RESEARCH PAPER

# Predictors of Decision-Making Confidence in AI-Assisted Brain Tumour Surgery: A Cross-Sectional Analysis of Reliability, Communication, and Uncertainty

**Hua Feng[1], Linda Chiuman[2], Dewi Sartika[3]**

[1,2,3]Universitas Prima Indonesia.

***Corresponding:** Linda Chiuman,
Universitas Prima Indonesia. Email: lindachiuman@unprimdn.ac.id

Received: 13-06-2026, Revised: 24-07-2026, Accepted: 28-07-2026, Published: 10-08-2026

## ABSTRACT

Artificial intelligence (AI) can support imaging interpretation, prognosis, and surgical planning in neuro-oncology, but confidence in AI-assisted decisions depends on more than technical performance. This quantitative-dominant study examined whether perceived AI reliability, communication quality, and decisional uncertainty were associated with decision-making confidence in AI-assisted brain tumour surgery. A cross-sectional survey of 300 stakeholders was analysed using descriptive statistics, Pearson correlations, and simultaneous multiple linear regression. The sample was evenly divided by gender and broadly distributed across five age categories. Mean ratings were above the neutral midpoint for perceived AI reliability ($M = 3.53$), communication quality ($M = 3.53$), decisional uncertainty ($M = 3.65$), and decision-making confidence ($M = 3.69$). All three predictors correlated positively with confidence: perceived AI reliability, $r = .538$; communication quality, $r = .536$; and decisional uncertainty, $r = .625$ (all $ps < .001$). The regression model was significant, $F(3, 296) = 82.013$, $p < .001$, explaining 45.4% of the variance in confidence. Perceived reliability ($B = .238$, beta $= .234$, $p < .001$), communication quality ($B = .139$, beta $= .157$, $p = .008$), and non-reversed decisional uncertainty ($B = .357$, beta $= .390$, $p < .001$) were significant predictors. The positive uncertainty coefficient should not be interpreted as evidence that uncertainty causes confidence. Rather, outcome uncertainty may coexist with confidence in a transparent, values-consistent decision process. The findings support human oversight, explicit risk communication, and careful validation of AI-supported neurosurgical decisions.

**Keywords:** Artificial Intelligence, Brain Tumour Surgery, Communication Quality, Decisional Uncertainty, Decision-Making Confidence

**How to cite this article:** Feng H, Chiuman L, Sartika D. Predictors of decision-making confidence in AI-assisted brain tumour surgery: a cross-sectional analysis of reliability, communication, and uncertainty. Int J Drug Deliv Technol. 2026;16(77s): 123-128. DOI: 10.25258/ijddt.16.77s.13

## INTRODUCTION

The surgery of brain tumours is a high-stakes decision situation where survival, neurological functioning, cognition, communication, independence, and quality of life can tug decisions towards various directions. Modern tumour classification incorporates histological and molecular data, and surgical planning can also involve an evaluation of tumour location, functional networks, anticipated extent of resection and the potential postoperative deficiency (Hervey-Jumper and Berger, 2016; Louis et al., 2021). The current epidemiological data in Indonesia is mostly hospital-oriented, which supports the idea that it is necessary to situate imported technologies and clinical decision-support systems under local service conditions and not expect a direct transfer of performance (Aninditha et al., 2021).

Neuro-oncology has added to the growing presence of AI in image segmentation, molecular classification, prognosis, estimation of risk, and surgical planning. Its popularity has been its ability to bring together big and disparate sets of data with much more regularity as compared to crowdsourcing. Nonetheless predictive accuracy in a development dataset alone does not imply calibration, fairness, external validity, workflow suitability, or clinical benefit. Data quality, institutional accountability, as well as interaction between computational output and professional judgement must all be considered to realize responsible clinical implementation (Kelly et al., 2019; Reddy et al., 2020; Topol, 2019; Wiens et al., 2019).

