CHAPTER

Trust in artificial intelligence for medical diagnoses

14

Georgiana Juravle[a,*], Andriana Boudouraki[b], Miglena Terziyska[c], Constantin Rezlescu[c]

[a]*Faculty of Psychology and Educational Sciences, Alexandru Ioan Cuza University, Iasi, Romania*
[b]*School of Computer Science, University of Nottingham, Nottingham, United Kingdom*
[c]*Department of Experimental Psychology, University College London, London, United Kingdom*
**Corresponding author: Tel.: +40-232201297, e-mail address: georgiana.juravle@uaic.ro*

Abstract

We present two online experiments investigating trust in artificial intelligence (AI) as a primary and secondary medical diagnosis tool and one experiment testing two methods to increase trust in AI. Participants in Experiment 1 read hypothetical scenarios of low and high-risk diseases, followed by two sequential diagnoses, and estimated their trust in the medical findings. In three between-participants groups, the first and second diagnoses were given by: human and AI, AI and human, and human and human doctors, respectively. In Experiment 2 we examined if people expected higher standards of performance from AI than human doctors, in order to trust AI treatment recommendations. In Experiment 3 we investigated the possibility to increase trust in AI diagnoses by: (i) informing our participants that the AI outperforms the human doctor, and (ii) nudging them to prefer AI diagnoses in a choice between AI and human doctors. Results indicate overall lower trust in AI, as well as for diagnoses of high-risk diseases. Participants trusted AI doctors less than humans for first diagnoses, and they were also less likely to trust a second opinion from an AI doctor for high risk diseases. Surprisingly, results highlight that people have comparable standards of performance for AI and human doctors and that trust in AI does not increase when people are told the AI outperforms the human doctor. Importantly, we find that the gap in trust between AI and human diagnoses is eliminated when people are nudged to select AI in a free-choice paradigm between human and AI diagnoses, with trust for AI diagnoses significantly increased when participants could choose their doctor. These findings isolate control over one's medical practitioner as a valid candidate for future trust-related medical diagnosis and highlight a solid potential path to smooth acceptance of AI diagnoses amongst patients.

Keywords

Trust, AI, Healthcare, Medical diagnosis, Medical decision-making

Progress in Brain Research, Volume 253, ISSN 0079-6123, https://doi.org/10.1016/bs.pbr.2020.06.006
© 2020 Elsevier B.V. All rights reserved.

1 Introduction

Having started as a basic tool to help the traditional practicing physician in the early 1950s (Yu et al., 2018), AI has now become a steady partner in the medical decision-making process. The projected growth of the AI health market is expected to reach a staggering 6.6 billion USD by 2021, as distributed across a wide range of health-related applications—from robot assisted surgery and virtual nursing assistants, through administrative workflow assistance, fraud detection, and dosage error reduction, up to key applications devoted to preliminary diagnosis, automated image diagnosis, or cyber security (Accenture, 2017). Healthcare practitioners are already dedicated users of AI with the purpose to ease data analysis, to formulate better diagnoses and treatment predictions, and importantly, to lower the amount of admin work. As such, AI is primarily treasured for *true workflow optimization*—As long as AI is used to streamline workflow and operations, clinicians are able to focus on developing specific tools for better diagnosis and treatment. In this respect, a recent report on healthcare practitioners who use AI describes a two thirds decrease of total time needed to write medical reports, with 79% of practitioners reporting that AI has helped them fight burnout in the workplace, and a significant 75% declaring improvement in predictions and treatment of disease following AI usage (GE Healthcare, 2019).

AI has real benefits for physicians, other healthcare professionals and patients. Physicians are able to dedicate more time to the direct relationship with the patient (De Fauw et al., 2018; Wrzeszczynski et al., 2017) and AI tools have significantly improved diagnosis and treatment prediction, helping to better patients' health. AI-based diagnosis is expected to be common in the near future (Ting et al., 2018), with machine-learning algorithms already used, for example, for various medical decisions such as skin cancer diagnosis (Esteva et al., 2017), more precise identification of body areas in need to receive radiotherapy (Chu et al., 2016), and pulmonary tuberculosis diagnosis (Lakhani and Sundaram, 2017). Such technologies are currently utilized with the help of medical staff. However, commercial companies such as Babylon, Ada, and Your.MD have developed AI systems that are set to provide patients with a diagnosis, without any input needed from a human medical doctor (Heaven, 2018).

Nevertheless, AI for healthcare is challenged by people's skepticism with respect to AI's provable benefit, responsibility attribution in case of an error, as well as the ever-growing concern over individual privacy that the AI brings (GE Healthcare, 2019). Arguably, the existent skepticism boils down to *trust*. Definitions of trust vary across fields (Rousseau et al., 1998), but in the medical domain trust can be defined as the expectation that a medical practitioner/technology will provide beneficial recommendations for a given patient's health, allowing for risks to be taken based on this expectation (Thom et al., 2011). Trust's crucial role in the relationship between patient and medical doctor has been long acknowledged, with numerous studies delineating how important trust is for accepting the doctor's advice and following the treatment plan (Hall et al., 2001; Thom, 2002), as well as for being satisfied with

the service and achieving the desired positive health outcome (Cook et al., 2004; Pearson and Raeke, 2000). Not surprisingly, trust was found to underlie willingness to try automation, making it crucial for the successful implementation of AI for healthcare (Lee and Moray, 1994; Lee and See, 2004). To date, however, there is little research on how patients' trust in a given diagnosis is affected when this is performed by AI instead of the traditional human doctors.

Trust in decisions made by automation technology is partly different from trust in decisions made by people. Whereas trust in humans most often relies on competence, benevolence, and integrity, it has been shown that trust in machines depends more on their perceived functionality and reliability (Mcknight et al., 2011). A consistent preference for humans' opinions over algorithms has been reported, even when the algorithms are known to be superior, a phenomenon called algorithm aversion (Önkal et al., 2009; Promberger and Baron, 2006). Relatedly, people are faster to lose trust in technologies such as AI, when they witness them erring, as compared to those situations where they would be witnessing another person make an error (Alvarado-Valencia and Barrero, 2014; Bisantz and Seong, 2000; Dietvorst et al., 2015; Dzindolet et al., 2003; Muir and Moray, 1996; Parasuraman and Riley, 1997; Promberger and Baron, 2006). Interestingly, people associate more words such as Absolute, Competence, Trustworthy, and Security in caring, with the idea of *trust in medical technology,* as compared to the simple notion of *trust in technology* (Montague et al., 2009). These results to date could be taken to highlight that people expect more from medical technology than from other forms of technology.

