

Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study

Digital Health
Volume 5: 1–12
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-
permissions
DOI: 10.1177/2055207619871808
journals.sagepub.com/home/dhj



Tom Nadarzynski¹ , Oliver Miles², Aimee Cowie³ and Damien Ridge¹

Abstract

Background: Artificial intelligence (AI) is increasingly being used in healthcare. Here, AI-based chatbot systems can act as automated conversational agents, capable of promoting health, providing education, and potentially prompting behaviour change. Exploring the motivation to use health chatbots is required to predict uptake; however, few studies to date have explored their acceptability. This research aimed to explore participants' willingness to engage with AI-led health chatbots.

Methods: The study incorporated semi-structured interviews (N=29) which informed the development of an online survey (N=216) advertised via social media. Interviews were recorded, transcribed verbatim and analysed thematically. A survey of 24 items explored demographic and attitudinal variables, including acceptability and perceived utility. The quantitative data were analysed using binary regressions with a single categorical predictor.

Results: Three broad themes: 'Understanding of chatbots', 'AI hesitancy' and 'Motivations for health chatbots' were identified, outlining concerns about accuracy, cyber-security, and the inability of AI-led services to empathise. The survey showed moderate acceptability (67%), correlated negatively with perceived poorer IT skills OR = 0.32 [CI^{95%}:0.13–0.78] and dislike for talking to computers OR = 0.77 [CI^{95%}:0.60–0.99] as well as positively correlated with perceived utility OR = 5.10 [CI^{95%}:3.08–8.43], positive attitude OR = 2.71 [CI^{95%}:1.77–4.16] and perceived trustworthiness OR = 1.92 [CI^{95%}:1.13–3.25].

Conclusion: Most internet users would be receptive to using health chatbots, although hesitancy regarding this technology is likely to compromise engagement. Intervention designers focusing on AI-led health chatbots need to employ user-centred and theory-based approaches addressing patients' concerns and optimising user experience in order to achieve the best uptake and utilisation. Patients' perspectives, motivation and capabilities need to be taken into account when developing and assessing the effectiveness of health chatbots.

Keywords

Acceptability, AI, Artificial Intelligence, bot, chatbot

Received 12 March 2019; accepted 3 August 2019

Introduction

Artificial intelligence (AI) is an umbrella term for computer software consisting of a complex mathematical algorithm that processes input information to produce any specific pre-defined outputs, which lead to relevant outcomes.¹ AI systems, which utilise large datasets, can be designed to enhance decision-making and analytical processes while imitating human cognitive functions. AI has been applied in medicine and various healthcare services such as diagnostic imaging and genetic diagnosis, as well as clinical laboratory, screening and health communications.^{2,3} These systems aid physicians by

providing pertinent medical information in order to reduce diagnostic or therapeutic errors, and alerts about any high-risk health outcomes. The recent

¹The University of Westminster, London, UK

²University College London, London, UK

³The University of Southampton, Southampton, UK

Corresponding author:

Tom Nadarzynski, The University of Westminster, 115 New Cavendish Street, London, W1W 6UW.

Email: T.Nadarzynski@westminster.ac.uk

Twitter: @TNadarzynski



digitalisation of healthcare services in the UK offers access to large pools of clinical data such as medical notes, electronic records, physical and laboratory examination or patient demographic and behavioural characteristics.⁴ It is anticipated that by 2024, every patient in England would have digital access to primary care consultations, with a reduced need for face-to-face outpatient visits. In addition, there is an ongoing transformation to provide fully digitalised acute, community and mental health services across all locations. AI systems can utilise such clinical data to enhance diagnostic accuracy and enable clinicians to offer patient-centred medical care, while eliminating variations across the country and helping patients in managing their conditions themselves.

Chatbots, as part of AI devices, are natural language processing systems acting as a virtual conversational agent mimicking human interactions.⁵ While this technology is still in its developmental phase, health chatbots could potentially increase access to healthcare, improve doctor–patient and clinic–patient communication, or help to manage the increasing demand for health services such as via remote testing, medication adherence monitoring or teleconsultations.^{6–8} The chatbot technology allows for such activities as specific health surveys, setting up personal health-related reminders, communication with clinical teams, booking appointments, retrieving and analysing health data or the translation of diagnostic patterns taking into account behavioural indicators such as physical activity, sleep or nutrition.⁹ Such technology could potentially alter the delivery of healthcare systems, increasing uptake, equity and cost-effectiveness of health services while narrowing the health and well-being gap,¹⁰ but these assumptions require further research.