The perceived AI reliability pertains since individuals will do what they think a system is able to accomplish, and not what its technical attributes claim. Trust-in-automation theory points a difference between proper reliance and under-trust and over-trust. Trust when depending on the technology is not well tuned to its real abilities (Hoff and Bashir, 2015; Lee and See, 2004). There are various indicators of reliability in clinicians and patients regarding health care, such as apparent accuracy, consistency, validation, transparency, institutional endorsement, and checking or override ability of the clinician to the recommendation (Asan et al., 2020; Vayena et al., 2018). Similar findings have been reported in neurosurgical research that proposes that AI can be more readily accepted as an advisory second opinion than as an independent decision-maker (Palmisciano et al., 2020).

The quality of communication is also critical since an algorithm does not transform likelihoods into its own sense-making outcomes. The process of shared decision-making should be clearly explained, offer balanced options, allow questions, focus on values, and clarify the person in charge of the recommendation (Barry and Edgman-Levitan, 2012; Elwyn et al., 2017; Makoul and Clayman, 2006). Surprisingly, risk numbers may generate spurious comfort in showing them without their assumptions or uncertainty, whereas information that is technically correct, but in general matches well with no connection to outcomes of interest to the patient, may not be used (Fagerlin et al., 2011; Street et al., 2009).

The relationship in case of decisional uncertainty is a more complex one. Decisional Conflict model explains that uncertainty concerning a course of action is a result of poor information, values, uncertainty of support, or the inability to make a decision (O'Connor, 1995). It may hence be anticipated to lower confidence. However belief in the process differs conceptually with assurance on the outcome. When brain tumour surgery is performed, one can still be confused about whether they will survive the surgery or if there will be a neurological survival on the end, but at the same time think that the chosen option is informed, ethically justified and is in line with the priorities of the patient. The difference is especially pertinent when AI introduces probabilistic information instead of eradicating uncertainty.

In neurosurgical attitude research history, most have investigated either acceptance or perceived benefits or perceived concerns instead of the modelling of the joint relationship of reliability, communication and uncertainty to a particular decision outcome. The supportive uses of AI are more acceptable to patients and relatives, whereas the complete autonomy is not, and the lack of sufficient qualitative and quantitative data marks issues of errors, bias, false accuracy, loss of hope, and clinical complacency (Palmisciano et al., 2020; Parsons et al., 2026). These results obtain relevance, but are not estimating the individual statistical contribution of each factor. This thus requires a multivariable analysis so that technic truthfulness is tested to stay related with confidence following the communication procedure and acknowledged uncertainty is likewise together later.

The current research investigated the statistical relationship of perceived AI reliability, communication quality, and decisional uncertainty and how these relate to decision-making confidence both together and separately. They were tested in three hypotheses: perceived AI reliability would have a positive relationship with confidence; quality of communication would have a positive relationship with confidence; and, with the no-reversed uncertainty coding that was used originally, more acceptance of decisional uncertainty would be associated with more confidence. The last hypothesis relates to association and coexistence not that uncertainty is better in making decisions.

## METHOD

### Design and Setting

The quantitative survey was a cross-sectional survey as part of an explanatory sequential research programme. The quantitative part aimed to determine how the three predictors contribute to decision-making confidence of stakeholders in AI-assisted brain tumour surgical decisions. Participants were invited using tertiary neurosurgical or neuro-oncology care in Indonesia and were adult patients, family surrogates, or primary caregivers, and neurosurgeons or actively engaged neuro-oncology clinicians. The respondents were allowed to base their answers on first-hand experience, observed utilisation, or were presented with a standardised vignette where AI was able to compare two viable surgical methods and offer advertising estimates of survival and functional risk. The AI was introduced as an aid to decision-making as opposed to supplying its own surgeon.

### Participants

The complete-case sample was a final sample of 300 respondents. The sampling plan tried to get available representation of patients, family surrogates, and neurosurgical clinicians but the final statistical result was not reported of the attained frequency of each stakeholder category. This omission does not allow group-specific interpretation and can be considered a significant limitation. The responses of all interviewees were complete in terms of gender, age, the 16 questionnaire items, and the four construct scores in correlation analysis, and the regression model. The respondents consisted of adults and were supposed to have either recently been involved in a surgical decision concerning a brain tumour or read the standardised decision vignette.

