Do People Favor Artificial Intelligence Over Physicians? A Survey Among the General Population and Their View on Artificial Intelligence in Medicine

Yakar, Derya; Ongena, Yfke P.; Kwee, Thomas C.; Haan, Marieke

Published in:
Value in Health

DOI:
10.1016/j.jval.2021.09.004

ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Themed Section: Artificial Intelligence

Do People Favor Artificial Intelligence Over Physicians? A Survey Among the General Population and Their View on Artificial Intelligence in Medicine

Derya Yakar, MD, PhD,* Yfke P. Ongena, PhD,* Thomas C. Kwee, MD, PhD, Marieke Haan, PhD

ABSTRACT

Objectives: To investigate the general population's view on artificial intelligence (AI) in medicine with specific emphasis on 3 areas that have experienced major progress in AI research in the past few years, namely radiology, robotic surgery, and dermatology.

Methods: For this prospective study, the April 2020 Online Longitudinal Internet Studies for the Social Sciences Panel Wave was used. Of the 3117 Longitudinal Internet Studies For The Social Sciences panel members contacted, 2411 completed the full questionnaire (77.4% response rate), after combining data from earlier waves, the final sample size was 1909. A total of 3 scales focusing on trust in the implementation of AI in radiology, robotic surgery, and dermatology were used. Repeated-measures analysis of variance and multivariate analysis of variance was used for comparison.

Results: The overall means show that respondents have slightly more trust in AI in dermatology than in radiology and surgery. The means show that higher educated males, employed or student, of Western background, and those not admitted to a hospital in the past 12 months have more trust in AI. The trust in AI in radiology, robotic surgery, and dermatology is positively associated with belief in the efficiency of AI and these specific domains were negatively associated with distrust and accountability in AI in general.

Conclusions: The general population is more distrustful of AI in medicine unlike the overall optimistic views posed in the media. The level of trust is dependent on what medical area is subject to scrutiny. Certain demographic characteristics and individuals with a generally positive view on AI and its efficiency are significantly associated with higher levels of trust in AI.

Keywords: artificial intelligence, general population, medicine, surveys, and questionnaires.

VALUE HEALTH. 2022; 25(3):374–381

Introduction

Artificial intelligence (AI), which refers to a wide variety of computer-executed tasks that simulate human intelligence, will improve and reshape the future of healthcare tremendously.[1-4] AI in healthcare, which includes mostly the fields of machine learning (the use of computer algorithms to perform specific tasks) and robotics,[1-4] is rapidly evolving and numerous applications have shown their potential value. For example, recent machine learning studies have shown to either equal or even outperform radiologists in the diagnosis of breast cancer on screening mammography, and dermatologists' performance in the detection of skin cancer.[5-8] Parallel to this, many examples of effective robotics-assisted surgery and newer techniques with autonomous robotic systems are underway.[9-11] Up until now, the technical development of AI systems has been at the center of attention. At present, there is very little experience with the general population's attitude and potential consequences of introducing these systems into the practice of patient care.[12,13] Ethical and legal issues are just as important as the technical performance of these systems for responsible and successful implementation. Among ethical priorities, human consent is one of the cornerstones of the patient-physician relationship for all investigations and treatments. Therefore, when vetting the proper context and defining the confines in which these AI systems should act, the consent of the public is essential. Involving the public will set practical conditions on how to put these new promising technologies into effect. Moreover, this will help predict how people will accept new technology, which provides a feedback loop to developers, thereby increasing the participation of the population.[14]

Previous studies on the acceptance of the use of AI in medicine were limited to specific specialty areas (such as radiology,[15,16] dermatology,[17] and robotics[18]) or were hampered by a low number of participants involved (varying between 20 and 264).[15–19] Given the fast-paced, new, and upcoming technologies in the

*Derya Yakar and Yfke P. Ongena contributed equally to this work.

entire field of medicine, there is a need for a larger study in the broader field of medicine. Outcomes of an analysis performed by the New York times on views expressed about AI involving the last 30 years show that discussions have been consistently more optimistic.[20] Though there was hope for the beneficial impact of AI on healthcare, it was not without specific concerns (eg, on the loss of control and ethical worries).[20] We, therefore, hypothesize that both excerpts of optimism and pessimism will be represented among the public, with an optimistic view being more dominant as time has passed. The purpose of this study was to investigate the general population's view on AI in medicine, with specific emphasis on 3 areas that have experienced major progress in AI research in the past few years, namely radiology, robotics, and dermatology.[5-11]