So far, chatbots have been applied in health education, diagnostics and mental health. A survey of conversational agents from 40 articles outlines chatbot taxonomy, specifies the main challenges and defines the types and contexts related to chatbots in health.¹¹ For example, chatbots can provide instant responses to health-related enquiries from patients while looking for specific patterns of symptoms in predicting disease, as presented by the internet-based Doc-Bot delivered via mobile phone or a Messenger-based chatbot for outpatient and translational medicine.¹² They can be tailored to specific populations, health conditions or behaviours. Crutzen et al. demonstrated the high engagement with a chatbot for adolescent students providing education on sex, drugs and alcohol.¹³ The users had positive views on the chatbot with emphasis on its anonymity as well as the quality and the speed of receiving information in comparison to popular search engines. The chatbot was seen as a reliable source of information; however, the ease of use was

rated as low, indicating challenges in implementing the technology on a larger scale. Other systems have also been proposed to act as a symptom checker,¹⁴ online triage service¹⁵ or health promotion assistant,¹⁶ providing live feedback in an interactive way. In addition, a number of studies have shown the usability of chatbot systems in mental health, in particular as a novel way of developing therapeutic and preventative interventions.^{17–19} For example, Ly et al. demonstrated the effectiveness of a chatbot based on cognitive behavioural therapy and positive psychology interactions in the non-clinical population.²⁰ There was a significant impact on well-being and perceived stress, with some participants reporting a specific ‘digital relationship’ with the chatbot. Nevertheless, chatbot systems are typically designed for specific functions, mainly to provide information. One of the main criticisms of chatbots is that they are not capable of empathy, notably to recognise users’ emotional states and tailor responses reflecting these emotions. The lack of empathy may therefore compromise the engagement with health chatbots.²¹

There is little research on health chatbot acceptability and motivations for its use. The acceptability of a healthcare intervention is a multi-faceted construct based on a range of dimensions including burden, values, effectiveness, cognition, and emotional responses.²² A study of 100 physicians in the US concluded that although the majority believed chatbots could assist with scheduling doctors’ appointments, locating health clinics and providing information about medication, over 70% also thought they cannot care for all patients’ needs, display emotion and could be a risk to patients due to incorrect self-diagnosis.²³ As these digital systems are capable of enhancing patient experiences of healthcare, and potentially influencing health behaviours, a theory-driven and person-centred approach is needed to inform their development and implementation. This study aimed to explore the acceptability of AI-led health chatbots to identify possible barriers and facilitators that could influence the delivery of these novel services. The findings are likely to inform the development of health chatbots using person-based approaches.

Method

Design

We used a mixed-methods approach²⁴ to assist in creative knowledge generation of a multi-layered issue.²⁵ Specifically, we incorporated face-to-face semi-structured interviews and an online survey format to explore the motivations for the use of chatbots in healthcare. Our interviews were guided by a topic schedule and informed the development of the exploratory survey distributed on social media. The study was approved

by the University of Southampton Ethics Committee (ref: 30986/31719).

Recruitment and data collection

Between November 2017 and January 2018, paper and digital adverts were distributed around the University of Southampton campus inviting students to take part in individual interviews in order to assess the attitudes towards new technologies in healthcare. Potential participants were asked to email researchers to arrange a suitable time and place for the interview. They were no specific exclusion criteria, although participants needed to be above the age of 18 years and capable of consenting to the study. There was no focus on any particular population in relation to healthcare utilisation, and this qualitative component of the study aimed to explore general views on health chatbots. As the advertisement strategy concentrated around the university settings, it was assumed that most participants had been familiar with conventional digital technologies.

The qualitative arm of the study was guided by a topic schedule, which was based on the theoretical framework of healthcare intervention acceptability²² that was adopted for health chatbots. The schedule consisted of five sets of open-ended questions exploring the understanding of chatbots, attitudes, usability and general concerns. The semi-structured interviews took place in a room at the University of Southampton. The interviews were conducted by two trained researchers who were also involved in transcription, analysis and data validation. All participants were asked to sign a consent form and were reminded about their rights for confidentiality and that they could withdraw from the study at any time without penalty. It was assumed that many participants had not experienced using a chatbot in the past; thus a chatbot demonstration was performed during the interview. The participants were asked to conduct a live conversation with a popular chatbot²⁶ in order to gather more credible views on chatbot acceptability. The interaction allowed participants to post any questions for the chatbot and gain immediate answers. The interviews lasted 20–30 minutes, were audio-recorded and transcribed verbatim. No incentive was offered to participants.

Between February and June 2018, an advert for the online survey was distributed on social media pages (i.e. university accounts on Facebook, Twitter and eFolio such as student union) inviting users to complete a short questionnaire about health chatbots. No particular health-specific pages were targeted for the advertisement; however, this quantitative arm of the study used a digital snowball sampling method encouraging users to share the study advert on their social media profiles. This method is likely to represent

views of the internet users more familiar with social media, although no specific populations were targeted. The participation was voluntary, and the respondents were offered a chance to enter a prize draw worth £50.

