |
| **Practice size** | |
| Up to 5,000 patients | 117 (11.63%) |
| 5,001 – 7,500 patients | 149 (14.81%) |

| | |
|---|---|
| 7,501 – 10,000 patients | 213 (21.17%) |
| 10,001 – 12,500 patients | 183 (18.19%) |
| 12,501 patients or more | 344 (34.19%) |

* For practice location, free-text answers to option *'Other'* included: HMS Prison (*n* = 1), Health Board (*n* = 1), Military Garrison (*n* = 1), National (*n* = 1)

**Table 3.** Regional distribution of sample GPs compared to those in GMC Registry.

| | **Sample** (*n* = 1,006) | **GMC Registry** (*n* = 71,900) |
|---|---|---|
| **England** | | |
| North West | 121 (12.03%) | 7,750 (10.78%) |
| North East and Yorkshire | 131 (13.02%) | 8,683 (12.08%) |
| Midlands | 162 (16.1%) | 10,328 (14.36%) |
| East of England | 84 (8.35%) | 6,179 (8.59%) |
| London | 113 (11.23%) | 10,293 (14.32%) |
| South East | 130 (12.92%) | 9,652 (13.42%) |
| South West | 99 (9.84%) | 6,979 (9.71%) |
| **Northern Ireland** | 34 (3.38%) | 2,075 (2.89%) |
| **Scotland** | 84 (8.35%) | 6,920 (9.62%) |
| **Wales** | 48 (4.77%) | 3,041 (4.23%) |

**Table 4.** Differences between GPs' characteristics based on use of Generative AI.

| | **Used Gen AI** (*n* = 205) | **Not used Gen AI** (*n* = 801) |
|---|---|---|
| **Gender** | | |
| Female | 73 (35.6%) | 394 (49.2%) |
| Male | 129 (62.9%) | 402 (50.2%) |
| Prefer not to say | 3 (1.5%) | 5 (0.6%) |

| | | |
|---|---|---|
| **Age** | | |
| 35 or under | 20 (9.8%) | 56 (7%) |
| 36 to 45 | 76 (37.1%) | 310 (38.7%) |
| 46 to 55 | 68 (33.2%) | 280 (35.0%) |
| 56 or over | 41 (20.0%) | 155 (19.4%) |
| **Practice location** | | |
| Major conurbation | 43 (21.0%) | 129 (16.1%) |
| Large town/city | 45 (22.0%) | 136 (17.0%) |
| Medium town/city | 56 (27.3%) | 168 (21.0%) |
| Small town/city | 50 (24.4%) | 255 (31.8%) |
| Village/hamlet | 11 (5.3%) | 104 (13.0%) |
| Other | 0 (0%) | 9 (1.1%) |
| **Practice size** | | |
| up to 5,000 patients | 21 (10.2%) | 96 (12.0%) |
| 5,001-7,500 patients | 28 (13.7%) | 121 (15.1%) |
| 7,501-10,000 patients | 45 (22.0%) | 168 (21%) |
| 10,001-12,500 patients | 37 (18.0%) | 146 (18.2%) |
| 12,501 patients or more | 74 (36.1%) | 270 (33.7%) |

**Table 5.** Distribution of responses to expected litigation risks comparing GPs based on reported use of generative AI.

| Item | **Used Gen AI** ($n$ = 205) | **Not used Gen AI** ($n$ = 801) |
|---|---|---|
| **These tools will …** | | |
| decrease my risk of having legal action taken against me. | 13 (6.3%) | 9 (1.1%) |
| neither decrease nor increase my risk. | 72 (35.1%) | 129 (16.1%) |
| increase my risk of having legal action taken against me. | 64 (31.2%) | 276 (34.5%) |
| Don’t know. | 56 (27.3%) | 387 (48.3%) |