Methods

Study Design and Subjects

For this study, we used the April 2020 Online Longitudinal Internet Studies for the Social sciences (LISS)-panel wave. The LISS panel is a nationally representative household panel study for people aged 16 years and older in The Netherlands (under Dutch law, the minimum required age for treatment consent is 17 years[21]) (see Table 1 for demographics). To establish the Longitudinal Internet Studies for the Social Sciences Panel, a traditional random sample was drawn from the population registers in collaboration with Statistics Netherlands.[22] Based on the 2019 key figures of Statistics Netherlands, the Longitudinal Internet Studies for the Social Sciences Panel distribution features (Table 1) are comparable with the distribution in the Dutch general population.[23] Another large research institute, the Netherlands Institute for Health Services, uses routinely recorded data from Dutch healthcare providers to evaluate the quality and effectiveness of healthcare. As shown by the Netherlands Institute for Health Services, 78.2% of Dutch people who are registered with a general practitioners (GPs) practice has had contact with a GP at least once in 2019,[24] which falls within the range of 72.18% (SD 14.47) (see Table 1) of the sample used in this study. In separate data collection rounds in this panel (ie, waves), different questions were asked of the same pool of respondents. We combined the April 2020 wave with an earlier wave including healthcare-related characteristics (eg, contact with medical professionals and hospitalizations). An informed

Table 1. Demographic characteristics of the sample.

Variable	n (%)	Mean (SD)
Age		55.07 (17.64)
Gender		
Male	908 (47.56)	
Female	1001 (52.44)	
Level of education		
Low (elementary school)	451 (23.62)	
High school or lower vocational	657 (34.42)	
College (BA, MA, Msc, MD, or PhD)	733 (38.40)	
Other	68 (3.56)	
Main occupation		
Employed	891 (46.67)	
Unemployed, looking for work	39 (2.04)	
Unemployed, not looking for work	89 (4.66)	
Student	113 (5.92)	
Housekeeping	170 (8.91)	
Retired	551 (28.86)	
Doing voluntary work	40 (2.10)	
Unknown	16 (0.84)	
Immigration background		
Dutch	1549 (81.14)	
Western immigration background	166 (8.7)	
Non-Western immigration background	139 (7.28)	
Unknown	55 (2.88)	
Consulting general practitioner		
0 times in past months	531 (27.82)	0
1 or more times in the past months	1378 (72.18)	2.52 (14.47)
Consulting specialist		
0 times in past months	113 (59.51)	
1 or more times in the past months	773 (40.49)	
Admitted to hospital		
0 times in past months	1728 (90.52)	
1 or more times in the past months	174 (9.11)	
Unknown	7 (0.4)	
Underwent surgery past 12 months	113 (5.92)	
Days in hospital (based on n = 174 admitted to hospital)		5.13 (19.90)
Days in hospital (based on n = 1909)		0.47 (6.17)

Table 2. Scales trust in different domains.

Trust in radiology	Trust in:
qv20a080 Even if computers are better at evaluating scans, I still prefer a doctor	Taking over the diagnostic interpretation of tasks
qv20a081 I think radiology is not ready for implementing artificial intelligence in evaluating scans	Taking over the diagnostic interpretation of tasks
qv20a082 It worries me when computers analyze scans without the interference of humans	Clarity about medical procedures
qv20a083 The sooner I get the results, even when this is from a computer, the more I am at ease	Patient communication
qv20a084 Through human experience a radiologist can detect more than the computer	Accuracy
qv20a085 It is unclear to me how computers will be used in evaluating scans	Clarity about medical procedures
qv20a091 I wonder how it is possible that a computer can give me the results of a scan	Clarity about medical procedures
Trust in surgery	
qv20a140 Even if computers are better in operating patients, I still prefer a doctor	Taking over operative procedures of the surgeon
qv20a141 I think hospitals are not ready for implementing artificial intelligence in operating patients	Taking over operative procedures of the surgeon
qv20a142 It worries me when computers operate patients without the interference of humans	Clarity about medical procedures
qv20a143 Through human experience a surgeon can detect more than the computer	Accuracy
qv20a144 It is unclear to me how computers will be used in conducting operations	Clarity about medical procedures
Trust in dermatology	
qv20a164 Even if computers are better at evaluating spots on the skin, I still prefer a doctor	Trust in AI in taking over diagnostic interpretation tasks of the dermatologist
qv20a165 I think dermatology is not ready for implementing artificial intelligence in evaluating spots on the skin	Trust in AI in taking over diagnostic interpretation tasks of the dermatologist
qv20a166 It worries me when computers analyze spots on the skin without the interference of humans	Clarity about medical procedures
qv20a167 The sooner I get the results, even when this is from a computer, the more I am at ease	Trust in AI in taking over diagnostic interpretation tasks of the dermatologist concerning clarity about medical procedures and patient communication
qv20a168 Through human experience a dermatologist can detect more than the computer	Accuracy
qv20a169 It is unclear to me how computers will be used in evaluating spots on the skin	Clarity about medical procedures