The internet users were directed to the survey after clicking on a pre-designed online advert. They were shown information about the study and asked to provide online consent by ticking a box. The survey took about 10 minutes to complete. The online survey consisted of 24 items, both demographic and attitudinal. It was developed based on the theoretical framework of acceptability²² and the findings from the qualitative interviews. The participants were asked general questions about the awareness and experience of chatbots, and more specifically about health chatbots. They were then presented with two sets of questions examining the perceived usefulness of health chatbots and their general attitudes. The perceived usefulness questions, assessed using a 5-point Likert scale (from ‘extremely unlikely’ to ‘extremely likely’), asked participants to rate their willingness to use chatbots for seeking general health information, information about medication, various diseases, potential symptoms, seeking results of medical tests, booking a medical appointment and looking for specialist medical services. The attitudinal questions, assessed using a 5-point Likert scale (from ‘strongly disagree’ to ‘strongly agree’), asked participants to indicate their agreement with 16 statements about their healthcare such as the worry about digital privacy, the accuracy of health information online, the preference for face-to-face interaction and trust in advice from a health chatbot. The main outcome measure – health chatbot acceptability – was assessed using one question: ‘How likely would you be to use a health chatbot in the next 12 months if it was available to you today?’ with five options (from ‘extremely unlikely’ to ‘extremely likely’).

Data analysis

Thematic analysis²⁷ was conducted on qualitative data to identify common patterns and trends. Two researchers familiarised themselves with the data by repeatedly reading the transcript to enhance understanding. The analyses were conducted independently using NVIVO software where data were coded and recorded to then be categorised into meaningful themes and subthemes. The results of the analysis were discussed between the researchers to find an agreement for the final set of findings. The themes were then validated by two researchers comparing quotes with the identified themes.

Descriptive and inferential statistics were conducted on quantitative data. All variables were dichotomised, and neutral values excluded in order to perform binary logistic regressions with a single categorical predictor

to determine the correlates of health chatbot acceptability. The model was not adjusted as it did not meet the statistical assumptions, due to multicollinearity and non-binomial distribution of responses. However, the regression allowed assessing the correlates of chatbot acceptability and their directions. The odds ratio and 95% confidence intervals were presented as the magnitude of association with the outcome variable in an explorative manner.

Results

Sample characteristics

In our qualitative sub-study, 29 participants (all university students, 24 self-identified as White and 15 as women) aged 18–22 years were interviewed. In our quantitative sub-study, 215 users completed the survey. The mean age was 30 years ($SD = 12$, range: 18–62) and the majority were women (61%), of White ethnicity (64%) and educated below the university degree (54%). Most (76%) rated their IT skills as ‘good’ or ‘very good’, and many reported looking for medical information online a few times per year (41%) or every month (33%).

Qualitative analysis

Qualitative data were organised into three themes: ‘Understanding of chatbots’, ‘AI Hesitancy’ and ‘Motivations for health chatbot’. Table 1 presents all themes and subthemes with corresponding quotes.

Understanding of chatbots

Most participants reported hearing about bots, notably in the context of social media or customer service, but were unsure how they functioned technologically. Due to limited experience with chatbots, the majority were unable to recall if they ever used one for their healthcare. After the chatbot demonstration, the participants appreciated the mainstream chatbot systems available, such as Alexa or Google Home, in particular in relation to information searches. However, they agreed that this technology was still emergent, and not part of the mainstream culture, despite large media coverage about AI. There was a general lack of familiarity and understanding of health chatbots amongst participants.

AI hesitancy

Many participants were hesitant about whether they would incorporate chatbots as part of their healthcare. They were uncertain about the quality, trustworthiness and accuracy of the health information provided by chatbots, as the sources underpinning such services

were not transparent. The majority of participants reported not being able to understand the technological complexity of chatbots, in particular how they are able to correctly respond to a health enquiry. There was a doubt about whether a chatbot could correctly identify symptoms of less common health conditions or diseases. A number of participants emphasised the potential for miscommunication between a chatbot and its users, who might not be able to accurately describe their health issues or name symptoms. There was a perception of a risk of harm if the information provided by a chatbot was inaccurate or inadequate. In general, there was a view that this technology was premature in terms of providing a diagnosis, as it was seen as an ‘unqualified’. However, most participants found receiving general health advice acceptable.

While a few participants thought that well-designed chatbots can be more accurate and logical compared with doctors, the lack of human presence was seen as the main limitation. In particular, participants worried about a lack of empathy and inability of chatbots to understand more emotional issues, notably in mental health. The responses given by chatbots were seen as depersonalised, cold and inhuman. They were perceived as inferior to doctor consultation, although several participants admitted that this technology offered a level of anonymity which could facilitate the disclosure of more intimate or uncomfortable aspects to do with health. Other participants were concerned about cyber-security and the ability (or not) of chatbots to maintain confidentiality so that their sensitive health-related information was protected from potential hacking or data leakage. There was also a concern that health chatbots could reduce the overall quality of healthcare if they were to replace experienced trained professionals.

Motivations for health chatbots

The majority of participants were willing to use chatbots for minor health concerns that would not require a physical examination. They were perceived as a convenient tool that could facilitate the seeking of health information online. Several participants compared chatbots to medical phone helplines, such as NHS Direct, that provide rapid guidance and health advice on minor health issues. They perceived chatbots to be particularly useful when they might struggle to comprehend the advice given via telephone, seeing written information as easier to understand. Some expressed preferences for a web-chat format of conversation. Thus, if free at the point of access, chatbots were seen as time-saving and useful platforms for triaging users to appropriate healthcare services.