With these considerations in mind, the current research had several aims. Given the scarcity of studies on the topic, our first aim was to better understand how patients' trust in their diagnosis is affected when the diagnosis is given by an AI doctor rather than by a human doctor, and whether this trust depends on severity of disease or familiarity with technology (Experiment 1). Our second aim was to investigate if patients require higher standards of performance from AI, as compared to human doctors, in order to trust a diagnosis (Experiment 2). With our third aim we were interested to explore if trust in AI can be increased by construction of preference (i.e., by nudging people to *choose* to be diagnosed by AI rather than being assigned to an AI diagnosis).

2 Experiment 1

In Experiment 1, we were interested how trust in AI compares to trust in a human doctor when AI is used as a primary or secondary diagnosis tool, offering a second (consistent or inconsistent) opinion to that of a human doctor. We also investigated any existent order effects on trust when patients are presented with both AI and human diagnoses.

Additionally, our goal was to assess whether and how *disease risk* impacts trust toward AI diagnoses. We refer to high-risk diseases as those diseases that are life-threatening if left untreated, whereas low-risk diseases signal no such life-threat

(Rolland, 1984). To date, no empirical research explicitly investigated disease severity and diagnosis acceptance, however, it has been found that patients gave their physicians more explicit mandates of trust for more complex diseases, while for simpler medical issues one mandate of trust was enough (Skirbekk et al., 2011). Further, numerous studies on patient-doctor trust have typically examined patients with high-risk diseases such as cancer, or chronic problems such as diabetes, fact which could be taken to suggest that trust appears to be an issue especially for such diseases (Cao et al., 2017; Piette et al., 2005; Ridd et al., 2009). Considering trust in connection to disease severity, we hypothesized a link between higher-risk disease and lower levels of trust.

Finally, we were interested to explore whether *familiarity with technology* can lead to greater trust toward AI. We expect people more familiar with technology to also better understand how AI works, or be more forgiving to the presence of a lack of explainability on its part (see, e.g., Doran et al., 2017; Holzinger et al., 2017). Therefore, we hypothesized that familiarity with technology will result in greater trust for AI diagnoses.

2.1 Method

2.1.1 Participants

182 participants were recruited through social media (Facebook and Instagram), Prolific, and the UCL Psychology participant pool. Six participants were excluded for incomplete answers. The reported analyses were conducted on data from 176 participants (101 female, age range: 18–85 years old).

2.1.2 Design and materials

All participants were presented with eight hypothetical medical scenarios in random order, each followed by two proposed diagnoses (first and second). The scenarios varied on: *disease risk* (high or low), *first diagnosis result* (positive or negative), and *second diagnosis result* (confirming or disconfirming the first diagnosis result); see Table 1. Diseases were classified as high-risk if they could be life threatening when left untreated (Rolland, 1984).

Depending on the source of the first and second diagnosis, participants were allocated to one of three groups: Human-AI, AI-Human, Human-Human. Except for this difference, all participants were presented with the same eight scenarios. Additionally, participants in the first two groups were given a text explaining how AI is used for medical diagnosis, text taken from an example of IBM's doctor Watson, with accuracy rates for AI and Human diagnoses for different types of cancer (Haenssle et al., 2018). See Appendix A Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006 for full texts of medical scenarios and AI description.

Familiarity with technology was measured with the Media and Technology Usage and Attitudes Scale (Rosen et al., 2013). This scale includes 35 questions about how often people perform certain actions with technology, such as *Check your personal e-mail* or *Check the news on a mobile phone* (see Appendix B in

Table 1 Eight medical scenarios were used for the 2x2x2 within-participants variables in Experiment 1.

Scenario	Medical problem	Risk	First diagnosis	Second diagnosis
1	Viral infection	Low	Positive	Agreement (Positive)
2	Tooth abscess	Low	Positive	Disagreement (Negative)
3	Sepsis	High	Positive	Agreement (Positive)
4	Diabetes	High	Positive	Disagreement (Negative)
5	Hormone imbalance	Low	Negative	Agreement (Negative)
6	Sexual disease	Low	Negative	Disagreement (Positive)
7	Heart disease	High	Negative	Agreement (Negative)
8	Lung cancer	High	Negative	Disagreement (Positive)

Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006, for the full scale). The answers were given on an 8-point visual frequency scale, from "Never" to "All the time."

2.1.3 Procedure

The study was conducted online using Qualtrics. Participants were randomly assigned to one of the three between-participants groups and were then presented with the eight scenarios, including the diagnoses. After *each* diagnosis, they were asked to indicate a percentage score (0–100%) for how much they trusted the initial diagnosis. At the end of the experiment, participants filled in the Media and Technology Usage and Attitudes Scales and they were debriefed about the purpose of the study.

2.1.4 Data analysis

For brevity and simplicity, we here report only those analyses for the scenarios including a positive first diagnosis (i.e., indicating the presence of a disease and recommending a treatment). Analyses for the scenarios with negative first diagnoses are detailed in the Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006 Appendix C.

To investigate trust in AI as a primary diagnosis tool, a 2×2 mixed analysis of variance (ANOVA) was conducted on the trust scores (dependent variable) given after the first diagnosis of low- and high-risk diseases (within-participants variable) by the participants in Human-AI and AI-Human groups (between-participants variable). Familiarity with technology was included in the analyses as a covariate.

To investigate trust in AI as a secondary diagnosis tool, two further ANOVAs were conducted on the differences in trust scores (calculated as trust after second diagnosis minus trust after first diagnosis, as the dependent variable), depending on whether the second diagnosis confirmed or disconfirmed the initial diagnosis. Specifically, a confirming second diagnosis was expected to increase trust, whereas a disconfirming one was expected to decrease trust. Each of the 2×2 mixed ANOVAs had disease risk (low vs high) as the within-participants factor. The between-factor was doctor type for the

second diagnosis (Human vs AI) as found in the Human-Human and Human-AI groups respectively. Note that a first Human diagnosis is the starting point in both subject groups.

Lastly, to test for any order effects in the combined Human and AI diagnoses, we conducted a 2×2 mixed ANOVA on the trust scores after the second diagnosis (within-participants variable) received for low- and high-risk diseases by those participants in the Human-AI and AI-Human groups (between-participants variable).

2.2 Results

Averages of trust scores together with standard deviations for all variables manipulated in Experiment 1, by subject group, are presented in the Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006, Appendix C Table S1.

2.2.1 Trust in AI as a primary diagnosis tool

A significant main effect of doctor type was found on the trust ratings data, $F(1,111) = 4.57$, $P = 0.035$, $\eta^2_p = 0.040$, with higher trust in Human ($M = 69.75$, $SD = 15.54$), as compared to AI diagnoses ($M = 63.01$, $SD = 17.66$). Further, a significant main effect of disease risk was evidenced on the ratings data, $F(1,111) = 85.56$, $P < 0.001$, $\eta^2_p = 0.435$, with participants placing higher trust in low-risk ($M = 75.08$, $SD = 17.91$), as compared to high-risk diagnoses ($M = 66.11$, $SD = 16.98$); see Fig. 1. There was no interaction effect between doctor type and disease risk, $F(1,111) < 0.01$, $P = 0.966$, $\eta^2_p = 0.000$.