AI indicates artificial intelligence.

consent procedure was used ensuring double consent by means of a reply card and an internet login (see Scherpenzeel et al[25] also for more methodological details). Ethical approval for the procedures in the LISS panel was given by the board of overseers (https://www.lissdata.nl/organization/board-overseers). All data are available at https://www.dataarchive.lissdata.nl.

Measurement of Attitude Scales

All attitude questions were 5-point agree-disagree scales, though experimental manipulation of labels for the scales was included, with half of the respondents answering on an agree-disagree scale, and the other half answering on a construct-specific scale. In all analyses, it was verified that these manipulations did not affect any of the outcomes (see Appendix 1 Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.004). From an existing scale on general trust in the implementation of AI in radiology,[15] we developed 3 new scales focusing on trust in the implementation of AI in radiology, robotic surgery, and dermatology. In each of these domains a specific task was included in the items for respondents to evaluate (ie, in radiology, "evaluating a scan"; in surgery, "operating patients"; and in dermatology, "evaluating my skin"), whereas for the items focusing on implementation in general medicine, only general terms ("medical tasks") were used. For an overview of all items and subcategories within the scale, see Table 2. Clarity about medical procedures was previously described by Haan et al.[16] as a need of participants finding it important to understand how AI would be used precisely during their visit to the radiology department (eg, the relation of AI to the caregiver, the scanning procedure itself, receiving results) and further adapted to this study (these scales and all data are available at https://lissdata.nl).

Measurement Predictor Variables

The level of education, immigration background, healthcare utilization (visits to GPs, medical specialists, and hospitalizations), and medical area (radiology, robotic surgery, and dermatology) was investigated as potential predictor variables. To measure the level of education, we used the LISS-panel item of highest earned

Table 3. Multivariate tests on 3 subscales of trust in AI (all significant at *P*<.001).

Effect	Λ	F	η^2	df1	df2	Radiology, F-value	Surgery, F-value	Dermatology, F-value
Age	0.93	45.84*	0.003	3	1733	111.38*	22.34*	86.55*
Gender	0.93	46.46*	0.001	3	1733	8103.84*	75.09*	88.86*
Education	0.92	25.76*	0.013	6	3468	70.11*	37.61*	36.05*
Immigration background	0.98	6.13*	0.006	6	3468	4.59†	12.05*	13.56*
Immigration background*gender	0.99	2.01†	0.003	6	3468	0.38	0.70	0.05†
Main occupation	0.96	2.02‡	0.007	18	5205	3.27‡	0.80	3.53‡
Consulting GP	0.99	1.69	0.002	3	1733	0.75	3.60§	3.85†
Consulting Specialist	0.99	1.96	0.002	3	1733	5.12†	0.97	0.15
Admitted to hospital	0.99	5.01‡	0.004	3	1733	1.64	6.56†	13.79*
Days in hospital	0.99	4.02‡	0.006	3	1733	0.52	4.08†	1.73
General attitude AI	0.75	187.97*	0.002	3	1733	376.66*	381.74*	331.57*
Distrust and accountability	0.66	291.78*	0.164	3	1733	641.31*	556.08*	476.85*
Personal Interaction	0.95	28.18*	0.032	3	1733	57.05*	63.81*	25.81*
Efficiency	0.74	144.13*	0.200	3	1733	244.23*	207.51*	364.31*

AI indicates artificial intelligence.
**P*<.001 (2-tailed).
†*P*<.050 (2-tailed).
‡*P*<.010 (2-tailed).
§*P*<.100 (2-tailed).

degree and categories were taken from the Dutch educational system (easiest for respondents to understand), which were converted into international categories: lower education (ie, primary education or lower vocational education), high school (preuniversity education or mediate vocational education), college (university or higher vocational education), and other (no degree, or degree not included among response options). Immigration background was asked in terms of the country of birth of the respondent and both parents. First- and second-generation immigrants were combined, and countries were recoded into Western and non-Western countries. Immigrants from Western countries included Europe (Turkey excluded), North America, Oceania (including Australia and New Zealand), Japan, and Indonesia—the latter was included because immigrants were mainly from former Dutch colonies. Non-Western immigrants in The Netherlands consisted mostly of those from Turkey, Morocco, Surinam, and the Dutch Antilles.