Table 1. Quotes from thematic analysis on the motivations for AI-led chatbots in healthcare.

| Theme (sub-theme) | Illustrative quotes |
|---------------------------------|---|
| Understanding of chatbots | |
| (Awareness) | <p>“I think that it’s online and you ask it questions and it can reply to you with information. It is not a real person. It is like stored information.”</p> <p>“There are those diagnosis bots and you can use. Those to get an idea of what is wrong with you or what your next step could be if it can just tell you a rough severity of what you may have. Then I suppose it is useful in those cases but, in terms of just chatting to something and telling it your problems or whatever and trying to get a diagnosis. If you are using it in that kind of way then I think that is a very limited thing to try and do.”</p> <p>“I’ve used one for banking before, and you just type in your query. And I think I’ve used some which you type to and it picks out keywords, and one which you can select the most appropriate response.”</p> |
| (Experience) | <p>“Where it talks back to you, it can be more specific, compared to Google where you have to search and look through.”</p> |
| AI hesitancy | |
| (Perceived accuracy) | <p>“It’s not that I think they [chatbots] would intentionally give me false information. I just don’t know how accurate they are. I don’t know whether they can be as accurate as doctors can be.”</p> <p>“I would find it hard to trust a health chatbot because it is just looking online at things. You would want a professional opinion.”</p> |
| (Premature technology) | <p>“I think it has a lot of potential in the future, but now certain places are releasing it before they have perfected their own system, then it could put people off. Because you can end up chatting for like half an hour and go back to being at the same question you were at in the first place. With that people get angry.”</p> <p>“I don’t know whether the technology is as adequate as a doctor.”</p> |
| (Non-human interaction) | <p>“I think a lot of people would be put off just being with a chatbot. It is like a segregation thing, I don’t think it will replace human interaction.”</p> <p>“If you are looking at a chatbot thinking that it is a replacement for a person then you are looking it in the wrong way. If you were looking for a deep and meaningful conversation you are not going to find one.”</p> |
| (Cyber-security) | <p>“Some people might find issues with confidentiality because if you were with the doctor it is just you and them, but with chatbots, you don’t know who is behind it all.”</p> <p>“Some things are confidential and you wouldn’t just type it on the internet. You would want the confidentiality of the GP practice.”</p> |
| Motivations for health chatbots | |
| (Anonymity) | <p>“I think for mental health it would be pretty useful because I think that it’s a lot harder to talk to a real person about that. Maybe sexual health too. I’m pretty open generally about both of those things but I can see where they might be seen as a better alternative due to privacy and not having to face a person and describe all of these problems.”</p> <p>“I think mental health would be a good thing to use a chatbot in, because some people with mental health issues they do not want to open up to an actual person, so it would be easier doing it over the internet in the comfort of their own home.”</p> |
| (Convenience) | <p>“You can use a chatbot instead of googling it and reading advice on the NHS page. If that information is integrated into the chatbot then yeah it would certainly save time.”</p> <p>“It would be good for healthcare because obviously, it is so hard to get a doctor’s appointment, so for people who have general queries like it would be good for them to quickly get advice on whether or not it is an urgent issue.”</p> |
| (Sign-posting) | <p>“Chatbot can tell you if you need further advice or you will worry. And it will reassure you so you don’t go to the doctors.”</p> <p>“People who have no awareness of or are perhaps too worried, it could be a good way to get in touch. Just like if you use 111, first of all, you go to a chatbot and they would figure out what is wrong and if there was a severity to it they would ring 999 and get an ambulance, otherwise, they can direct you to a GP and a clinic.”</p> |

Quantitative analysis

Table 2 presents sample characteristics and correlates of health chatbot acceptability amongst 215 participants. The analysis showed that while 6% had heard of a health chatbot and 3% had experience of using it, 67% perceived themselves as likely to use one within 12 months. None of the demographic variables was associated with acceptability, although those who perceived themselves to have poor or moderate IT skills showed lower acceptability. The majority of participants would use a health chatbot for seeking general health information (78%), booking a medical appointment (78%) and looking for local health services (80%). However, a health chatbot was perceived as less suitable for seeking results of medical tests and seeking specialist advice such as sexual health. All nine items measuring perceived utility were associated with chatbot acceptability with the highest levels reported for seeking general health information as well as the information about symptoms and medication. The analysis of attitudinal variables showed that most participants reported their preference for discussing their health with doctors (73%) and having access to reliable and accurate health information (93%). While 80% were curious about new technologies that could improve their health, 66% reported only seeking a doctor when experiencing a health problem and 65% thought that a chatbot was a good idea. Interestingly, 30% reported dislike about talking to computers, 41% felt it would be strange to discuss health matters with a chatbot and about half were unsure if they could trust the advice given by a chatbot. Nine attitudinal items were associated with acceptability, with perceived trust and the belief that a chatbot was a good idea being the strongest predictors.