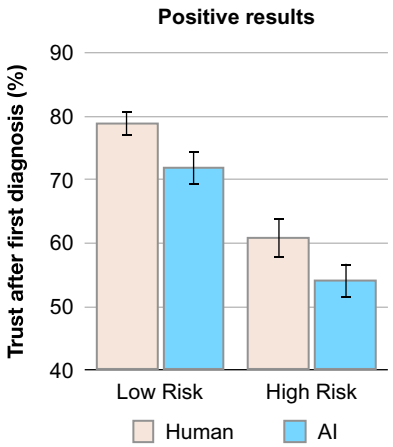

FIG. 1

Mean trust scores (as percentage) in Experiment 1 after the primary diagnosis provided by Human vs AI doctors, for low-risk and high-risk diseases. Error bars represent the standard error of the mean (SEM).

When adding familiarity with technology as a covariate, $F(1,111) = 0.20$, $P = 0.759$, $\eta^2_p = 0.001$, only the main effect of doctor type remained significant, $F(1,111) = 4.52$, $P = 0.036$, $\eta^2_p = 0.039$, but not the main effect of risk, $F(1,111) = 1.68$, $P = 0.198$, $\eta^2_p = 0.015$. Familiarity with technology did not correlate significantly with trust in either low-risk, $r(61) = 0.042$, $P = 0.746$, or high-risk diagnoses made by AI, $r(61) = -0.034$, $P = 0.792$.

2.2.2 Trust in AI as a secondary diagnosis tool

When the second diagnosis confirmed the initial one, we found a significant main effect of disease risk, $F(1,110) = 10.28$, $P = 0.002$, $\eta^2_p = 0.083$, with significantly larger increases in trust ratings for high-risk ($M = 17.72$, $SD = 23.73$), as compared to low-risk diseases ($M = 9.86$, $SD = 11.75$). The main effect of doctor type, $F(1,110) = 3.46$, $P = 0.065$, $\eta^2_p = 0.031$, and the interaction effect between risk and doctor type, $F(1,110) = 3.47$, $P = 0.065$, $\eta^2_p = 0.028$, suggested that trust increased more when the second confirming opinion was provided by a Human doctor ($M = 16.31$, $SD = 11.33$) vs an AI doctor ($M = 11.08$, $SD = 16.78$), with this difference being larger for high-risk diseases. For high-risk diseases, the difference between Human ($M = 21.98$, $SD = 17.98$) and AI ($M = 12.64$, $SD = 28.52$) for a secondary diagnosis was confirmed significant by a follow-up independent t-test, $t(112) = 2.13$, $P = 0.036$, Cohen's $d = 0.40$. For low-risk diseases, the difference between Human ($M = 10.15$, $SD = 9.73$) and AI ($M = 9.52$, $SD = 13.81$) as secondary diagnosis was not significant, $t(110) = 0.28$, $P = 0.778$, Cohen's $d = 0.05$.

When the second diagnosis disconfirmed the initial one, neither the main effect of disease risk, $F(1,112) = 0.07$, $P = 0.791$, $\eta^2_p = 0.001$, nor the main effect of doctor type, $F(1,110) = 0.10$, $P = 0.921$, $\eta^2_p = 0.000$, were significant. A significant interaction was nevertheless found between disease risk and doctor type, $F(1,110) = 4.39$, $P = 0.038$, $\eta^2_p = 0.038$. The average impact of a second disconfirming Human diagnosis was $M = -25.74$ ($SD = 26.74$) for low-risk, and $M = -19.73$ ($SD = 25.66$) for high-risk diseases, $t(61) = 1.86$, $P = 0.068$, Cohen's $d = 0.24$. The average impact of a second disconfirming AI diagnosis was $M = -19.96$ ($SD = 25.89$) for low-risk and $M = -24.63$ ($SD = 30.44$) for high-risk diseases, $t(51) = 1.16$, $P = 0.250$, Cohen's $d = 0.16$. See Fig. 2 for a depiction of trust in AI as a secondary diagnosis.

2.2.3 Order effects

Relevant to our research question, neither the main effect of order, $F(1,110) = 1.74$, $P = 0.190$, $\eta^2_p = 0.016$, nor the interaction effect of group by risk, $F(1,110) = 0.57$, $P = 0.451$, $\eta^2_p = 0.003$, were significant. Solely the main effect of disease risk proved significant, $F(1,110) = 98.98$, $P < 0.001$, $\eta^2_p = 0.472$, consistent with our previous results that trust is higher for diagnoses of low-risk ($M = 11.08$, $SD = 16.78$), as compared to high-risk diseases ($M = 11.08$, $SD = 16.78$).

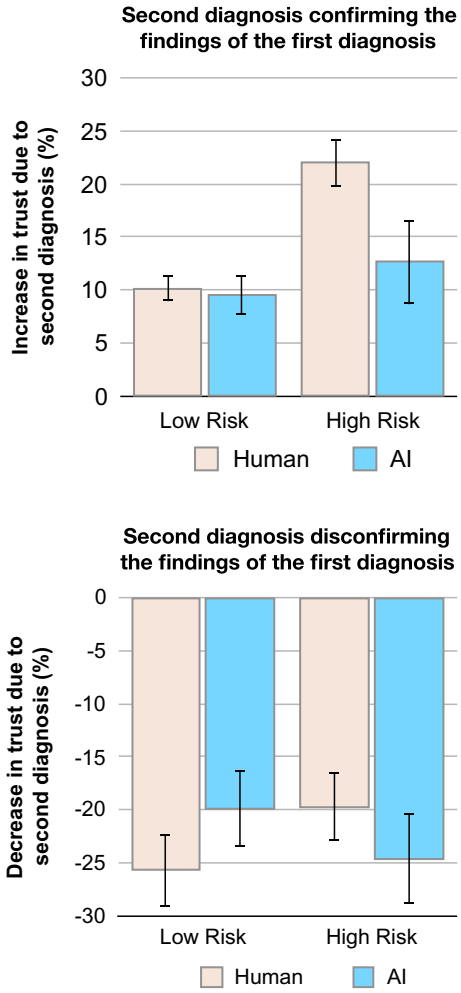

FIG. 2

Change in trust (as percentage) from a first diagnosis given by a human to second diagnosis coming from either a Human or AI doctor for both low-risk and high-risk medical conditions, with the second diagnosis confirming the first positive diagnosis (upper panel), and the second diagnosis disconfirming the first positive diagnosis (lower panel). Error bars represent SEM.

