BJPsych | The British Journal of Psychiatry (2026)
1–11. doi: 10.1192/bjp.2026.10555

Original Article

Communication of information on benefits and harms of multiple competing medical interventions: three-group, open-label, randomised controlled trial

Edoardo G. Ostinelli, Orestis Efthimiou, Adeola Agunbiade, Huseyin Naci and Andrea Cipriani

Background

Information about a treatment's benefits and harms available to a patient often relies on text. However, for many medical conditions, patients must trade off benefits and harms across multiple competing treatments. It remains unknown how to appropriately communicate information on benefits and harms to patients.

Aims

We compared three communication tools using textual information (Cochrane summary of findings table) or increasing combinations of textual and graphical information (Kilim and Vitruvian plots, respectively) to convey the available evidence.

Method

Communication of Benefit–Risk Information, an online randomised controlled trial, is a three-group, parallel, open-label, automated, randomised controlled trial (no. NCT05917639). We recruited participants aged between 18 and 65 years from the general population. Participants were randomly allocated (1:1:1) to one of the three communication tools providing information on competing fictional treatments for social anxiety, and were asked to choose one based on externally provided preferences. The primary outcome was the perceived level of decisional conflict when selecting a treatment (decisional conflict scale (DCS): 0 = best, 100 = worst). Because this was an all-or-nothing, single-visit trial, only those participants providing data contributed to the primary analyses (modified intention to treat).

Results

We recruited 2178 adults between 1 June and 27 November 2023. Vitruvian and Kilim plots outperformed the Cochrane summary of findings table on the primary outcome (adjusted mean difference −10.9, 95% CI −13.5 to −8.2, $P < 0.0001$ and −9.7, 95% CI −12.4 to −7.1, $P < 0.0001$), respectively). Results varied by participants' literacy and numeracy skills, lived experience of the condition of interest, ethnic group, gender assigned at birth and age.

Conclusions

Combining graphical and textual information, as opposed to text only, improved communication and reduced decisional conflict when choosing across multiple competing medical interventions. Organisations involved in disseminating scientific evidence should consider endorsing a combined graphical and textual approach and adopting more intuitive and accessible communication methods. We identified several prognostic factors that should inform the development of future patient decision aids and communication of scientific findings.

Trial Registration Number

NCT05917639.

Keywords

Shared decision-making; communication of benefits and harms; precision medicine; patient decision aids; digital medicine.

Effective communication of the benefits and harms of medical interventions equips people with the evidence base needed during shared decision-making (SDM) discussions with a healthcare professional.[1,2] Patient decision aids (PDAs), including evidence communication tools, are valuable instruments before, during and after SDM in clinical practice. Their effectiveness is affected by how people understand and engage with the format in which these are delivered.[2,3] When multiple competing interventions are available for the medical condition of interest, trade-offs both within each intervention (i.e. between benefits and harms) and across interventions must be considered.[4] This complex task may translate into decisional conflict, jeopardising engagement and empowerment in one's healthcare.[2,5]

Equity in accessing and understanding information is essential: people accessing medical information have variable levels of numeracy and literacy skills, preferences and needs[6]. If not adequately supported by accessible and user-friendly communication, people may be at risk of misunderstanding and distrusting the available evidence, or may face barriers in accessing optimal levels of care.[2,7,8] To effectively equip people with the essential tools to make choices about their health, simply providing information is not enough[9]. After the supporting evidence is processed and understood, people should feel sufficiently confident in the acquired knowledge to actively engage in SDM, particularly when the treatment options involve significant harms or other types of disutility.[1,2] The combination of graphical and textual communication of benefits and harms has long been advocated to adapt and tailor the presentation of information to the preferences, proficiencies and unique sociocultural backgrounds of users.[1,7,8]

Among the communication tools available for multiple competing medical interventions, the summary of findings table is widely used by Cochrane and GRADE systematic reviews to communicate evidence to wide audiences, including the lay public, using textual information.[10,11] Alternatives, such as Kilim and Vitruvian plots, combine graphical and textual information to communicate how well treatments work, uncertainties and the

availability of evidence.[12,13] However, it remains unclear whether these more advanced tools offer any added benefit for the effective communication of information on the benefits and harms of multiple interventions.

Communication of Benefit–Risk Information (CICERO) is an open-label, three-group, parallel, online, automated, randomised trial aimed at determining how well the three above-mentioned communication tools, which use different levels of textual and/or graphic languages, conveyed benefits and harms information on multiple competing medical interventions to the general population.