Discussion

To our knowledge, this is the first study exploring the acceptability of AI-led chatbot systems for healthcare from the perspective of the general public with no pre-existing medical conditions. The awareness and experience of health chatbots were low amongst our participants, and most had mixed attitudes towards these novel technologies. The qualitative analysis showed that a substantial proportion was hesitant to AI and health chatbots, mainly because of concerns about the accuracy and security of these services. There was also a view that chatbots could enable some users to discuss their intimate and perhaps embarrassing health issues, promoting access to professional health services. Although they were seen as a convenient and anonymous tool for minor health issues that may carry a level of stigma, the lack of empathy and professional human approach made chatbots less acceptable to

some users. The survey demonstrated that the participants were more willing to use these systems to find general health information over finding out the results of medical tests or specialist advice. Amongst the strongest predictors of acceptability were positive attitudes towards health chatbots and the curiosity about new technologies that could improve health. Also, those who showed dislike for talking to chatbots and preferred to discuss their health face-to-face with a clinician were less likely to accept chatbots. Although these innovative services were acceptable by the majority of participants, we propose that 'AI hesitancy' would have a negative influence on the engagement and effectiveness of those technologies. Therefore, the patient perspective needs to be taken into consideration when developing AI-enabled health services.

These findings are consistent with previous research and theoretical frameworks on the acceptability of novel health interventions. The acceptability rate in the present study is comparable to the acceptability level of 49% within a Dutch cohort offered use of a chatbot for smoking cessation.¹⁶ Nevertheless, Laufer has already argued that the social acceptability of AI systems is compromised by the ambiguous status of 'artificial', which has negative and 'inferior-to-natural' connotations.²⁸ According to the Diffusion of Innovation theories, the implementation of new technologies is a process in which the adoption is dependent on widespread awareness, understanding and utilisation.²⁹ Adopters are generally divided into innovators, early adopters, majority users and laggards. While the passage of time is necessary for any innovation to be adopted, certain characteristics of social systems such as governmental endorsement, mass media campaigns or personal views of social role models are likely to influence potential adopters. The theoretical framework of acceptability²² outlines that the burden of engaging with the intervention, ethical consequences and negative user experiences are likely to increase hesitancy or even lead to the failure of the intervention. Hence, the concerns about accuracy, trustworthiness and privacy, as well as the perceived lack of empathy, are likely to compromise the adoption of AI systems in healthcare. Therefore, user-centred approaches, for example incorporating qualitative methodologies or 'A/B testing' techniques, are necessary to overcome potential barriers to engagement.³⁰ These approaches require a thorough investigation into the awareness, comprehension and motivation for the use of novel health interventions. The specific personal agency, intervention content and quality as well as the 'user experience', notably the interaction and perceived support, need to be studied and optimised for best uptake.

It is important to acknowledge that users perceived several benefits of health chatbots, notably in relation

Table 2. Sample characteristics and predictors of health chatbot acceptability.

| Variable | Total of the sample (%) | (%) of those 'likely' to use health chatbot within 12 months/Odds ratio [95% CI] |
|---|-------------------------|--|
| Age [mean, SD] | [30, 12] | |
| Below 25 years | 113 (53) | (89)/2.07 [0.87–4.93] |
| 25 years and above | 102 (47) | (80)/Ref |
| Gender | | |
| Male | 84 (39) | (87)/1.36 [0.55–3.38] |
| Female | 131 (61) | (84)/Ref |
| Ethnicity | | |
| White | 138 (64) | (83)/0.49 [0.17–1.39] |
| Non-white | 77 (36) | (90)/Ref |
| Education | | |
| Below university degree | 116 (54) | (84)/0.79 [0.33–1.87] |
| University degree | 99 (46) | (86)/Ref |
| Perceived IT skills | | |
| Poor or moderate | 51 (24) | (72)/0.32 [0.13–0.78]* |
| Good or very good | 164 (76) | (89)/Ref |
| Health information seeking | | |
| Several times per year | 103 (48) | (83)/0.72 [0.31–1.70] |
| Every month or more often | 112 (52) | (87)/Ref |
| Health chatbot awareness | | |
| Yes | 12 (6) | Assumption not met ^a |
| No | 203 (94) | |
| Past health chatbot use | | |
| Yes | 7 (3) | Assumption not met ^a |
| No | 202 (97) | |
| Likelihood of using health chatbot within 12 months if available (acceptability) | | |
| Likely | 143 (67) | |
| Neutral | 45 (21) | |
| Unlikely | 25 (12) | |

(continued)

Table 2. Continued.