2.3 Discussion

The main goal of Experiment 1 was to investigate trust for AI as a diagnosis tool. When looking at AI *as a primary diagnosis*, the results of Experiment 1 highlighted that participants trusted AI doctors less than human doctors in order to confirm a medical condition and recommend treatment. Such results are in line with the traditional findings on algorithm aversion pointing out that we are more likely to trust a

human over an algorithm (Bisantz and Seong, 2000; Dietvorst et al., 2015; Dzindolet et al., 2003; Parasuraman and Riley, 1997). This finding signals that people may have a built-in preference for interacting with other humans over interacting with other types of agents, such as AI machines. Considering that throughout life we interact most often with other humans than with AI, trusting the human medical practitioner over the AI agent may simply reflect a learned preference (e.g., the affective judgments literature, see Zajonc, 1980). Note that AI technology has only recently started to become mainstream (i.e., in the form of apps we use on mobile phones and operating systems like Apple's Siri). Moreover, it is also a lot less likely for people to already have consistently encountered AI in the healthcare. Therefore, lack of exposure to and interaction with AI could be taken as a solid start to explain the distrust in AI, as found in Experiment 1. In line with this explanation, it has been shown that familiarity with algorithms can increase acceptance (Kramer et al., 2018). Social norms may also play a role: Informing participants that an algorithm is already being used by a large part of the population can significantly boost acceptance (Alexander et al., 2018). However, in our Experiment 1 we found that familiarity with technology could not explain the lower trust in AI.

AI technology could be used as a *secondary diagnosis tool*, to confirm or disconfirm an initial diagnosis provided by a human doctor. To examine how much people trust AI for this purpose, we measured its positive/negative impact on people's trust on the initial human diagnosis and compared it with the positive/negative impact a second human diagnosis would have had. The increases in trust following a second confirming diagnosis were significantly larger for high-risk diseases. In particular, we found that this boost in trust comes from Human confirming diagnoses, which have almost double the impact of AI confirming diagnoses; no similar differences in impact were observed for Human and AI disconfirming second diagnoses.

Lastly, the finding that the order of two given diagnoses, with either human doctor or AI algorithm first, does not impact trust ratings, may suggest that implementing AI as a medical diagnosis tool could potentially be accommodated at any stage in the medical healthcare process.

To summarize, Experiment 1 provided evidence that people tend to trust Human diagnoses more than AI diagnoses. One explanation for this result may be that people expect higher standards of AI diagnoses, in order to be trusted as much as human diagnoses (e.g., see Montague et al., 2009). When AI and human doctors are known or implied to be equally good at diagnosing illnesses, one possibility is that people will still place more trust in human doctors. In Experiment 2 we were interested to test this possibility by addressing directly participants' expectations with regard to the accuracy of medical diagnoses by Human and AI doctors.

3 Experiment 2

To measure expectations relative to AI as opposed to human medical practitioners, we presented participants with scenarios including diagnoses by AI and human doctors and asked them to indicate how confident they would *require the doctors* to be in

their diagnoses, such that the participants themselves will accept the diagnosis and follow through with the recommended treatment.

3.1 Methods

3.1.1 Participants

44 participants were recruited for this experiment through the Testable Minds platform (minds.testable.org), where the experiment took place. Of those, three were excluded for providing incomplete or unusual answers indicating that they misunderstood the questions, such that the final analysis included 41 participants (18 male, age range: 20–73 years old, $M = 38.90$, $SD = 12.86$).

3.1.2 Design and materials

We had two within-participant variables: *doctor type* (Human vs AI) and *disease risk* (Low vs High). Each participant responded to four imagined medical scenarios covering each combination of the variables. The low-risk scenarios were *bacterial infection* and *tooth abscess,* and the high-risk scenarios were *lung cancer* and *sepsis*. Scenarios were counterbalanced, such that each scenario featured an equal number of times with a Human doctor and an AI doctor. Each scenario described a series of symptoms, followed by a doctor's diagnosis (always positive, i.e., illness confirmed) and a recommended treatment. Participants were asked to indicate the required level of confidence from the doctor for them to accept the given diagnosis and follow through with the recommended treatment. Confidence ratings provided the dependent variable. All materials are available as Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006; see Appendix D.

3.1.3 Procedure

After they gave their informed consent to participate and answered demographics questions (i.e., age, gender, ethnicity, and level of education), participants were presented with the instructions and a text explaining that in some scenarios their doctor would be an AI rather than a Human, followed by a short explanation of what an AI doctor is (i.e., a sophisticated algorithm that can make medical diagnoses). Participants were then presented with the four diagnosis scenarios, in a randomized sequential order, and had to indicate after each one how confident they required the doctor to be in order for them to follow through with the recommended treatment. A demo of the experiment can be accessed at this link: testable.org/t/29215de1b6.

3.2 Data analysis and results

A 2×2 repeated measures ANOVA revealed a significant main effect of disease risk, $F(1,40) = 16.03$, $P < 0.001$, $\eta^2_p = 0.27$, with participants expecting higher confidence for high-risk ($M = 86.02$, $SD = 12.64$) than low-risk diseases ($M = 77.20$, $SD = 13.73$); see Fig. 3. The main effect of doctor type and the interaction effect were not significant ($Ps > 0.139$).

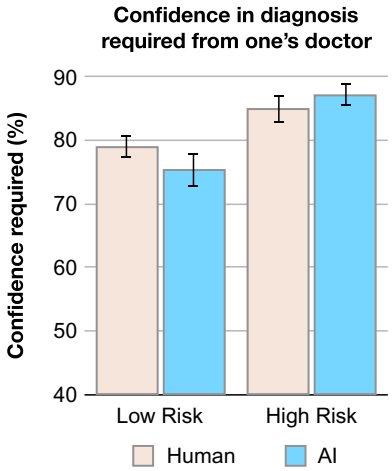

FIG. 3

Confidence levels (as percentage) that participants would require from a medical practitioner (AI or human doctor) in order to follow the recommended treatment for low-risk and high-risk diseases in Experiment 2. Error bars represent SEM.

3.3 Discussion

The results of Experiment 2 indicate that while participants expected their doctors to have higher confidence in the diagnoses of high-risk diseases (i.e., cancer or sepsis) as compared to low-risk diseases (i.e., bacterial infection or tooth abscess) in order to follow through with the recommended treatment, they did so irrespectively of whether the diagnosis was given by a Human or AI. Considering that people seem to have similar performance standards for Human and AI doctors in order to trust them, it would be interesting to learn if information that AI is the better tool will increase their trust in AI diagnoses. We will test this hypothesis in Experiment 3.