### Measures

There were four items that measured four constructs with the help of the 5-point Likert scale (between 1 (strongly disagree) and 5 (strongly agree). Perceived AI reliability measured apparent accuracy, patient-specific application, and similar cases consistency, and perceived clinical validation. The quality of communication evaluated the proper explanation of the role of AI, the difference between AI output and the opinion of neurosurgeon, the balanced presentation of advantages, dangers and alternatives, and the possibility to ask a question. Decisional uncertainty evaluated uncertainty regarding the preferred choice, perceived incomplete information, problems interpreting information provided by AI, and problems in making a commitment to the choice. All of these items retained their original direction meaning that the higher scores were the more the acknowledged uncertainty. The level of confidence in decision-making measured a sense of knowing, agreement with the patient, confidence in integrating clinical judgement and AI data, and readiness to support or continue with the decision. The construct scores were presented on the original 1-to-5 scale.

**Statistical Analysis**

The analysis was done using the IBM SPSS Statistics, descriptive and inferential. Demographic characteristics were characterized by frequencies and percentages. Means, standard deviations, and percentages of agreements were used in summarising the item distributions and construct-level values were calculated based on the four item means. Pearson correlations were used to test the zero-order correlations between the four constructs. Concurrent multiple linear regression was conducted by combining perceived reliability of AI, communication quality and uncertainty about decisions combined to determine decision making confidence. Statistical significance was determined at p =.05 and unstandardised coefficients, standard errors, standardised coefficients, t statistics, approximate 95% interval and the R squared, adjusted R squared, and the overall F test were provided. Since the study was cross-sectional, the word predictor is done in the statistical sense only and does not mean temporal or causal effects. There are a few verbatim follow-up quotations, which are only used in order to interpret the surprise coefficient of uncertainty; they do not modify estimates of the quantitative effects.

## RESULTS

**Participant Characteristics**

All 300 cases were retained in the analyses. There was male-female balance in terms of 150 men and women. Age was broadly distributed: 16.7% were aged 18 to 29 years, 21.7% were 30 to 39 years, 23.3% were 40 to 49 years, 21.7% were 50 to 59 years, and 16.7% were 60 years or older. The equal representation between genders is a descriptive tool, yet the relationships were not tested on the basis of the age and gender.

**Table 1**

*Demographic Characteristics of the Analytic Sample (N = 300)*

| Characteristic | n | % |
|---|---|---|
| **Gender** | | |
| Female | 150 | 50.0 |
| Male | 150 | 50.0 |
| **Age group** | | |
| 18–29 years | 50 | 16.7 |
| 30–39 years | 65 | 21.7 |
| 40–49 years | 70 | 23.3 |
| 50–59 years | 65 | 21.7 |
| 60 years or older | 50 | 16.7 |

*Note.* Percentages total 100% within each demographic characteristic. Stakeholder-group frequencies were not available in the aggregate output.

**Descriptive Findings**

Construct mean scores were all above 3.00 which was the neutral midpoint. The greatest mean was decision-making confidence (M = 3.69) and the second one is decisional uncertainty (M = 3.65). The perceived quality of AI was 3.53 and the perceived communication was 3.53. The reliability items implied conditional instead of unqualified trust: perceived patient specific relevance was the weakest item (M = 3.29), and perceived clinical validation was the strongest (M = 3.72). The opportunity to ask questions was rated most positively (M = 3.63) whereas clarity of the role of AI and full presentation of risks and alternatives was rated less strongly. The close proximity of the items of uncertainty (M range = 3.61 to 3.71) revealed that there was uncertainty on the informational, interpretive, choice, and commitment dimensions.

The coexistence of comparatively high confidence and high uncertainty was the most important descriptive pattern. To illustrate an instance, 72.0% responded yes or strongly yes that they had been sufficiently informed on the surgical options and 66.7% further responded that there was unimportant information that they still did not know. The responses do not have to be mutually exclusive. Decision is provided with sufficient information without necessarily removing the biological, prognostic, or procedural uncertainty.