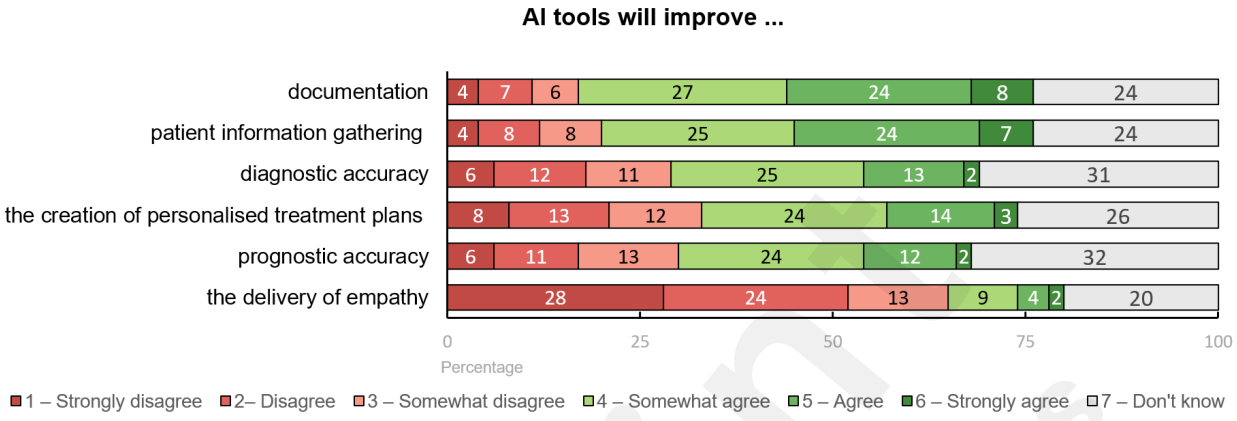


**Figure 1.** Distribution of responses to the potential for ChatGPT/Bard/Bing AI to improve clinical tasks, across the whole sample.

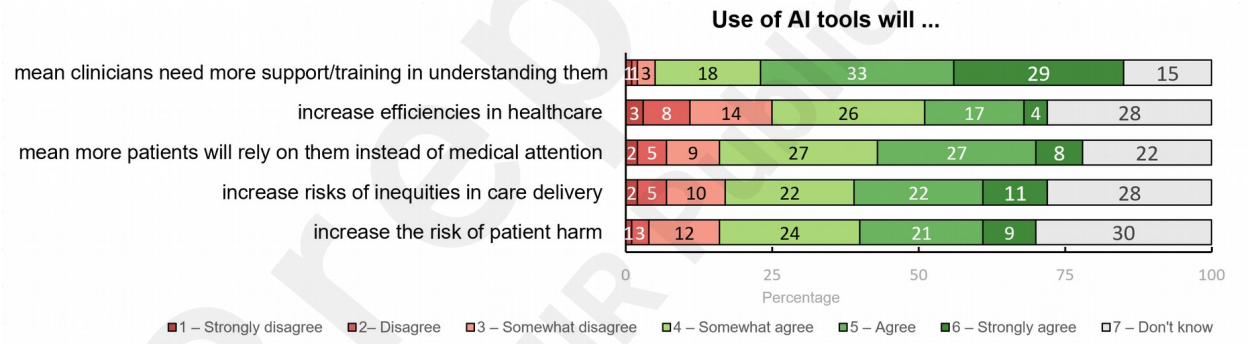


**Figure 2.** Distribution of responses to the potential impact of ChatGPT/Bard/Bing AI on care delivery, across the whole sample.

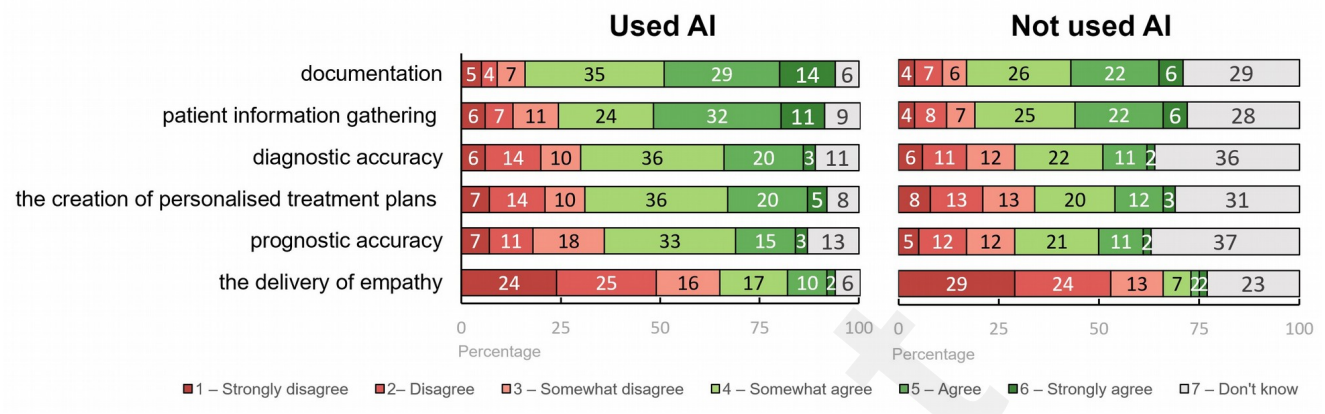


**Figure 3.** Distribution of agreement levels to the survey item "AI tools will improve..." between GPs who have used generative AI to assist in clinical practice and GPs who have not.