Another way to increase trust in AI may be through construction of preference toward AI (Lichtenstein and Slovic, 2006). Specifically, we tested the situation of simply nudging participants toward choosing (rather than being assigned to) AI over Human diagnoses. We hypothesized that involvement in the diagnosis will result in enhanced trust especially for those situations where trust is relatively low, such as for AI diagnoses. A strong relationship between trust and choice has already been described (Kao et al., 1998), with involvement acting to empower patients and helping them to trust their doctor more (Say et al., 2006).

4 Experiment 3

In Experiment 3 we asked whether it is possible *to increase trust in AI* by: (i) informing our participants that AI doctors are better at diagnosing a certain medical condition of interest, or (ii) nudging them to select an AI doctor over a Human

doctor in a free choice setup. We therefore manipulated the information given to our participants regarding what type of doctor is more accurate (Human or AI) and participants' control over choosing their doctor type. We expect that trust in AI when learning that AI provides the more accurate diagnosis will still be lower than trust in Humans when learning that Human doctors provide the more accurate diagnosis. However, we expect that trust in AI diagnoses will increase significantly when participants are nudged to select (rather than being assigned to) the AI doctor, and that this effect does not affect Human doctors' diagnoses.

4.1 Method

4.1.1 Participants

197 participants were recruited from the Testable Minds platform (minds.testable.org). Five participants with very short response times to any of the questions, suggesting lack of involvement with the task, were excluded (response threshold: $RT_{z\text{-}score} > 3$; Pukelsheim, 1994). We also excluded 5 participants in the Human and 25 participants in the AI groups who did not submit to our nudges and did not select the expected doctor (i.e., the doctor presented as more accurate; see Section 4.1.3). Therefore, our final analyses included 162 participants (82 male, 78 female, 2 other; age range: 18–74, $M = 35.30$ years old, $SD = 12.29$), with between 38 and 43 participants in each group.

4.1.2 Design and materials

We had a 2x2x2 mixed design, with *disease risk* (low- vs high-risk) as a within-subjects factor, and *doctor type* (Human vs AI) and decision maker over doctor type (patient vs authority) as between-participants factors. The scenarios included the description of symptoms, a medical diagnosis, and a recommended treatment for bacterial infection and lung cancer, respectively. All materials are available as Supplementary materials in the online version at https://doi.org/10.1016/bs.pbr.2020.06.006; see Appendix E.

4.1.3 Procedure

Participants were informed that they will be presented with two medical scenarios and that the existing medical system can use either Human or AI doctors to diagnose their problems, followed by a brief description of what an AI doctor is. The scenarios were presented in a random order. Participants were asked to imagine that these happened to them and estimate the probability (0–100%) that they would start the recommended treatment, given that it was not possible to get a second opinion. Depending on the between-participants condition to which they were randomly distributed, participants were then told that either the Human or AI doctor was more accurate and thus preferred by people, and that they were randomly assigned to this more accurate Human/AI doctor (authority as decision maker condition), or they can rather select which doctor they prefer (participant as decision maker condition). Following this manipulation, most participants in the Participant-selected conditions chose the more

accurate doctor, but a few still chose the other one. We excluded the data from those participants. A demo of the experiment is available at this link: testable.org/t/292189c9a4

4.2 Data analysis and results

We found a significant main effect of doctor type, $F(1,158)=14.32$, $P<0.001$, $\eta^2_p=0.083$, with the probability to follow the treatment significantly lower when recommended by the AI doctor ($M=69.66$, $SD=23.17$), as compared to a human doctor recommendation ($M=81.69$, $SD=18.95$). A significant main effect of disease risk was also found, $F(1,158)=11.28$, $P<0.001$, $\eta^2_p=0.067$, with a higher probability to follow treatment in the low-risk condition ($M=80.17$, $SD=25.12$), than in the high-risk condition ($M=71.19$, $SD=30.70$). Most importantly, we found a significant main effect of decision maker, $F(1,158)=4.40$, $P=0.038$, $\eta^2_p=0.027$, and a significant interaction effect between decision maker and doctor type, $F(1,158)=6.40$, $P=0.012$, $\eta^2_p=0.039$. Overall, there was a higher probability to follow treatment in the Participant-selected condition ($M=78.81$, $SD=18.20$), than in the Authority-selected condition ($M=72.62$, $SD=24.80$), and this difference was significantly higher for AI diagnoses ($M=76.88$, $SD=18.23$ vs $M=61.88$, $SD=25.53$), as compared to Human diagnoses ($M=80.95$, $SD=18.17$ vs $M=82.35$, $SD=19.79$); see Fig. 4. The other interaction effects failed to reach

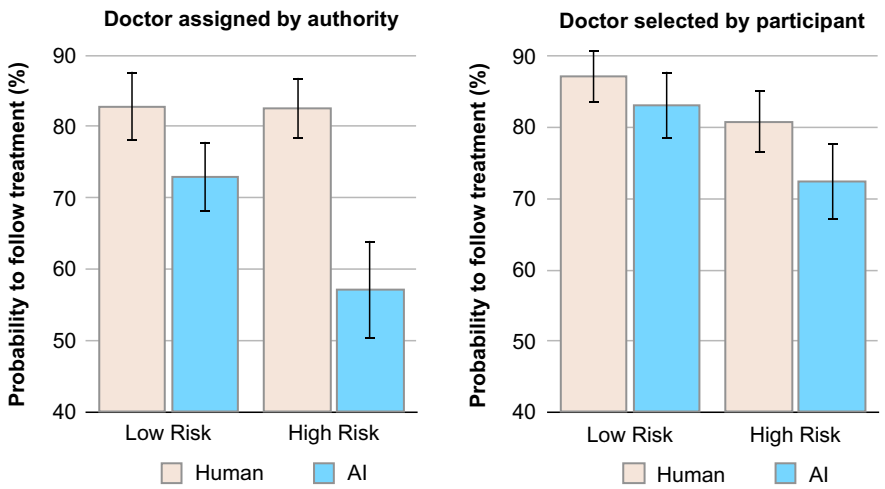

FIG. 4

Probability to follow treatment recommended by the Human and AI doctors, for low- and high-risk medical conditions. Left panel presents probability to follow treatment when the doctor was assigned by authority, whereas the right panel presents probability to follow treatment when participants selected their doctor. Error bars represent SEM.

significance ($Ps > 0.27$), with the interaction effect between disease risk and doctor type marginally approaching significance, $F(1,158) = 3.81$, $P = 0.053$, $\eta^2_p = 0.024$.