**Table 2**

*Construct-Level Descriptive Summary*

| Construct | Mean | Item mean range | Interpretation |
|---|---|---|---|
| Perceived AI reliability | 3.53 | 3.29–3.72 | Moderately positive, conditional trust |
| Communication quality | 3.53 | 3.49–3.63 | Moderately positive |
| Decisional uncertainty | 3.65 | 3.61–3.71 | Substantial acknowledged uncertainty |
| Decision-making confidence | 3.69 | 3.61–3.79 | Moderately high confidence |

*Note.* Values are averages of the four reported item means for each construct on a 1-to-5 scale. Composite-score standard deviations were not available.

**Correlations**

Every one of the six pairwise correlations was significant with a p of less than.001. The perceived AI reliability had a positive correlation with decision-making confidence, r = .538, and the correlation with communication quality was also almost similar, r = .536. The close relationships between decisional uncertainty and confidence show a strong zero-order correlation (r = .625). The predictors also were correlated with each other: reliability versus communication quality, r =.528, and versus uncertainty, r =.566; communication quality versus uncertainty, r =.655. Even though the predictor correlations were not at standard warning levels of severe bivariate multicollinearity, the size of the overlap indicates that communication, perceived credibility and awareness of uncertainty are possibly related to the same underlying decision-making experience.

**Table 3**

*Pearson Correlations Among Study Constructs (N = 300)*

| Variable | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. Perceived AI reliability | — | | | |

| | | | | |
|---|---|---|---|---|
| 2. Communication quality | .528 | — | | |
| 3. Decisional uncertainty | .566 | .655 | — | |
| 4. Decision-making confidence | .538 | .536 | .625 | — |

*Note.* All reported correlations were positive and statistically significant at $p < .001$, two-tailed. Decisional uncertainty retained its original, non-reversed coding.

**Multiple Regression**

The simultaneous regression model was found to be significant, $F(3, 296) = 82.013$, $p <.001$. The three predictors were correlated, R =.674, R 2 =.454, adjusted R 2 =.448; thus, a combination of the three predictors explained 45.4 percent of the variance in decision making confidence. The error of the estimate was 0.543. Perceived AI reliability was a significant positive predictor, B = .238, SE = .055, beta = .234, t = 4.346, p < .001, 95% CI [.130, .346]. Communication quality also contributed positively, B = .139, SE = .052, beta = .157, t = 2.674, p = .008, 95% CI [.037, .241]. Decisional uncertainty had the largest unique coefficient, B = .357, SE = .055, beta = .390, t = 6.447, p < .001, 95% CI [.249, .465]. The model was fitted as confidence = 1.053 + 0.238(reliability) + 0.139(communication) + 0.357(uncertainty). The statistical relationships between these three directional hypotheses were supported by these coefficients.

**Table 4**

*Simultaneous Multiple Regression Predicting Decision-Making Confidence*

| Predictor | B | SE B | β | t | p | 95% CI for B |
|---|---|---|---|---|---|---|
| Perceived AI reliability | .238 | .055 | .234 | 4.346 | < .001 | [.130, .346] |
| Communication quality | .139 | .052 | .157 | 2.674 | .008 | [.037, .241] |
| Decisional uncertainty | .357 | .055 | .390 | 6.447 | < .001 | [.249, .465] |

*Note.* $R = .674$, $R^2 = .454$, adjusted $R^2 = .448$, $F(3, 296) = 82.013$, $p < .001$. Decisional uncertainty was scored in its original direction.

## DISCUSSION

The results indicate that trust in AI-based brain tumour surgical decisions is linked to a set of technological credibility, inter-personal communication, and the open acknowledging of uncertainty. The model described a significant percentage of variability of a self-reported outcome of perception. However, one should not conclude that AI enhanced the survival, neurological outcomes, the diagnostic accuracy, or objective decision quality. Perceived confidence was the dependent variable which may be determined by informed deliberation but also be due to deference to authority, optimism in technology or a general positive judgment of the institution.

AI perceived reliability maintained an independent relationship with confidence following the control of communication and uncertainty. This is in line with the theory of trust-in-automation, which suggests that trust relies upon the beliefs regarding system functions, processes, and goals (Hoff and Bashir, 2015; Lee and See, 2004). The descriptive pattern, nevertheless, is vital: it was identified that the respondents were more positive as regards consistency and apparent validation as compared to patient-specific relevance. This difference is of importance in brain tumour surgery, in which the identical statistical model can behave differently in different tumour types, imaging regimens, patients, and priorities of functionality. The manner in which a system seems sound due to institutional approval or even precise numbers, even when there has not yet been external Check, and subgroup performance does not yet exist. The primarily used type of evidence in clinical confidence should not be interface authority, but contested evidence (Ghassemi et al., 2021; Vasey et al., 2022).