The use of healthcare was defined as the yearly number of visits to a GP or a medical specialist, whether the respondent was admitted to a hospital, and if so, whether he or she underwent surgery, and the number of days spent in a hospital, all within a period of over the past 12 months.

The general attitude toward AI was measured using items developed from a scale by selecting 3 previously validated factors,[15] namely, distrust and accountability (distrust in AI in taking over tasks of doctors concerning patient communication and confidentiality, Cronbach's α 0.73; M 2.9; SD 0.47; a higher score means less trust in AI), personal interaction (preference of personal interaction over AI-based communication, α 0.85; M 4.3; SD 0.60; a higher score means finding personal interaction more important), efficiency (belief of whether AI will improve diagnostic workflow, Cronbach's α 0.74; M 3.2; SD 0.57; a higher score means finding AI more efficient), and a scale measuring the attitude toward AI in medicine (Cronbach's α 0.87; M 3.6; SD 0.78; a higher score means a more positive attitude toward AI). The results of these general attitude scales were also used partially in a previously published study with a different purpose namely, assessing AI in mammpgraphy screening and women's preferences (in the April 2020 wave).[26]

Statistical Analysis

Two different types of analysis of variance were used (for a general explanation of the analysis of variance [ANOVA] and F-values, see Altman et al[27]). First, a repeated-measures ANOVA was used to compare the scores of the attitude scales on AI in radiology, surgery, and dermatology, and to account for the fact that these concepts were measured in a specific order (first radiology, then surgery, and finally, dermatology). Second, a multivariate ANOVA was used, taking the 3 attitude scales as dependent variables to allow comparison of the 3 scales and differences in different predictor variables in 1 analysis (see Altman et al[27] for a more detailed explanation of these tests). All statistical analyses were conducted in R version 4.0.3 (https://www.r-project.org/).

Results

Out of the 3117 LISS panel members contacted, 2411 completed the full questionnaire (77.4% response rate). In the analysis, data were combined with data from an earlier wave that involved measurement of relevant predictor variables, such as the use of healthcare. Because 502 respondents in the last data collection did not participate in that earlier wave, this combination of waves reduced the final sample size to 1909 respondents. The scales on trust in the implementation of AI were reliable (based on the Cronbach's α of 0.75 in radiology, Cronbach's α of 0.76 in robotic surgery, and Cronbach's α of 0.79 in dermatology).

Multivariate Analysis

A repeated-measures ANOVA showed a significant effect of the specialty area—that is, radiology, surgery, or dermatology (F [2, 4832] 162.2; *P*<.010)—on trust in AI. Mauchly's test was significant

Table 4. Pearson correlations between the 3 subscales of trust in AI, general scales and age.

Measure	Trust AI in radiology	Trust AI in surgery	Trust AI in dermatology
General attitude AI	0.397*	0.403*	0.384*
Distrust and accountability	−0.610*	−0.594*	−0.556*
Personal Interaction	−0.286*	−0.278*	−0.229*
Efficiency	0.531*	0.523*	0.564*
Age	−0.177*	−0.083*	−0.159*
Consulting GP (#times)	−0.03	−0.04	−0.05
Consulting Specialist (#times)	−0.07	−0.05	−0.05
Days in hospital	−0.03	−0.03	0.02

# indicates number; AI, artificial intelligence; GP, general practitioner.
*$P<.001$.

(W 0.99; $P<.001$), and, therefore, we used Greenhouse-Geisser and Huynh-Feldt corrections (corresponding corrective coefficients ε 0.995 and ε 0.996, respectively).

In a multivariate ANOVA main effects and interaction, effects were analyzed. All main effects, as well as an interaction effect between immigration background and gender (here Wilk's lambdas), were significant (Table 3). Bartlett's tests, investigating the equality of variance/covariance matrices of the different groups analyzed, were found as not significant for all 3 measures (ie, trust of AI in radiology, robotic surgery, and dermatology), indicating that the groups analyzed have roughly equal variances. Effect sizes were largest for belief in the efficiency of AI ($\eta2$ 0.200; 95% confidence interval 0.167-0.230) and distrust and accountability ($\eta2$ 0.164; 95% confidence interval 0.115-0.174). Given that the effects for the AI in medicine attitude scales, age, number of consultations with a GP and/or medical specialist, and number of

Table 5. Means and SDs of scores of trust in AI in 3 different medical areas for general sample and per gender, education, and immigration background.