Method

Study design and participants

CICERO is a three-group, parallel, open-label, automated, randomised controlled trial. Eligible participants were adults aged between 18 and 65 years (inclusive), willing and able to give informed consent for participating in the study and sufficiently fluent in English. Because the trial was online, potentially eligible participants were required to have access to the internet. There were no country-specific restrictions, and participation in the trial was not restricted to any specific health status (i.e. no definition of the presence or absence of a specific medical condition was defined in the eligibility criteria), with no further eligibility criteria to maximise representativeness within a real-world population. Information about the trial, and the participant information sheet, was distributed by several patient-oriented initiatives and charities to their user base via mailing lists, social media and websites (in alphabetical order): Affa Sair, Bipolar UK (https://www.bipolaruk.org), British Menopause Society (https://thebms.org.uk), Crohn's & Colitis UK (https://crohnsandcolitis.org.uk), Diabetes UK (https://www.diabetes.org.uk), McPin Foundation (https://mcpin.org), MQ Mental Health (https://www.mqmentalhealth.org), NIHR Be Part of Research (https://bepartofresearch.nihr.ac.uk), Pain Concern (https://painconcern.org.uk), Pain UK (https://painuk.org), Parkinson's UK (https://www.parkinsons.org.uk), Psoriasis Association (https://www.psoriasis-association.org.uk), Versus Arthritis (https://www.versusarthritis.org) and Women's Health Concern (https://www.womens-health-concern.org). We selected patient-oriented initiatives and charities focusing on conditions for which multiple competing medical interventions are available, either for the conditions *per se* or their common comorbidities. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation, and with the Helsinki Declaration of 1975 as revised in 2013. All procedures involving human subjects/patients were approved by the University of Oxford (no. R86270). The local ethical committee confirmed the opening of the study on 31 May 2023, and we registered the CICERO trial on Clinicaltrials.gov on the same day (31 May 2023). The first participant was recruited on 1 June 2023, in line with the International Committee of Medical Journal Editors (ICMJE) timeline definition that registration should occur 'at or before the time of first patient enrolment'. Participants were not remunerated for their involvement in the trial.

Randomisation and masking

An automated, computer-based algorithm randomly assigned eligible participants in real time (1:1:1) to one communication tool (i.e. summary of findings table, Kilim plot or Vitruvian plot).[14] All questionnaires were administered using an online platform and followed a standardised process, with no human-to-human interaction throughout the study. The trial consisted of a single online visit ('all-or-nothing' design), starting with the first user-led interaction within the system (informed consent) and ending with the online administration of the questionnaires. Participants could pause the visit at any time and resume their session at their convenience. Because blinding was not possible due to the nature of the interventions, participant information sheets did not include specific information regarding the communication tools investigated in the trial, to limit bias due to knowledge of which intervention participants were allocated to. Similarly, information sheets did not indicate any of the investigated communication strategies as potentially superior to the others. Informed consent from potentially eligible participants was received before the delivery of trial procedures.

Procedures

Eligible participants were asked to complete the baseline assessment, including basic demographic data and an assessment of their health literacy skills (defined as a combination of numeracy and literacy) using a validated questionnaire, the Newest Vital Sign (NVS), UK version).[15,16] We also asked participants whether they had ever received or sought medical attention and medical treatment for social anxiety disorder (social phobia), to control for the potential impact of personal experiences on the health condition used in the fictional scenario. Participants were provided with written instructions to familiarise themselves with the communication tool to which they had been allocated. There was no time restriction on completing the trial-related procedures, and participants could access the instructions at any time (Supplementary Material pp. 8–15 available at https://doi.org/10.1192/bjp.2026.10555). Participants were then presented with a clinical scenario in which they were asked to impersonate a fictional patient, choosing between five different medications for social anxiety after looking at their profiles over five benefit- and harm-related outcomes: response, quality of life, acceptability, headache and nausea. Participants were provided with a given set of preferences on these outcomes and were instructed to carefully consider their trade-offs (Supplementary Material p. 20). The data on the five outcomes for the five medications, and a reference (placebo), were identical across the groups but were differently communicated as per the allocated communication tool. At the end, participants were asked to answer several questions to evaluate the performance of the communication tool, as well as the choice they would have made with their own set of preferences.