This aspect of reliability can be seen as conditional trust and not seeking more automation. A summary of this stance was given by one respondent: I liked the fact that the system matched my scan with lots of previous cases, but with the surgeon still having to tell me why that match was to my tumour and my life (P01, patient). This assertion relates the technical breadth to the interpretation in cases. It is also linked to evidence that neurosurgical stakeholders tend to be more lenient toward the use of AI as an advisor tool, rather than as a surgeon replacement (Palmisciano et al., 2020). Human oversight is not also a strategy of reassurance. It serves as a protection against automation bias and a way to provide professional responsibility in cases of divergence in computational and clinical judgment (Goddard et al., 2012).

The quality of communication was important and had the lowest unique standardised coefficient. This cannot be interpreted to imply that communication is clinically insignificant. It had an almost zero-order relationship with confidence and communication was strongly correlated with uncertainty. Explanations can be clearly given in a manner that generates confidence in integrity of the process and also raise awareness of limitations. Communication can thus be functioning partly just as the understanding of reliability and uncertainty and then a smaller unique coefficient remains with the overlap variance being held constant.

The most conceptually challenging result is that the coefficient of uncertainty is positive. Since all item uncertainty remained in the original (non-reversed) state, the outcome indicates that those respondents who more recognized unresolved uncertainty also felt higher confidence even with the perceived reliability and communication held constant. It does not demonstrate that confusion brings out confidence and information misses are good. The fact that outcome uncertainty and process confidence are different is the more justifiable interpretation. A respondent might not be aware of what occurs after surgery but still conclude that the decision was open, cautious, justified and in line with patient priorities.

There are a number of restrictions that should be interfere with. To begin with, the aggregate output did not show the stakeholder-group frequencies. The joint regression can also mask the varied meanings and various coefficient magnitudes differently among the patients, family surrogates and neurosurgeons. Second, there was no incorporation of the planned reliability coefficients, item-total statistics, factor analysis, and measurement-invariance testing. The internal consistency and cross group equivalence of the four item scales are not proven hence (Tavakol and Dennick, 2011). Third, the tolerance, variance inflation factor, residual plots, independence tests, assessment of heteroscedasticity, and influence statistics (formal regression diagnostics) were not available. The large F test does not prove that all the assumptions of regression were met.

Fourth, the setting of measurement of all the predictors and the outcome was in the same self-report setting, which resulted in a potential common-method variance and acquiescent response patterns. Fifth, the model adopted was simultaneous as opposed to the predetermined hierarchical model and failed to control the stakeholder group, age and gender, previous experience with AI, direct and vignette exposure, tumour, and institutional trust. It can also be caused in reverse: respondents who initially felt confident might have graded the whole decision making process more desirably. The scale validation, success in assembling the stakeholders, subgroup or interaction analysis, the model diagnostics, and results of prospective studies should include and report scale, as well as achieved stakeholder composition, subgroup or interaction analysis, model diagnostics and the results in terms of decisional regret, knowledge, treatment concordance, and postoperative function.

Nevertheless, the results have implications in practice despite these limitations. Neurosurgical consultations based on AI should introduce the system as someone consultative and approachable, reveal the purposes of its use, and demonstrate the boundaries of validation, and clarify the discrepancy between model and clinician. Communication must reduce probabilities with possible impacts on speech, movement, cognition, independence, rehabilitation, and care giving. Before routine use, hospitals ought to develop local validation, audit trails, overrides procedures and develop clear accountability. Such measures have a higher probability of adopting calibrated confidence, as compared to uncritical technological advocacy or wholesale dismissal of AI.