Variable	n	Mean (SD) trust AI in radiology	Mean (SD) trust AI in surgery	Mean (SD) trust AI in dermatology
Overall	1909	2.82 (0.66)	2.75 (0.72)	2.90 (0.73)
Males	908	2.92 (0.66)	2.85 (0.69)	3.02 (0.71)
Females	1001	2.73 (0.66)	2.65 (0.73)	2.80 (0.74)
Low (elementary school)	451	2.58 (0.61)	2.56 (0.71)	2.70 (0.72)
High school or lower vocational	657	2.81 (0.66)	2.75 (0.73)	2.89 (0.72)
College (BA, MA, Msc, MD or PhD)	733	3.01 (0.65)	2.90 (0.67)	3.08 (0.70)
Unknown	68	2.56 (0.63)	2.44 (0.79)	2.57 (0.76)
Immigration background				
Dutch	1549	2.82 (0.66)	2.76 (0.71)	2.91 (0.74)
Western immigration background	166	2.85 (0.66)	2.70 (0.70)	2.93 (0.70)
Non-Western immigration background	139	2.78 (0.61)	2.59 (0.79)	2.75 (0.76)
Unknown	55	2.97 (0.58)	2.86 (0.70)	3.06 (0.67)
Main occupation				
Employed	891	2.92 (0.64)	2.82 (0.69)	3.02 (0.69)
Unemployed, looking for work	39	2.88 (0.63)	2.83 (0.70)	2.97 (0.71)
Unemployed, not looking for work	89	2.78 (0.66)	2.72 (0.71)	2.87 (0.74)
Student	113	3.06 (0.58)	2.83 (0.70)	3.08 (0.71)
Housekeeping	170	2.57 (0.66)	2.57 (0.76)	2.68 (0.77)
Retired	551	2.71 (0.66)	2.68 (0.72)	2.75 (0.76)
Doing voluntary work	40	2.64 (0.65)	2.64 (0.66)	2.74 (0.72)
Unknown	16	2.58 (0.80)	2.56 (0.84)	3.25 (0.82)
Admitted to hospital				
0 times in past months	174	2.84 (0.68)	2.81 (0.69)	3.01 (0.71)
1 or more times in past months	1728	2.82 (0.66)	2.74 (0.71)	2.90 (0.74)
Unknown	7	2.69 (0.78)	2.66 (0.93)	2.66 (0.83)

Note. Mean numbers represent a 5-point agree-disagreement scale.
AI indicates artificial intelligence.

Table 6. Means and SDs of scores on 3 subscales of trust in AI for gender and immigration background.

Immigration background	Mean trust AI in radiology		Mean trust AI in surgery		Mean trust AI in dermatology	
	Male	Female	Male	Female	Male	Female
Dutch	2.92 (0.67)	2.73 (0.64)	2.86 (0.69)	2.67 (0.71)	3.01 (0.73)	2.82 (0.73)
Western	2.98 (0.64)	2.75 (0.66)	2.78 (0.65)	2.64 (0.73)	3.03 (0.61)	2.85 (0.75)
Non-Western	2.86 (0.61)	2.67 (0.60)	2.77 (0.82)	2.39 (0.69)	2.96 (0.69)	2.51 (0.76)
Unknown	3.08 (0.62)	2.81 (0.49)	2.97 (0.71)	2.69 (0.68)	3.25 (0.66)	2.77 (0.57)

Note. Mean numbers represent a 5-point agree-disagreement scale.
AI indicates artificial intelligence.

hospitalization days are measured as numeric variables, they can best be interpreted from correlations (Table 4). The positive correlations for the general attitude in AI and efficiency show that respondents who have a positive view of AI in medicine, and those who find it more efficient have more trust in the implementation of AI in radiology, surgery, and dermatology. Correlations for distrust and accountability, personal interaction, and age are negative, which means that respondents who distrust AI in medicine, who find personal interaction important, and are older, have less trust in the implementation of AI in radiology, surgery, and dermatology.

The overall means (Table 5) show that respondents have slightly more trust in AI in dermatology (M 2.90; SD 0.73) than in radiology (M 2.82; SD 0.66) and especially surgery (M 2.75; SD 0.73). Nevertheless, notably, all means are quite close to the middle point of the scale (ie, neither agree nor disagree).

The means for gender show that males, higher educated persons, those who are employed or students, respondents with Western immigration or Dutch background, and those who were not admitted to a hospital in the past 12 months, have more trust in AI than females, lower educated persons, and those with a non-Western immigration background. The significant interaction effect between gender and immigration background shows that the trust among females with a non-Western immigration background is particularly low (Table 6), although this effect was only significant for the trust of AI in dermatology.