To better understand the effects of decision maker on Human and AI diagnoses, we conducted two more ANOVAs. For participants in the Human conditions only, a 2×2 mixed ANOVA with factors disease risk (low vs high) and decision maker (authority vs participant) revealed no significant main effects or interactions (all $Ps > 0.409$). For participants in the AI conditions only, we found a significant main effect of risk, $F(1,79) = 12.58$, $P < 0.001$, $\eta^2_p = 0.137$, with participants more likely to follow the treatment in the low-risk condition ($M = 76.88$, $SD = 24.60$), as compared to the high-risk condition ($M = 62.44$, $SD = 33.68$). Most importantly, a significant main effect of decision maker was also found, $F(1,79) = 9.36$, $P = 0.035$, $\eta^2_p = 0.106$, with participants more likely to follow treatment when nudged to select the AI doctor ($M = 76.88$, $SD = 18.23$), as opposed to those situations when they were only assigned to it ($M = 61.88$, $SD = 25.53$). The interaction effect between decision maker and disease risk, $F(1,79) = 0.07$, $P = 0.799$; $\eta^2_p = 0.137$, was not significant.

To examine whether information about the more accurate doctor can eliminate the difference in trust between Human and AI, we performed a further 2×2 mixed ANOVA on the data from the Authority conditions only. Results confirmed the significant main effect of doctor type, $F(1,80) = 16.62$, $P < 0.001$, $\eta^2_p = 0.172$, suggesting participants place significantly more trust in Human as opposed to AI diagnoses, when they were informed their assigned doctor was the more accurate one. There was also an interaction effect between doctor type and disease risk, $F(1,80) = 4.50$, $P = 0.037$, $\eta^2_p = 0.053$, indicating a larger gap between trust in Human and AI diagnoses for the high-risk disease ($M = 82.91$, $SD = 22.77$, and $M = 54.13$, $SD = 37.08$, respectively), than for the low-risk disease ($M = 81.79$, $SD = 26.94$, and $M = 69.94$, $SD = 27.27$, respectively). The main effect of disease risk did not reach significance, $F(1,80) = 3.37$, $P = 0.070$, $\eta^2_p = 0.040$.

4.3 Discussion

Consistent with our findings from Experiment 1, we found that participants were less likely to follow recommended treatments for high-risk diseases, as compared to low-risk diseases, and from an AI, rather than from a human doctor. This preference for Human over AI diagnoses was not eliminated when participants were informed that the doctor to which they were assigned (Human or AI) was the most accurate one.

However, we found evidence that giving participants the option to choose the doctor type while heavily nudging them to opt for an AI diagnosis dramatically increased trust in AI, when chosen. This effect was specific to AI diagnoses and was evident for both low-risk and high-risk diseases. The same manipulation did not have an effect on human diagnoses.

Giving participants the choice of medical practitioner in a situation where they did not feel particularly trusting and as such, making them feel more in control, might have alleviated their hesitations about the AI. Involvement may act to give the opportunity to those patients who have not yet established trust to feel more empowered

and through that, trust the diagnosis more (Say et al., 2006). This result is potentially exciting, as it provides a path to address the inclusion of AI in traditional healthcare plans, as well as patients' accepting as a standard participant to the healthcare process.

4.4 General discussion

In the present study we investigated how much trust people have in using AI as a potential diagnosis tool in the medical healthcare field with the goal to find ways to increase this trust.

Lower trust in AI algorithms compared to human diagnoses was found in Experiment 1 and further confirmed in Experiment 3. These findings are in line with previous literature, highlighting algorithm aversion (Alvarado-Valencia and Barrero, 2014; Bisantz and Seong, 2000; Dietvorst et al., 2015; Dzindolet et al., 2003; Muir and Moray, 1996; Önkal et al., 2009; Parasuraman and Riley, 1997; Promberger and Baron, 2006). Initial trust in diagnoses for life-threatening conditions was generally lower, as compared to low-risk conditions, but the gap in trust between diagnoses made by human and AI was similar across conditions.

The lower trust in AI algorithms as compared to their human counterpart was evident for first diagnoses. Additionally, our participants trusted more a second opinion coming from a Human than from an AI agent to confirm the initial (human) diagnosis of a life-threatening disease. This result led us to believe that people may have more demanding criteria for trusting AI medical technology, as compared to trusting human doctors, a hypothesis tackled by Experiment 2.

Surprisingly, against our prediction, Experiment 2 found that people had comparable standards of expertise for AI and human doctors, as measured by the required level of confidence from their AI/human doctors about the diagnosis. We acknowledge that the word *confidence* may have different meanings when referring to human and AI doctors. When it comes to a human doctor's confidence in their diagnosis, one may assume that it refers to a mere estimation based on the doctor's gut feeling, whereas when an AI reports its confidence, a better guess would be that it is derived from calculations based on concrete data. As such, a confidence report from an AI may be more meaningful to people, than one from a human.

Because algorithm appreciation has been demonstrated for certain concrete tasks in which algorithms are expected to outperform humans (Logg et al., 2019), in our last investigation in Experiment 3 we attempted to boost trust in AI by explicitly informing participants about the AI superiority. However, we found that trust in AI did not increase when people were told that the AI outperformed the human doctor. Another manipulation seemed way more successful: Participants exhibited higher trust in AI when they were nudged to select the AI for a diagnosis, as opposed to being automatically assigned to an AI diagnosis. The probability to follow the recommended treatment increased significantly when participants were given a choice and this effect was specific to the AI diagnoses.

There is considerable evidence in psychology that choices influence preferences. Numerous studies using free-choice paradigms have demonstrated that participants increased their preference for the chosen option in a wide range of situations (Coppin et al., 2010; Gerard and White, 1983; Sharot et al., 2009; Shultz et al., 1999). Surprisingly, even people who were *led to believe* that they had chosen a non-preferred option came to prefer the new option, the opposite of their chosen alternative (e.g., Hall et al., 2010; Johansson et al., 2014). Preference increases for the selected choice are typically explained by our need to reduce cognitive dissonance (Festinger, 1957, 1962). Choosing between alternatives creates cognitive dissonance because the non-chosen option has desirable aspects that we now must ignore. The way to reduce dissonance is to increase our evaluation of the selected choice and devalue the ignored option (Shultz et al., 1999). Through the same mechanism, people who are generally less trusting of AI but nevertheless are successfully nudged to choose it over a human doctor will look for ways to justify this decision to themselves and increase their preference for AI in the process.

We highlight two potential limitations of the current study: First, our experiments used hypothetical medical scenarios and evaluated participants' trust as they had to only *imagine* having the described disease. There is always the question of how well our findings would replicate in real-life situations. Second, it is possible that our results highlighting that choice increased trust in AI were driven (at least partly) by specifically excluding participants whose trust in AI was so low that nudging failed to make a difference, and by consequently selecting only those who already had higher trust in AI. However, this is unlikely to be the case, because we did not record the same results for Human diagnoses (although admittedly there were fewer participants excluded in the Human condition).