For the fictional scenario we chose social anxiety, because this is a common condition among the general population. This condition often starts during teenage years, but can emerge at any age.[17] Although cognitive–behavioural therapy (CBT) is generally considered the optimal treatment for social anxiety, we did not include it in our scenario, the reason being that our goal was to evaluate multiple competing treatments for which sufficient information was available on their overall benefits and harms, whereas harms associated with psychological therapies remain under-explored and under-reported.[18,19] Additionally, CBT may be less well known and intuitively understood compared with 'medications' (we purposedly did not emphasise a specific treatment or treatment class: e.g. we avoided using terms such as selective serotonin reuptake inhibitors or escitalopram).[17]

Interventions

The three interventions differed in whether the various domains related to the communication of benefits and harms information were conveyed using graphical and/or textual information (Box 1). An overview of the administered interventions is

Box 1 Cochrane summary of findings table

Medication 1 versus placebo for social anxiety disorder (social phobia)
Patient or population: people with social anxiety disorder (social phobia)
Intervention: medication 1
Comparison: placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Comments
	Assumed risk	Corresponding risk		
	Placebo	Medication 1		
Response	16 per 100	24 per 100 (19 to 30)	OR 1.69 (1.26 to 1.27)	
Quality of life	16 per 100	23 per 100 (20 to 27)	OR 1.60 (1.32 to 1.94)	
Acceptability	11 per 100	11 per 100 (8 to 15)	OR 0.97 (0.67 to 1.40)	
Headache	2 per 100	3 per 100 (2 to 4)	OR 1.10 (0.64 to 1.89)	
Nausea	5 per 100	4 per 100 (2 to 10)	OR 0.78 (0.30 to 2.07)	

Medication 2 versus placebo for social anxiety disorder (social phobia)
Patient or population: people with social anxiety disorder (social phobia)
Intervention: medication 2
Comparison: placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Comments
	Assumed risk	Corresponding risk		
	Placebo	Medication 2		
Response	16 per 100	24 per 100 (19 to 30)	OR 1.69 (1.26 to 1.27)	
Quality of life	16 per 100	23 per 100 (20 to 27)	OR 1.60 (1.32 to 1.94)	
Acceptability	11 per 100	11 per 100 (10 to 11)	OR 0.97 (0.93 to 1.01)	
Headache	2 per 100	3 per 100 (2 to 3)	OR 1.10 (1.02 to 1.20)	
Nausea	5 per 100	4 per 100 (2 to 10)	OR 0.78 (0.30 to 2.07)	

Medication 3 versus placebo for social anxiety disorder (social phobia)
Patient or population: people with social anxiety disorder (social phobia)
Intervention: medication 3
Comparison: placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Comments
	Assumed risk	Corresponding risk		
	Placebo	Medication 3		
Response	16 per 100	24 per 100 (19 to 30)	OR 1.69 (1.26 to 1.27)	
Quality of life	16 per 100	23 per 100 (20 to 27)	OR 1.60 (1.32 to 1.94)	
Acceptability	11 per 100	9 per 100 (7 to 11)	OR 0.81 (0.63 to 1.04)	
Headache	2 per 100	3 per 100 (2 to 4)	OR 1.10 (0.64 to 1.89)	
Nausea	5 per 100	9 per 100 (5 to 16)	OR 2.02 (1.11 to 3.67)	

Medication 4 versus placebo for social anxiety disorder (social phobia)
Patient or population: people with social anxiety disorder (social phobia)
Intervention: medication 4
Comparison: placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Comments
	Assumed risk	Corresponding risk		
	Placebo	Medication 4		
Response	16 per 100	25 per 100 (20 to 30)	OR 1.75 (1.33 to 2.32)	
Quality of life	See comment	See comment	Not estimable	No data available.
Acceptability	11 per 100	11 per 100 (8 to 15)	OR 0.97 (0.67 to 1.40)	
Headache	2 per 100	3 per 100 (2 to 5)	OR 1.29 (0.79 to 2.10)	
Nausea	5 per 100	4 per 100 (2 to 10)	OR 0.78 (0.30 to 2.07)	