| Variable | Total of the sample (%) | (%) of those 'likely' to use health chatbot within 12 months/Odds ratio [95% CI] |
|--|-------------------------|--|
| Perceived utility of health chatbots[#] | | |
| To seek general health information | | |
| Likely | 168 (78) | (98)/5.10 [3.08-8.43]* |
| Unlikely | 24 (11) | (8)/Ref |
| To seek information about medication | | |
| Likely | 128 (60) | (99)/3.21 [1.92-5.37]* |
| Unlikely | 52 (24) | (49)/Ref |
| To seek information about diseases | | |
| Likely | 148 (69) | (97)/2.97 [2.10-4.10]* |
| Unlikely | 38 (18) | (33)/Ref |
| To seek information about symptoms | | |
| Likely | 144 (67) | (98)/3.44 [2.25-5.24]* |
| Unlikely | 28 (13) | (28)/Ref |
| To seek results of medical tests | | |
| Likely | 83 (39) | (92)/1.42 [1.08-1.85]* |
| Unlikely | 81 (38) | (75)/Ref |
| To book a medical appointment | | |
| Likely | 167 (78) | (92)/1.88 [1.46-2.42]* |
| Unlikely | 34 (16) | (50)/Ref |
| To look for local medical services (e.g. pharmacy) | | |
| Likely | 172 (80) | (92)/2.25 [1.64-3.08]* |
| Unlikely | 17 (8) | (33)/Ref |
| To seek specialist advice (e.g. sexual health) | | |
| Likely | 104 (48) | (98)/2.69 [1.61-4.49]* |
| Unlikely | 65 (30) | (61)/Ref |
| Beliefs associated with chatbot acceptability[#] | | |
| Worried about health | | |
| Agree | 107 (50) | (81)/0.94 [0.74-1.19] |
| Disagree | 70 (33) | (85)/Ref |

(continued)

Table 2. Continued.

| Variable | Total of the sample (%) | (%) of those 'likely' to use health chatbot within 12 months/Odds ratio [95%; CI] |
|---|-------------------------|---|
| Worried about privacy using a health chatbot | | |
| Agree | 100 (47) | (93)/1.42 [1.08-1.86]* |
| Disagree | 71 (33) | (77)/Ref |
| Worried about the security of information | | |
| Agree | 100 (47) | (93)/1.36 [1.02-1.81]* |
| Disagree | 71 (33) | (80)/Ref |
| Confident in finding accurate health information online | | |
| Agree | 104 (48) | (87)/1.31 [1.03-1.67]* |
| Disagree | 50 (23) | (70)/Ref |
| Confident in identifying own health symptoms | | |
| Agree | 146 (68) | (88)/1.20 [0.90-1.60] |
| Disagree | 32 (15) | (78)/Ref |
| Comfortable with outlining symptoms to a chatbot | | |
| Agree | 131 (61) | (91)/1.46 [1.07-1.99]* |
| Disagree | 26 (12) | (68)/Ref |
| Prefer to talk face to face with a doctor about health | | |
| Agree | 158 (73) | (83)/Assumption not meta |
| Disagree | 19 (9) | (100) |
| I don't like talking to computers | | |
| Agree | 64 (30) | (77)/0.77 [0.60-0.99]* |
| Disagree | 103 (48) | (90)/Ref |
| It would be strange to talk to a chatbot about health | | |
| Agree | 88 (41) | (76)/0.72 [0.54-0.97]* |
| Disagree | 68 (32) | (93)/Ref |
| Health chatbot could help to make better decisions | | |
| Agree | 65 (30) | (100)/Assumption not meta |
| Disagree | 42 (19) | (62) |

(continued)

Table 2. Continued.

| Variable | Total of the sample (%) | (%) of those 'likely' to use health chatbot within 12 months/Odds ratio [95% CI] |
|---|-------------------------|--|
| Would trust advice from a health chatbot | | |
| Agree | 59 (27) | (98)/1.92 [1.13-3.25]* |
| Disagree | 54 (25) | (78)/Ref |
| A health chatbot is a good idea | | |
| Agree | 139 (65) | (93)/2.71 [1.77-4.16]* |
| Disagree | 13 (6) | (20)/Ref |
| Willing to enter symptoms on an online form | | |
| Agree | 136 (63) | (91)/1.34 [0.95-1.86] |
| Disagree | 24 (11) | (76)/Ref |
| Curious how new technologies could improve health | | |
| Agree | 172 (80) | (88)/1.60 [1.15-2.21]* |
| Disagree | 15 (7) | (54)/Ref |
| Reliable and accurate information is important | | |
| Agree | 199 (93) | (85)/Assumption not met |
| Disagree | 4 (2) | (0) |
| Only seek a doctor if I have a health problem | | |
| Agree | 141 (66) | (84)/0.81 [0.55-1.19] |
| Disagree | 33 (15) | (92)/Ref |

*Significant at $p < 0.05$; SD: standard deviation; CI: confidence intervals; #Neutral values removed for the binary regression analysis; Ref: reference category for binary regression; ^aStatistical assumptions required to perform binary logistic regression with a single categorical predictor were not met.

to anonymity, convenience and faster access to relevant information. This is in line with previous research showing that users might be as likely to disclose emotional and factual information to a chatbot as they do to a human partner.³¹ The perceived understanding, disclosure intimacy and cognitive reappraisal were similar in conversations conducted with chatbots and humans, indicating that people psychologically engage with chatbots as they do with people. The perceived anonymity was noted by a few participants in sexual health and mental health settings, although the preferences for particular chatbot use in healthcare settings needs to be explored further. Our analysis also supports the findings from a qualitative study exploring user expectations of chatbots in terms of their understanding and preferences.³² Users are generally unclear

about what chatbots can do, although they foresee this technology as improving their experience by providing immediate access to relevant and valuable information. It was also shown that users saw the lack of judgement as a unique aspect of this technology, although it was noted that building rapport with a chatbot would require trust and meaningful interactions. These motivations for the use of chatbots need to be explored in more detail in order to understand how this technology could be safely incorporated into healthcare.