In terms of practical implications, we consider our results to be exciting. They suggest a potential working path toward increased acceptance and inclusion of AI within the traditional healthcare system. Follow up research could explore other behavioral nudges to influence patients to choose AI for medical diagnoses and examine how these nudges differ in terms of effectiveness. It has been shown that patients are traditionally keen to be involved in decision-making tasks (e.g., choosing one doctor over another), although they expect the doctor to solve their specific medical problem that required treatment in the first place (Deber, 1994; Deber et al., 1996). This reduced interest in how doctors make diagnoses and decide on treatments is good news for AI, considering its complexity and that the computations leading to a certain outcome are often impenetrable to humans (e.g., the black box problem, see Castelvecchi, 2016).

5 Conclusions

We find that people have lower trust in AI diagnoses, as compared to human diagnoses. However, this trust gap can be almost eliminated if we move away from enforcing AI diagnoses to the *libertarian paternalism* proposed by Thaler and Sunstein

(2003). That is, if we preserve patients' freedom of choice, while at the same time steering them toward the better option (for individuals and/or the society). Giving people a choice between human and AI diagnoses while heavily nudging them to choose the AI increases their trust in AI diagnoses to levels typically found for human diagnoses.

References

Accenture, 2017. Artificial intelligence: healthcare's new nervous system—accenture report. In: Accenture Report, 1–8. Retrieved from https://www.accenture.com/us-en/insight-artificial-intelligence-future-growth.

Alexander, V., Blinder, C., Zak, P.J., 2018. Why trust an algorithm? Performance, cognition, and neurophysiology. Comput. Hum. Behav. 89, 279–288. https://doi.org/10.1016/j.chb.2018.07.026.

Alvarado-Valencia, J.A., Barrero, L.H., 2014. Reliance, trust and heuristics in judgmental forecasting. Comput. Hum. Behav. 36, 102–113. https://doi.org/10.1016/j.chb.2014.03.047.

Bisantz, A.M., Seong, Y., 2000. Assessment of operator trust in and utilization of automated decision aids under different framing conditions. In: Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Association, "Ergonomics for the New Millennium", 28, 5–8. https://doi.org/10.1016/S0169-8141(01)00015-4.

Cao, W., Qi, X., Yao, T., Han, X., Feng, X., 2017. How doctors communicate the initial diagnosis of cancer matters: cancer disclosure and its relationship with patients' hope and trust. Psychooncology 26 (5), 640–648. https://doi.org/10.1002/pon.4063.

Castelvecchi, D., 2016. Can we open the black box of AI? Nature 538 (7623), 20–23. https://doi.org/10.1038/538020a.

Chu, C., De Fauw, J., Tomasev, N., Paredes, B.R., Hughes, C., Ledsam, J., … Cornebise, J., 2016. Applying machine learning to automated segmentation of head and neck tumour volumes and organs at risk on radiotherapy planning CT and MRI scans [version 1; referees: 1 approved with reservations]. F1000Research 5, 2104. https://doi.org/10.12688/F1000RESEARCH.9525.1.

Cook, K.S., Kramer, R.M., Thom, D.H., Stepanikova, I., Mollborn, S.B., Cooper, R.M., 2004. Trust and distrust in patient-physician relationships: perceived determinants of high- and low-trust relationships in managed-care settings. In: Trust and Distrust in Organization: Dilemmas and Approaches. Russell Sage Foundation, pp. 65–98.

Coppin, G., Delplanque, S., Cayeux, I., Porcherot, C., Sander, D., 2010. I'm no longer torn after choice: how explicit choices implicitly shape preferences of odors. Psychol. Sci. 21 (4), 489–493. https://doi.org/10.1177/0956797610364115.

De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., … Ronneberger, O., 2018. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat. Med. 24 (9), 1342–1350. https://doi.org/10.1038/s41591-018-0107-6.

Deber, R.B., 1994. Physicians in health care management: 8. The patient-physician partnership: decision making, problem solving and the desire to participate. CMAJ 151 (4), 423–427.

Deber, R.B., Kraetschmer, N., Irvine, E.J., 1996. What role do patients wish to play in treatment decision making? Arch. Intern. Med. 156 (13), 1414–1420. https://doi.org/10.1001/archinte.156.13.1414.

Dietvorst, B.J., Simmons, J.P., Massey, C., 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. J. Exp. Psychol. Gen. 144 (1), 114–126. https://doi.org/10.1037/xge0000033.

Doran, D., Schulz, S., Besold, T.R., 2017. What does explainable AI really mean? In: A New Conceptualization of Perspectives. Retrieved from http://arxiv.org/abs/1710.00794.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P., 2003. The role of trust in automation reliance. Int. J. Hum. Comput. Stud. 58 (6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542 (7639), 115–118. https://doi.org/10.1038/nature21056.

Festinger, L., 1957. An introduction to the theory of dissonance. In: A Theory of Cognitive Dissonance. Stanford University Press. https://doi.org/10.1037/10318-001.

Festinger, L., 1962. Cognitive dissonance. Sci. Am. 207, 93–102. https://doi.org/10.7312/columbia/9780231175081.003.0017.

GE Healthcare, 2019. The AI effect. How artificial intelligence is making Healthcare more human. In: MIT Technology Review, (October), pp. 1–38.

Gerard, H.B., White, G.L., 1983. Post-decisional reevaluation of choice alternatives. Pers. Soc. Psychol. Bull. 9 (3), 365–369. https://doi.org/10.1177/0146167283093006.

Haenssle, H.A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., … Zalaudek, I., 2018. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann. Oncol. 29 (8), 1836–1842. https://doi.org/10.1093/annonc/mdy166.

Hall, M.A., Dugan, E., Zheng, B., Mishra, A.K., 2001. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? Milbank Q. 79 (4), 613–639. https://doi.org/10.1111/1468-0009.00223.

Hall, L., Johansson, P., Tärning, B., Sikström, S., Deutgen, T., 2010. Magic at the marketplace: choice blindness for the taste of jam and the smell of tea. Cognition 117 (1), 54–61. https://doi.org/10.1016/j.cognition.2010.06.010.

Heaven, D., 2018. Your next doctor's appointment might be with an AI. In: MIT Technology Review. Retrieved from https://www.technologyreview.com/s/612267/your-next-doctors-appointment-might-be-with-an-ai/.

Holzinger, A., Biemann, C., Pattichis, C.S., Kell, D.B., 2017. What do we need to build explainable AI systems for the medical domain? arxiv:1712.09923[cs, AI] Retrieved fromhttp://arxiv.org/abs/1712.09923.

Johansson, P., Hall, L., Tärning, B., Sikström, S., Chater, N., 2014. Choice blindness and preference change: you will like this paper better if you (believe you) chose to read it!. J. Behav. Decis. Mak. 27 (3), 281–289.