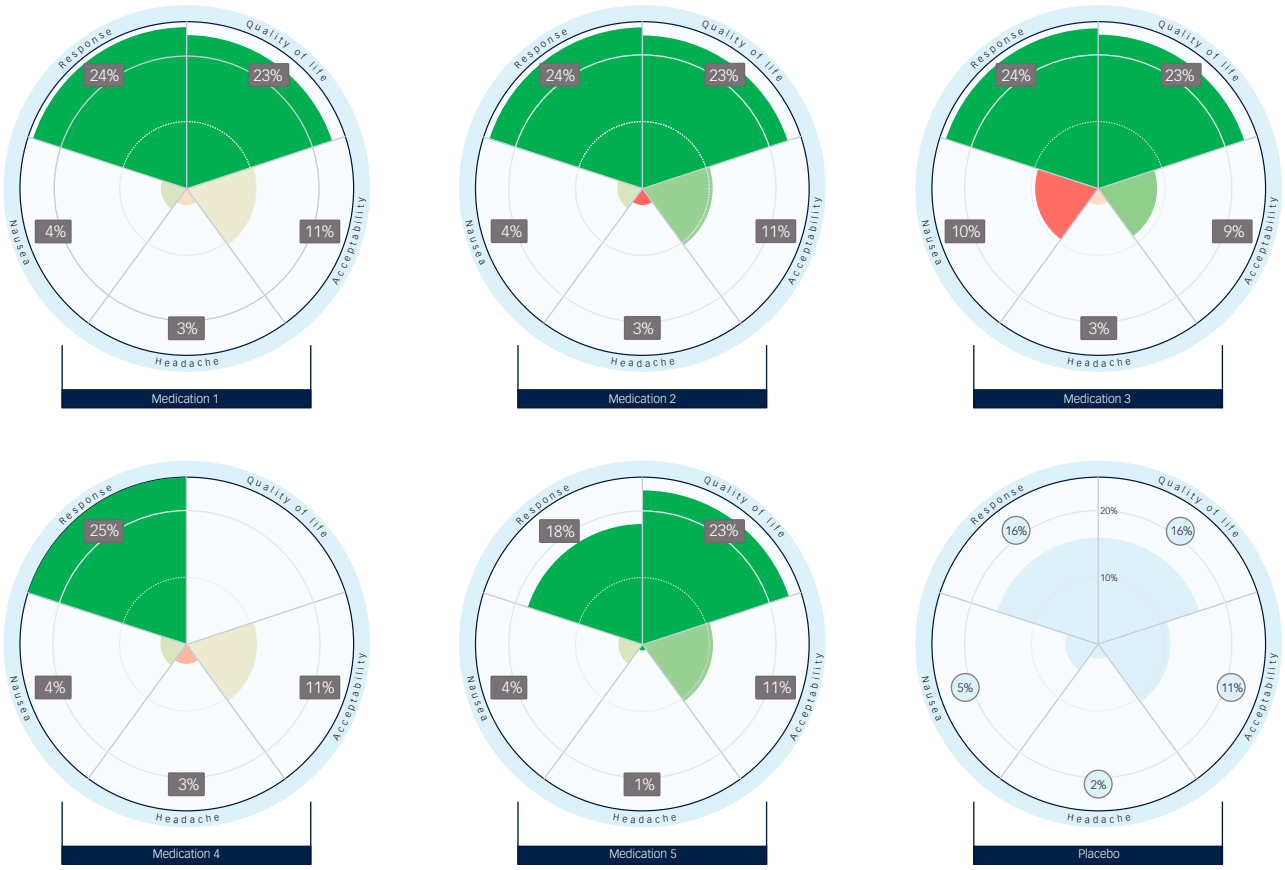
Medication 5 versus placebo for social anxiety disorder (social phobia)
Patient or population: people with social anxiety disorder (social phobia)
Intervention: medication 5
Comparison: placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Comments
	Assumed risk	Corresponding risk		
	Placebo	Medication 5		
Response	16 per 100	18 per 100 (17 to 20)	OR 1.69 (1.26 to 1.27)	
Quality of life	16 per 100	23 per 100 (20 to 27)	OR 1.60 (1.32 to 1.94)	
Acceptability	11 per 100	11 per 100 (10 to 11)	OR 0.97 (0.94 to 1.01)	
Headache	2 per 100	1 per 100 (1 to 2)	OR 0.61 (0.55 to 0.68)	
Nausea	5 per 100	4 per 100 (2 to 10)	OR 0.78 (0.30 to 2.07)	

Kilim plot

	Response	Quality of life	Acceptability	Headache	Nausea
Placebo	15.9%	15.9%	10.9%	2.3%	4.9%
Medication 1	24.0%	23.0%	11.0%	3.0%	4.3%
Medication 2	24.0%	23.0%	11.0%	3.1%	4.3%
Medication 3	24.0%	23.2%	9.2%	3.0%	9.1%
Medication 4	25.2%	–	11.3%	2.9%	3.9%
Medication 5	18.0%	23.0%	11.0%	1.3%	4.3%

Vitruvian plot



Please read the following scenario, and assume you are Mrs Jones.
Mrs Jones has social anxiety disorder (social phobia) and discussed it with her doctor. Mrs Jones is willing to start a medication that can help her feel better.