The present study had a number of methodological issues. The use of mixed methods allowed new concepts to be tested in the online survey, as the qualitative analysis of views on AI-led chatbots fed into the quantitative arm of the study. In addition, the demonstration of the popular chatbot and the opportunity for participants to

directly interact with it likely strengthen the validity of the findings, as the explored views were not purely hypothetical but based on experience during the study. The qualitative sub-study also informed the development of the exploratory survey, increasing its reliability, although a further investigation into the development of measurement tool is needed. Nevertheless, the survey responses were mainly drawn from students and internet users, in particular, a young and educated cohort that might not be representative of the population that might be asked to use health chatbots. It is likely that these answers are more typical of people who are relatively experienced with digital technologies. User 'perceived IT skill' was a correlate of chatbot acceptability; thus future studies need to assess the willingness to use these technologies in clinical and community-based populations. A perspective on patient acceptability of chatbots in those who experience acute and chronic conditions would enhance the understanding of the feasibility of this intervention within healthcare systems. In addition, the digital snowball sampling method through social media, as used in the present study, is likely to compromise the generalisability of the findings if participants were selected based on comparable characteristics. Thus, subsequent assessments of chatbot acceptability should employ robust methodological design able to capture diverse views representative of potential healthcare users. It would also be useful to examine the acceptability of specialist chatbots that serve a particular population or in relation to a specific condition, as well as general chatbots used as a triage tool. Various designs of chatbots, notably when the health information is stored and retrieved or when chatbots are fully conversational, could also affect acceptability and engagement.

There are several implications of this study. AI intervention designers need to include opinions of users and health professionals to maximise engagement and retention. No AI-led health chatbot should be implemented without rigorous piloting that can address patients' concerns and remove any potential barriers. As a large number of participants reported the preference for face-to-face interaction, health chatbots should be a supplementary service rather than a replacement of the professional health force. While for some users they might be perceived as a reduction in the quality of care, others might see chatbots as an improvement, notably in overcoming 'shameful' issues. Thus, their mechanism of action and clinical effectiveness as an intervention should be clearly and transparently communicated to all users. Intervention designers should reassure users of the human dimensions of AI systems that are developed to improve health and well-being in order to increase the acceptability of these services.

This study has identified the concept of 'AI hesitancy'. As outlined, concerned about accuracy, cyber-

security, lack of empathy and technological maturity were reported as potential factors associated with the delay in acceptability or refusal. The construct of 'hesitancy' has been applied in various acceptability studies, notably in vaccinations, mainly referred to the level of confidence, complacency and convenience.³³ Although the constructs from the vaccine hesitancy model could potentially overlap with AI hesitancy, future research is needed to further define and operationalise this concept in order to have a precise understanding of motivations for patient engagement with AI systems. As there is a substantial investment in the development of AI in healthcare, purely driven by the need for cost-effectiveness, it is essential to produce a theory contributing to its design.

In conclusion, as the application of AI chatbot services in healthcare is becoming more apparent, service users' motivation, uptake and engagement need to be evaluated to maximise the benefits from these technologies. At present, we identified that many are receptive to health chatbots, but a substantial number may feel hesitant to use AI modules. Intervention designers need to apply user-centred and theory-based approaches in order to address user concerns and develop effective and ethical services, capable of reducing the gap in health and well-being. Future studies are required to explore how health chatbots could be used in preventative medicine and healthcare utilisation, notably by allowing patients to engage with their health.

Acknowledgements: We thank Wisman Siew and Elizabeth Simpson from the University of Southampton for their help with data collection.

Conflict of interests: The authors declare that there is no conflict of interest.

Contributorship: TN, OM and AC designed the study and were involved in the development of the study protocol and gaining ethical approval. All authors were involved in data collection and analysis as well as the write-up of the final manuscript.

Ethical approval: The ethics committee of the University of Southampton approved this study (REC number: 30986/31719)

Funding: The authors received no specific funding for this work.

Guarantor: Dr Tom Nadarzynski

ORCID iD

Tom Nadarzynski  <https://orcid.org/0000-0001-7010-5308>

Peer review: This manuscript was reviewed by a single individual, who has chosen to remain anonymous.