Kao, A.C., Green, D.C., Davis, N.A., Koplan, J.P., Cleary, P.D., 1998. Patients' trust in their physicians: effects of choice, continuity, and payment method. J. Gen. Intern. Med. 13 (10), 681–686. https://doi.org/10.1046/j.1525-1497.1998.00204.x.

Kramer, M.F., Schaich Borg, J., Conitzer, V., Sinnott-Armstrong, W., 2018. When do people want AI to make decisions? In: AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, pp. 204–209. https://doi.org/10.1145/3278721.3278752.

Lakhani, P., Sundaram, B., 2017. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284 (2), 574–582. https://doi.org/10.1148/radiol.2017162326.

Lee, J.D., Moray, N., 1994. Trust, self-confidence, and operators' adaptation to automation. Int. J. Hum. Comput. Stud. 40 (1), 153–184. https://doi.org/10.1006/ijhc.1994.1007.

Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46, 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Lichtenstein, S., Slovic, P. (Eds.), 2006. The Construction of Preference. Cambridge University Press, New York, NY.

Logg, J.M., Minson, J.A., Moore, D.A., 2019. Algorithm appreciation: people prefer algorithmic to human judgment. Organ. Behav. Hum. Decis. Process. 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005.

Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F., 2011. Trust in a specific technology: an investigation of its components and measures. ACM Trans. Manag. Inf. Syst. 2 (2), 1–15. https://doi.org/10.1145/1985347.1985353.

Montague, E., Kleiner, B.M., Winchester, W.W., 2009. Empirically understanding trust in medical technology. Int. J. Ind. Ergon. 39 (4), 628–634. https://doi.org/10.1016/j.ergon.2009.01.004.

Muir, B.M., Moray, N., 1996. Trust in automation. Part ii. Experimental studies of trust and human intervention in a process control simulation. Ergonomics 39 (3), 429–460. https://doi.org/10.1080/00140139608964474.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., Pollock, A., 2009. The relative influence of advice from human experts and statistical methods on forecast adjustments. J. Behav. Decis. Mak. 22 (4), 390–409. https://doi.org/10.1002/bdm.637.

Parasuraman, R., Riley, V., 1997. Humans and automation: use, misuse, disuse, abuse. Hum. Factors 39 (2), 230–253. https://doi.org/10.1518/001872097778543886.

Pearson, S.D., Raeke, L.H., 2000. Patients' trust in physicians: many theories, few measures, and little data. J. Gen. Intern. Med. 15 (7), 509–513. https://doi.org/10.1046/j.1525-1497.2000.11002.x.

Piette, J.D., Heisler, M., Krein, S., Kerr, E.A., 2005. The role of patient-physician trust in moderating medication nonadherence due to cost pressures. Arch. Intern. Med. 165 (15), 1749–1755. https://doi.org/10.1001/archinte.165.15.1749.

Promberger, M., Baron, J., 2006. Do patients trust computers? J. Behav. Decis. Mak. 19 (5), 455–468. https://doi.org/10.1002/bdm.542.

Pukelsheim, F., 1994. The three sigma rule. Am. Stat. 48, 88–91.

Ridd, M., Shaw, A., Lewis, G., Salisbury, C., 2009. The patient-doctor relationship: a synthesis of the qualitative literature on patients' perspectives. Br. J. Gen. Pract. 59 (561), 268–275. https://doi.org/10.3399/bjgp09X420248.

Rolland, J.S., 1984. Toward a psychosocial typology of chronic and life-threatening illness. Fam. Syst. Med. 2 (3), 245–262. https://doi.org/10.1037/h0091663.

Rosen, L.D., Whaling, K., Carrier, L.M., Cheever, N.A., Rokkum, J., 2013. The media and technology usage and attitudes scale: an empirical investigation. Comput. Hum. Behav. 29 (6), 2501–2511. https://doi.org/10.1016/j.chb.2013.06.006.

Rousseau, D.M., Sitkin, S.B., Burt, R.S., Camerer, C., 1998. Not so different after all: a cross-discipline view of trust. Acad. Manage. Rev. 23 (3), 393–404. https://doi.org/10.5465/AMR.1998.926617.

Say, R., Murtagh, M., Thomson, R., 2006. Patients' preference for involvement in medical decision making: a narrative review. Patient Educ. Couns. 60 (2), 102–114. https://doi.org/10.1016/j.pec.2005.02.003.

Sharot, T., Martino De, B., Dolan, R.J., 2009. How choice reveals and shapes expected hedonic outcome. J. Neurosci. 29 (12), 3760–3765. https://doi.org/10.1523/JNEUROSCI.4972-08.2009.

Shultz, T.R., Léveillé, E., Lepper, M.R., 1999. Free choice and cognitive dissonance revisited: Choosing "lesser evils" versus "greater goods". Pers. Soc. Psychol. Bull. 25 (1), 40–48. https://doi.org/10.1177/0146167299025001004.

Skirbekk, H., Middelthon, A.L., Hjortdahl, P., Finset, A., 2011. Mandates of trust in the doctor-patient relationship. Qual. Health Res. 21 (9), 1182–1190. https://doi.org/10.1177/1049732311405685.

Thaler, R.H., Sunstein, C.R., 2003. Libertarian paternalism. Am. Econ. Rev. 93 (2), 175–179. https://doi.org/10.1257/000282803321947001.

Thom, D.H., 2002. Patient trust in the physician: relationship to patient requests. Fam. Pract. 19 (5), 476–483. https://doi.org/10.1093/fampra/19.5.476.

Thom, D.H., Wong, S.T., Guzman, D., Wu, A., Penko, J., Miaskowski, C., Kushel, M., 2011. Physician trust in the patient: development and validation of a new measure. Ann. Fam. Med. 9 (2), 148–154. https://doi.org/10.1370/afm.1224.

Ting, D.S.W., Liu, Y., Burlina, P., Xu, X., Bressler, N.M., Wong, T.Y., 2018. AI for medical imaging goes deep. Nat. Med. 24 (5), 539–540. https://doi.org/10.1038/s41591-018-0029-3.

Wrzeszczynski, K.O., Frank, M.O., Koyama, T., Rhrissorrakrai, K., Robine, N., Utro, F., … Darnell, R.B., 2017. Comparing sequencing assays and human-machine analyses in actionable genomics for glioblastoma. Neurol. Genet. 3 (4), 1–8. https://doi.org/10.1212/NXG.0000000000000164.

Yu, K.H., Beam, A.L., Kohane, I.S., 2018. Artificial intelligence in healthcare. Nat. Biomed. Eng. 2, 719–731. https://doi.org/10.1038/s41551-018-0305-z.

Zajonc, R.B., 1980. Feeling and thinking: preferences need no inferences. Am. Psychol. 35 (2), 151–175.