An exploratory analysis aiming to explain the associations between utilization of healthcare and trust in AI showed that the age of respondents was weakly positively associated with consulting a GP (Pearson correlation 0.06; $P<.050$), consulting a specialist (Pearson correlation 0.12; $P<.010$), and the number of days admitted in a hospital (Pearson correlation 0.06; $P<.01$). Education was not associated with consulting a GP (F [2, 1831] 1.22; $P=.295$), consultation of medical specialists (F [2, 1831] 1.94; $P=.144$), and the number of days admitted in a hospital (F [2, 1831] 1.32; $P=.266$). Gender was not associated with consulting a GP (t [986.85] −0.39; $P=.701$), and the number of days admitted in a hospital (t [1101.4] 0.54; $P=.590$), but a significant difference was found for consultation of medical specialists on an annual basis, with females consulting more often (M 1.38; SD 2.88) than males (M 0.96; SD 2.08; t [1809.1] −3.65; $P<.010$) (see Appendix 2 Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.004).

Discussion

In this Dutch national online survey study performed among 1909 participants, we found that the general population's view on AI in medicine is leaning more toward a higher level of distrust. This is opposite from what we hypothesized based on publications in the mainstream media. We also found that the level of trust may be dependent on what medical area is subject to scrutiny, and that demographic characteristic and a generally positive view on AI and its efficiency in daily life and society (without specifically considering medicine) are significantly associated with higher levels of trust in AI in medicine.

Given the fact that no previous study has been performed in such a large group of participants, and that AI applications in several different medical areas were investigated at once, comparing our results with other studies is somewhat challenging. For instance, in a study that is comparable in terms of the general focus of AI involving 229 patients in Germany, it was found that patients favored physicians over AI in most clinical settings except when basing treatment decisions on the most current scientific evidence.[19] In contrast, a study[18] among 264 visitors of the Minnesota State Fair (Minnesota), most participants expressed confidence in AI providing medical diagnosis (with a considerable proportion putting more trust in AI than the doctor), which seems to contradict the findings of the German and the present study. This is likely because the study sample of Stai et al[18] consisted largely (70%) of higher educated people (bachelor's degree or higher), whereas, in our study and the German study by Lennartz et al,[19] the percentage of higher educated people was 38.4% and 36.7%, respectively. In both studies, a substantially lower number (and within the range of average percentages [25.9%-39.9%]) of higher educated people of a total population in other Western countries[28] were used. Further supporting this explanation is the fact that being higher educated was a predictor for higher levels of trust in the present study. Furthermore, Stai et al[18] also reported that most respondents were uncomfortable with automated robotic surgery (which matches our findings), but also that most respondents mistakenly believed that partially autonomous surgery was already happening. This emphasizes the need for more patient education and informing them about procedural knowledge about the technique itself, which has also emerged as a specific desire of patients in a previously published qualitative study.[16] Future studies with a narrower focus are necessary to define which topics qualify and should be prioritized for targeted education. In another recent study by Nelson et al[17] in 48 patients who presented at the Brigham and Women's Hospital and melanoma clinics at the Dana-Farber Cancer Institute (Massachusetts), the authors concluded that patients seem to be receptive to the use of AI for skin cancer screening if the integrity of the human physician-patient relationship is preserved. The percentage of higher educated participants, however, was also high in that study (77%).[17] Therefore, similar to the study by Stai et al,[18] Nelson et al[17] overestimated the trust of the average citizen in AI and medicine because of a low number of participants and overrepresentation of higher educated participants. Both of these previous studies highlight the necessity of having a representative and large sample with a more balanced composition of higher and lower educated, Western and non-Western immigrants, older and younger participants, that allows making valid conclusions.

Despite the fact of growing positive attention for AI in medicine in the media, the public does not share the same opinion. This means that we should not assume that journalists or reporters automatically reflect the view of the average citizen and that positive media attention should not be mistaken for public consent. Another relevant finding of this study was that people think differently of AI in radiology and dermatology compared with robotic surgery, with the latter being more distrusted than the former 2. An explanation for this could be that image-based specialties such as radiology and dermatology are considered to be less invasive and have less direct implications in case of fallibility. Practically, this would be best interpreted as that each medical area should be investigated on its own for the implementation of AI in healthcare. Even though all means for the level of trust for radiology, robotic surgery, and dermatology were close to the middle in the 5-point scale for level of trust, it should not be assumed that the results of surveys performed in different medical specialty areas can simply be translated to each other.