(a) For Mrs Jones, the treatment must work well (response) and make her feel better (quality of life).
(b) Mrs Jones wants to be certain that she can continue the treatment. Therefore, she prefers certainty over uncertain treatment effects when considering the likelihood of stopping the treatment (acceptability).
(c) Regarding side-effects, Mrs Jones is mainly concerned about not experiencing nausea (nausea). She does not care about headaches (headache) and considers this irrelevant to making a choice.

Set of multiple competing fictional medications across the three investigated communication tools. Participants were provided with instructions for the tool to which they had been allocated, and were asked to explore five possible competing intervention medications. We provided the comparator (placebo) for all the communication tools to contextualise absolute effects. For the Kilim and Vitruvian plots, vibrant green/red (print version: dark blue/black) colours indicate strong statistical evidence of a beneficial/detrimental effect compared with the comparator; light green/red (print version: light blue/grey) indicate weaker evidence of a beneficial/detrimental effect; colours closer to yellow (print version: blue with diagonal pattern) indicate an increasing lack of evidence on whether the intervention performed better or worse than placebo; light blue (print version: pale grey) identifies placebo as the common comparator. OR, odds ratio.

provided in the extended data, with further details in Supplementary Material pp. 16–21.

Box 1 communicates benefit–harm information in a comprehensive format, mainly using textual communication. It is intended for a broad audience, including end-users of systematic reviews and guidelines. The use of summary of findings tables is recommended by Cochrane, and these are included in most of their published systematic reviews. We focused on the following columns in the summary of findings template: outcomes, illustrative comparative risks, relative effects (95% confidence interval) and number of participants and studies.[10]

The Kilim plot conveys benefit–harm information in a tabular format, using a mix of textual and graphical elements. All evidence is combined in a single table, facilitating comparison between treatments across multiple outcomes.[12]

The Vitruvian plot communicates benefit–harm information in both textual and graphical format. The evidence from all outcomes is condensed into a mono-coloured plot per treatment, allowing gauging of its overall profile and exploring trade-offs within and across treatments.[13]

Measures

The communication of information about benefits and harms involves three key steps: (a) understanding the available information; (b) feeling confident about one's own understanding; and (c) feeling prepared to actively participate in SDM. Additional details are provided in Supplementary Material pp. 23–44.

DCS, low-literacy version (primary outcome)

The decisional conflict scale (DCS) measures perceptions of difficulty in making decisions, including lack of information in helping choose between options and feeling unclear about one's own preferences. We used DCS as an overall measure to evaluate

the perceived effectiveness of the communication of information process. DCS can be further organised into four subscales: feeling uninformed (knowledge), feeling uncertain about best choice (certainty), feeling unclear about personal values for benefits and harms (values clarity) and feeling unsupported in decision-making (support).

Information comprehension (secondary outcome)

Following the presentation to participants of a short fictional clinical scenario, they were asked to indicate which medication they would opt for. One of five possible answers had been defined *a priori* as correct during trial development and had been previously tested with five members of the public (Supplementary Material p. 22).[20,21] We also explored the frequency of wrong answers: 'medication 1', which prompted participants to consider the certainty of effects between multiple outcomes; 'medication 3', which prompted participants to jointly consider magnitude, direction and certainty of effects; 'medication 4', which had missing information for one relevant outcome (availability of evidence); 'medication 5', which prompted participants to consider the direction and magnitude of effects; and hesitancy ('I don't know').

DSE (secondary outcome)

The decision self-efficacy (DSE) scale measures self-confidence and belief in a person's ability to make decisions about their own health.

PDM scale (secondary outcome)

The preparation for decision-making (PDM) scale measures a user's perception of how useful a decision support tool is in preparing them to talk to a health professional about a health decision. The original scale has ten items, but one item ('Help you recognise that a decision needs to be made?') was omitted because it was considered not relevant to this study, as done by Hopkin and colleagues.[22]

Time to completion (secondary outcome)

We measured the time required by each participant to familiarise themselves with the intervention (starting time: administration of the intervention) and complete the questionnaires (ending time: completion of all questionnaires). The overall time spent was considered a proxy measure of feasibility.