References

- Bostrom N and Yudkowsky E. The ethics of artificial intelligence. In: Frankish K and Ramsey WM (eds) *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, 2014, pp.316–334.
- Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc Neurol* 2017; 2(4): 230–243.
- He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019; 25(1): 30–36.
- Iacobucci G. NHS long term plan: Care to be shifted away from hospitals in “21st century” service model. *BMJ* 2019; 364: 185.
- Ivanovic M and Semnic M. The Role of Agent Technologies in Personalized Medicine. In: *2018 5th International Conference on Systems and Informatics (ICSIAI)*, 2018, IEEE, pp.299–304.
- Hoermann S, McCabe KL, Milne DN, et al. Application of synchronous text-based dialogue systems in mental health interventions: Systematic review. *J Med Internet Res* 2017; 19(8): e267.
- Fadhil A and Gabrielli S. Addressing challenges in promoting healthy lifestyles: the ai-chatbot approach. In: *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, ACM, 2017, pp.261–265.
- Comendador BEV, Francisco BMB, Medenilla JS, et al. Pharmabot: A pediatric generic medicine consultant chatbot. *J Automat Control Eng* 2015; 3(2): 137–140.
- Abashev A, Grigoryev R, Grigorian K, et al. Programming tools for messenger-based chatbot system organization: Implication for outpatient and translational medicines. *BioNanoScience* 2017; 7(2): 403–407.
- Harwich E and Laycock K (2018). Thinking on its Own—AI in the NHS. (Reform). <https://reform.uk/research/thinking-its-own-ai-nhs> (accessed 2 November 2018).
- Montenegro JLZ, da Costa CA and da Rosa Righi R. Survey of conversational agents in health. *Exp Syst Appl* 2019; 129: 56–67.
- Tripathy AK, Carvalho R, Pawaskar K, et al. Mobile based healthcare management using artificial intelligence. In: *2015 International Conference on Technologies for Sustainable Development (ICTSD)*, IEEE, 2015, pp.1–6.
- Crutzen R, Peters GJY, Portugal SD, et al. An artificially intelligent chat agent that answers adolescents’ questions related to sex, drugs, and alcohol: An exploratory study. *J Adolesc Health* 2011; 48(5): 514–519.
- Ghosh S, Bhatia S and Bhatia A. Quro: Facilitating user symptom check using a personalised chatbot-oriented dialogue system. *Stud Health Technol Inform* 2018; 252: 51–56.
- Razzaki S, Baker A, Perov Y, et al. A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis. *arXiv preprint arXiv* 2018; 1806: 10698.
- Grolleman J, van Dijk B, Nijholt A, et al. Break the habit! Designing an e-therapy intervention using a virtual coach in aid of smoking cessation. In: *Persuasive Technology*. Springer: Berlin, Heidelberg, 2006, pp.133–141.
- Oh KJ, Lee D, Ko B, et al. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In *18th IEEE International Conference on Mobile Data Management (MDM)*, 2017. IEEE, 2017, pp.371–375.
- Elmasri D and Maeder A. A conversational agent for an online mental health intervention. In *International Conference on Brain and Health Informatics*. Springer: Cham. 2016, pp.243–251.
- Suganuma S, Sakamoto D and Shimoyama H An embodied conversational agent for unguided internet-based cognitive behavior therapy in preventative mental health: Feasibility and acceptability pilot trial. *JMIR Ment Health* 2018; 5(3): e10454.
- Ly KH, Ly AM and Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interv* 2017; 10: 39–46.
- Morris RR, Kouddous K, Kshirsagar R, et al. Towards an artificially empathic conversational agent for mental health applications: System design and user perceptions. *J Med Internet Res* 2018; 20(6): e10148.
- Sekhoni M, Cartwright M and Francis JJ. Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017; 17(1): 88.
- Palanica A, Flaschner P, Thommandram A, et al. Physicians’ perceptions of chatbots in healthcare: A cross-sectional web-based survey. *J Med Internet Res* 2019; 21(4): e12887.
- Denscombe M Communities of practice: A research paradigm for the mixed methods approach. *J Mixed Method Res* 2008; 2(3): 270–283.
- Mason J. Mixing methods in a qualitatively driven way. *Qual Res* 2006; 6(1): 9–25.
- Gehl RW. Teaching to the Turing Test with Cleverbot. *Transformations* 2014; 24(1-2): 56–66.
- Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3(2): 77–101.
- Laufer R. The social acceptability of AI systems. *Artif Intell Crit Concept* 2000; 4(1992): 343.
- Ward R. The application of technology acceptance and diffusion of innovation models in healthcare informatics. *Health Policy Technol* 2013; 2(4): 222–228.
- Yardley L, Morrison L, Bradbury K, et al. The person-based approach to intervention development: Application to digital health-related behavior change interventions. *J Med Internet Res* 2015; 17(1): e30.
- Ho A, Hancock J and Miner AS. Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *J Commun* 2018; 68(4): 712–733.
- Zamora J. I’m Sorry, Dave, I’m Afraid I Can’t Do That: Chatbot Perception and Expectations. In: *Proceedings of the 5th International Conference on Human Agent Interaction*. ACM, 2017, pp.253–260.
- Dubé E, Laberge C, Guay M, et al. Vaccine hesitancy: An overview. *Human Vaccin Immunother* 2013; 9(8): 1763–1773.