Furthermore, being a male, higher educated, employed, and coming from a Dutch or Western background were all found to be associated with a higher level of trust in AI, as opposed to being a female, lower educated, and coming from a non-Western immigration background. This knowledge is important to avoid healthcare inequalities between specific demographic groups that might be created by the introduction of AI technologies in healthcare. Patient education should take the differences between these groups into account. In addition, persons who have a positive general attitude to AI or find it more efficient have more trust in the implementation of AI in radiology, robotic surgery, and dermatology. Respondents who distrust AI in medicine (ie, at the general level) who find personal interaction important and are older have less trust in the specific implementation of AI in radiology, surgery, and dermatology. If we want the population to be more willing and accepting of newer technologies such as AI in healthcare, these aforementioned findings underline the importance of informing people about how these techniques work. Nevertheless, this notion remains speculative because it also assumes that a positive general attitude and being convinced of the higher efficiency of AI are automatically associated with a deeper and correct knowledge of AI in healthcare. This requires further research. Finally, utilizing less healthcare was found to be associated with a higher level of trust in AI. In our data, this could not be explained by associations between age, education level, and use of healthcare. Nevertheless, there was a significant association of gender and utilization of consulting a medical specialist, with females consulting more often than males. Thus, an explanation could be that men utilize less healthcare, and, as a consequence, are less affected by changes in healthcare and, therefore, also have an overall higher level of trust in AI in medicine.

In summary, the results of this study should not be interpreted as a barrier to translate AI-based technologies into clinical practice. Instead, it is the beginning of a shared decision between the physician and patient, which starts with a conversation about a person's preferences and thoughts.[29]

This study was limited because it was performed in the Dutch population and results may not be generalizable to every other population worldwide. In The Netherlands, all inhabitants have maximum access to healthcare because of a national compulsory insurance system. In countries or populations with less access to healthcare, one might be more open to the use of AI in healthcare as this may increase the availability of healthcare services. Furthermore, other variables such as education, immigration, occupational status, might have different impacts in other countries or cultures. Another study limitation is that the view of healthcare professionals (which also plays an important role when considering the implementation of AI technologies in clinical practice) could not be investigated because healthcare professionals constituted a very small and heterogeneous minority in the national household panel that was used. Interestingly, in a previous study done among radiologists with 1041 respondents from 54 countries,[30] it was shown that this group fears being replaced by AI and that this was associated with limited knowledge of AI. This indicates that there is also a need to educate healthcare professionals. Furthermore, there are many more applications of AI in healthcare than on disease detection in medical imaging, dermatology, and robotics in surgery (eg, prognostication/risk management, image processing, healthcare operations or management, natural language processing, etc[4]), and people may have a different view on other medical applications that were not specifically addressed in the present study. For example, AI is likely to have a disruptive impact on the risk management of patients across healthcare providers. This could potentially have a considerable impact on the trust levels between patient and doctor. It can be hypothesized that the deployment of such an AI-based risk management tool (both at an individual and population level) can pose considerable distrust in the general population.

Conclusions

Unlike overall optimistic views posed in the media about AI in medicine, the general population is more distrustful of AI in medicine. The level of trust is dependent on what medical area is subject to scrutiny. Demographic characteristics and a generally positive view of AI and its efficiency are significantly associated with higher levels of trust in AI.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.09.004.

Article and Author Information

Accepted for Publication: September 6, 2021

Published Online: October 12, 2021

doi: https://doi.org/10.1016/j.jval.2021.09.004

Author Affiliations: Department of Radiology, Medical Imaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands (Yakar, Kwee); Center of Language and Cognition, University of Groningen, Groningen, The Netherlands (Ongena); Department of Sociology, University of Groningen, Groningen, The Netherlands (Haan).

Correspondence: Derya Yakar, MD, PhD, Department of Radiology, Medical Imaging Center, University Medical Center Groningen, University of Groningen, Hanzeplein 1, PO Box 30.001, Groningen, The Netherlands 9700 RB. Email: d.yakar@umcg.nl

Author Contributions: *Concept and design:* Yakar, Ongena, Kwee, Haan
Acquisition of data: Yakar, Ongena, Kwee, Haan
Analysis and interpretation of data: Yakar, Ongena, Kwee, Haan
Drafting of the manuscript: Yakar, Ongena, Kwee, Haan
Critical revision of the paper for important intellectual content: Yakar, Ongena, Kwee, Haan
Statistical analysis: Ongena
Provision of study materials or patients: Ongena, Haan

Obtaining funding: Yakar, Ongena, Kwee, Haan
Administrative, technical, or logistic support: Yakar
Supervision: Yakar

Conflict of Interest Disclosures: The authors reported no conflicts of interest.