## CONCLUSION

These include perceived AI reliability, communication quality and perceived decision-making uncertainty, which were all positively related to decision-making confidence in AI-based brain tumour surgery. The two were able to account for 45.4 percent of confidence variance. The results are in line with the socio-technical account whereby confidence is not merely due to the performance of the algorithm, but due to believable evidence, explicable communication, clear uncertainty management, and human responsibility. The surprising positive uncertainty coefficient can best be explained as the presence of the uncertain results along with the belief in the transparent and values-congruent process. Before the implementation of the model to dictate implementation policy, further study is needed to validate the reliability of scales, the assumption of regression, the effects specific to stake holders, and the clinical outcomes.

## REFERENCES


Aninditha, T., Ranakusuma, T. A. S., Malueka, R. G., et al. (2021). Adult brain tumour in Cipto Mangunkusumo General Hospital: A descriptive epidemiology. *Romanian Journal of Neurology*, *20*(4), 480–484. https://doi.org/10.37897/RJN.2021.4.13

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. *Journal of Medical Internet Research*, *22*, e15154. https://doi.org/10.2196/15154

Barry, M. J., & Edgman-Levitan, S. (2012). Shared decision making: Pinnacle of patient-centered care. *New England Journal of Medicine*, *366*(9), 780–781. https://doi.org/10.1056/NEJMp1109283

Elwyn, G., Durand, M. A., Song, J., et al. (2017). A three-talk model for shared decision making: Multistage consultation process. *BMJ*, *359*, j4891. https://doi.org/10.1136/bmj.j4891

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2011). Helping patients decide: Ten steps to better risk communication. *Journal of the National Cancer Institute*, *103*(19), 1436–1443. https://doi.org/10.1093/jnci/djr318

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digital Health*, *3*(11), e745–e750. https://doi.org/10.1016/S2589-7500(21)00208-9

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Hervey-Jumper, S. L., & Berger, M. S. (2016). Maximizing safe resection of low- and high-grade glioma. *Journal of Neuro-Oncology*, *130*(2), 269–282. https://doi.org/10.1007/s11060-016-2110-4

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*, 195. https://doi.org/10.1186/s12916-019-1426-2

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Louis, D. N., Perry, A., Wesseling, P., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: A summary. *Neuro-Oncology*, *23*(8), 1231–1251. https://doi.org/10.1093/neuonc/noab106

Makoul, G., & Clayman, M. L. (2006). An integrative model of shared decision making in medical encounters. *Patient Education and Counseling*, *60*(3), 301–312. https://doi.org/10.1016/j.pec.2005.06.010

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making*, *15*(1), 25–30. https://doi.org/10.1177/0272989X9501500105

Palmisciano, P., Jamjoom, A. A. B., Taylor, D., Stoyanov, D., & Marcus, H. J. (2020). Attitudes of patients and their relatives toward artificial intelligence in neurosurgery. *World Neurosurgery*, *138*, e627–e633. https://doi.org/10.1016/j.wneu.2020.03.029

Parsons, M. V., Buckley, O., Ebadi, H., Leuthardt, E., & McIntosh, T. (2026). Machine learning-assisted prognosis prediction and surgical decision-making for glioblastoma: Perceived benefits and concerns of patients, caregivers, and neurosurgeons. *Frontiers in Neurology*, *17*, 1818670. https://doi.org/10.3389/fneur.2026.1818670

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491–497. https://doi.org/10.1093/jamia/ocz192

Stacey, D., Legare, F., Lewis, K., et al. (2017). Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews*, *4*, CD001431. https://doi.org/10.1002/14651858.CD001431.pub5

Street, R. L., Makoul, G., Arora, N. K., & Epstein, R. M. (2009). How does communication heal? Pathways linking clinician-patient communication to health outcomes. *Patient Education and Counseling*, *74*(3), 295–301. https://doi.org/10.1016/j.pec.2008.11.015

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*, 44–56. https://doi.org/10.1038/s41591-018-0300-7

Vasey, B., Nagendran, M., Campbell, B., et al. (2022). Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nature Medicine*, *28*, 924–933. https://doi.org/10.1038/s41591-022-01772-9

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, *15*(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689

Wiens, J., Saria, S., Sendak, M., et al. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, *25*, 1337–1340. https://doi.org/10.1038/s41591-019-0548-6