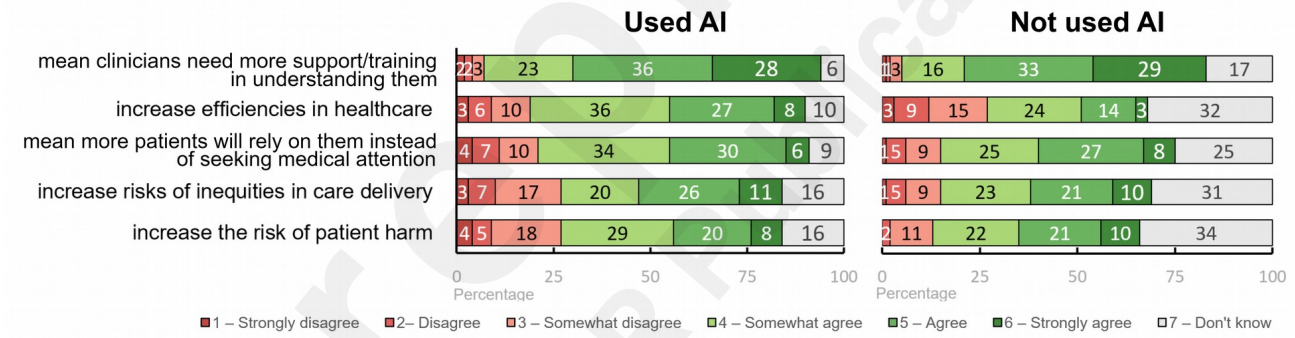


**Figure 4.** Distribution of agreement levels to the survey item "Use of AI tools will..." between GPs who have used generative AI to assist in clinical practice and GPs who have not.

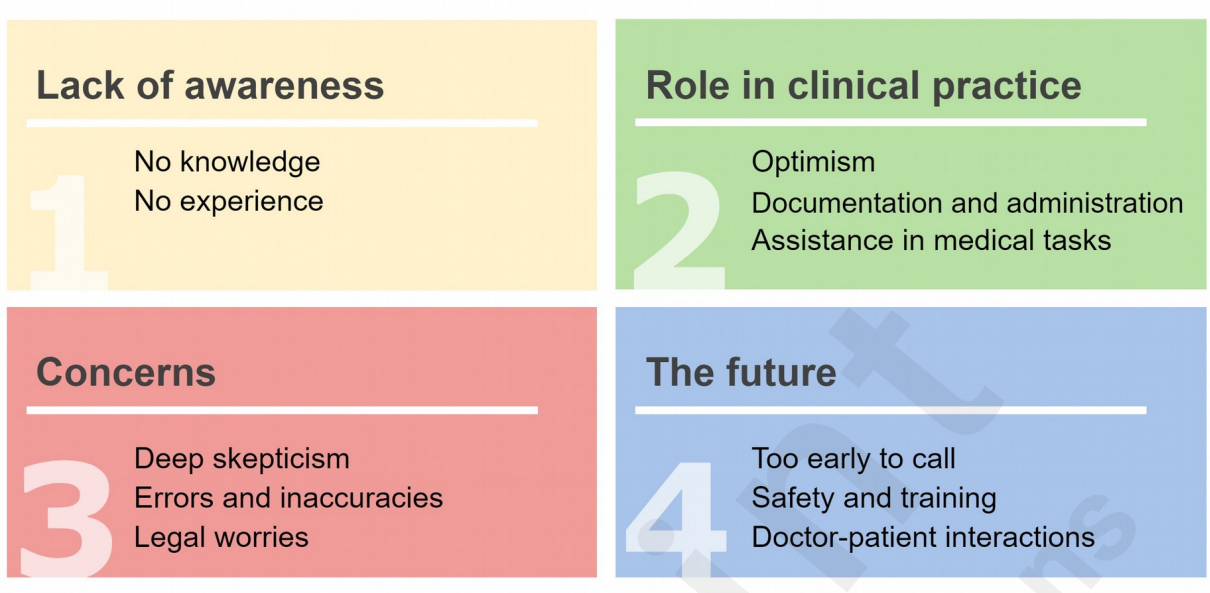


***Figure 5.*** *The 4 major categories of GPs' views in relation to generative AI in clinical practice.*

# Supplementary Files

# Multimedia Appendixes

Additional analyses reports.
URL: http://asset.jmir.pub/assets/3bd31634502848882c1a27206e8cdb95.docx

Online survey.
URL: http://asset.jmir.pub/assets/a36903a7be1f79569f7c92c495a36e65.docx

# CONSORT (or other) checklists

Appendix 2 Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
URL: http://asset.jmir.pub/assets/e10ab3b11abbdaeeb2444996dd638464.pdf