Funding/Support: This research is part of a project funded by a grant of the Open Data Infrastructure for Social Science and Economic Innovations in The Netherlands.

Role of Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436.
2. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317–1318.
3. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health*. 2018;8(2):020303.
4. Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. *BMC Med Inform Decis Mak*. 2021;21(1):125.
5. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology*. 2019;290(2):305–314.
6. Wu N, Phang J, Park J, et al. Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging*. 2020;39(4):1184–1194.
7. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks [published correction appears in Nature. 2017;546(7660):686]. *Nature*. 2017;542(7639):115–118.
8. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol*. 2018;29(8):1836–1842.
9. Peters BS, Armijo PR, Krause C, Choudhury SA, Oleynikov D. Review of emerging surgical robotic technology. *Surg Endosc*. 2018;32(4):1636–1655.
10. Ficuciello F, Tamburrini G, Arezzo A, Villani L, Sicilian B. Autonomy in surgical robots and its meaningful human control. *Paladyn J Behav Robot*. 2019;10(1):30–43.
11. Rassweiler JJ, Autorino R, Klein J, et al. Future of robotic surgery in urology. *BJU Int*. 2017;120(6):822–841.
12. Geis JR, Brady AP, Wu CC, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Radiology*. 2019;293(2):436–440.
13. Bhandari M, Zeffiro T, Reddiboina M. Artificial intelligence and robotic surgery: current perspective and future directions. *Curr Opin Urol*. 2020;30(1):48–54.
14. Mathieson K. Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. *Inf Syst Res*. 1991;2(3):173–191.
15. Ongena YP, Haan M, Yakar D, Kwee TC. Patients' views on implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiology*. 2020;30:1033–1040.
16. Haan M, Ongena YP, Hommes S, Kwee TC, Yakar D. A Qualitative study to understand patient perspective on the use of artificial intelligence in radiology. *J Am Coll Radiol*. 2019;16:1416–1419.
17. Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. *JAMA Dermatol*. 2020;156(5):1–12.
18. Stai B, Heller N, McSweeney S, et al. Public perceptions of artificial intelligence and robotics in medicine. *J Endourol*. 2020;34(10):1041–1048.
19. Lennartz S, Dratsch T, Zopfs D, et al. Use and control of artificial intelligence in patients across the medical workflow: single-center questionnaire study of patient perspectives. *J Med Internet Res*. 2021;17(23):e24221.
20. Fast E, Horvitz E. Long-term trends in the public perception of artificial intelligence. Posted online September 16, 2016. arXiv:1609.04904. http://arxiv.org/abs/1609.04904.
21. Civil code book 7: article 447. Overheid.nl. https://wetten.overheid.nl/BWBR0005290/2020-07-01/#Boek7_Titeldeel7_Afdeling5_Artikel447. Accessed July 1, 2021.
22. Sample and recruitment. CentERdata, Institute for Data Collection and Research. https://www.lissdata.nl/about-panel/sample-and-recruitment. Accessed August 18, 2021.
23. Population; key figures. StatLine. https://opendata.cbs.nl/statline/#/CBS/en/dataset/37296eng/table?fromstatweb. Accessed August 18, 2021.
24. Use primary care. Public Health care.info. https://www.volksgezondheidenzorg.info/onderwerp/eerstelijnszorg/cijfers-context/gebruik-eerstelijnszorg#!node-gebruik-huisartsenzorg-naar-leeftijd. Accessed August 18, 2021.
25. Scherpenzeel AC, Das M. "True" longitudinal and probability-based internet panels: evidence from The Netherlands. In: Das M, Ester P, Kaczmirek L, eds. *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. New York, NY: Routledge; 2010:77–104.
26. Ongena YP, Yakar D, Haan M, Kwee TC. Artificial intelligence in screening mammography: a population survey of women's preferences. *J Am Coll Radiol*. 2021;18:79–86.
27. Altman Douglas G, Martin BJ. Statistics Notes: comparing several groups using analysis of variance. *BMJ*. 1996;312(7044):1472.
28. http://appsso.eurostat.ec.europa.eu/nui/show.do. Accessed February 12, 2021.
29. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361–1367.
30. Huisman M, Ranschaert E, Parker W, et al. An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education [published online May 11, 2021]. *Eur Radiol*. https://doi.org/10.1007/s00330-021-07782-4.