Drop-out rate (additional analysis)

We measured (binary) discontinuation from the intervention following allocation to an arm of interest (*post hoc* analysis).

Associations among outcomes

We estimated the associations among outcomes throughout the chain of communication of information on benefits and harms: first, participants needed to understand the available information, then to rate how confident they felt about the available information and, finally, how these prompted having an active role in SDM.

Statistical analysis

Information about sample size and power is provided in Supplementary Material p. 4. The primary analyses were conducted following a modified intention-to-treat approach for all the outcomes as described by the risk of bias 2 guidance, because this was a simultaneous trial (all-or-nothing adherence): those who did not initiate the unique study visit were excluded from the analyses.[23] Relative effects between interventions were evaluated via a linear regression model for continuous outcomes and a logistic regression model for binary outcomes, using age, gender assigned at birth, ethnic group (based on United Kingdom Office for National Statistics macro-categories; see Supplementary Material p. 6), literacy and numeracy, education and exposure to social anxiety as covariates. Relative intervention effects were expressed as adjusted mean difference (aMD) and adjusted odds ratio (aOR). To explore the association across outcomes, we estimated the Pearson correlation (ρ) between the considered outcomes and the difference in rating questionnaires across participants who correctly understood the information and those who did not. Additional information on model definition is provided in Supplementary Material pp. 23–61. All analyses were performed in R version 4.3.1 on macOS (R Core Team, Vienna, Austria; https://cran.r-project.org).

Results

Participant characteristics

We randomly assigned 2178 participants: 703 (32.3%), 740 (34.0%) and 735 (33.7%) to the summary of findings, Kilim plot and Vitruvian plot groups, respectively (Fig. 1).

The mean age of the sample was 48.5 ± 15.4 years, with 67% of participants ($n = 1137$) identifying as female (Table 1). Overall, 34% ($n = 582$) reported their highest educational level to be lower than university, 92% ($n = 1562$) identified themselves as White and 20% ($n = 339$) reported NVS scores suggestive of limited literacy and numeracy (Table 1). Overall, 476 participants (22%) did not complete the trial procedures after providing informed consent: 196 (28%) in the summary of findings group, 137 (19%) in the Kilim plot group and 143 (19%) in the Vitruvian plot group (Fig. 1).

Relative effects between interventions

DCS, low-literacy version

The Kilim plot (mean 21.9, s.d. 22.0, adjusted mean difference versus summary of findings −9.7, 95% CI −12.4 to −7.1, $P = 6 \times 10^{-13}$) and Vitruvian plot (mean 20.9, s.d. 21.6, adjusted mean difference −10.9, 95% CI −13.5 to −8.2, $P = 2 \times 10^{-15}$) groups were associated with lower decisional conflict compared with the summary of findings group (mean 32.1, s.d. 24.5). There was no strong evidence of differential effects between participants allocated to the Vitruvian plot and those in the Kilim plot group (aMD −1.1, 95% CI −3.7 to 1.4, $P = 0.38$) (Table 2).

Factors contributing to improved performances on DCS (i.e. prognostic factors) were increasing age (regression coefficient per year of age −0.24, 95% CI −0.3 to −0.2, $P = 2 \times 10^{-9}$) and increasing literacy and numeracy (regression coefficient per point increase −1.4, 95% CI −2.4 to −0.5, $P = 0.003$). Self-identification with a Black ethnic group (coefficient versus White 15.8, 95% CI 3.7 to 27.9, $P = 0.01$) and familiarity with the medical condition used in the clinical scenario (coefficient versus no familiarity 3.0, 95% CI 0.02 to 5.90, $P = 0.05$) were associated with lower performance (Supplementary Material pp. 24–25). For the remaining factors, this study found no evidence of an association. We found similar findings regarding the performance of the interventions when focusing on the four subscales (Supplementary Material pp. 26–33).

Information comprehension

Twenty-six per cent of participants (130 out of 507, Table 2) allocated to the summary of findings group chose the correct medication, as opposed to 39% in the Kilim plot group (238 out of 603; aOR versus summary of findings 1.85, 95% CI 1.42 to 2.41, $P = 5 \times 10^{-6}$) and 53% in the Vitruvian plot group (311 out of

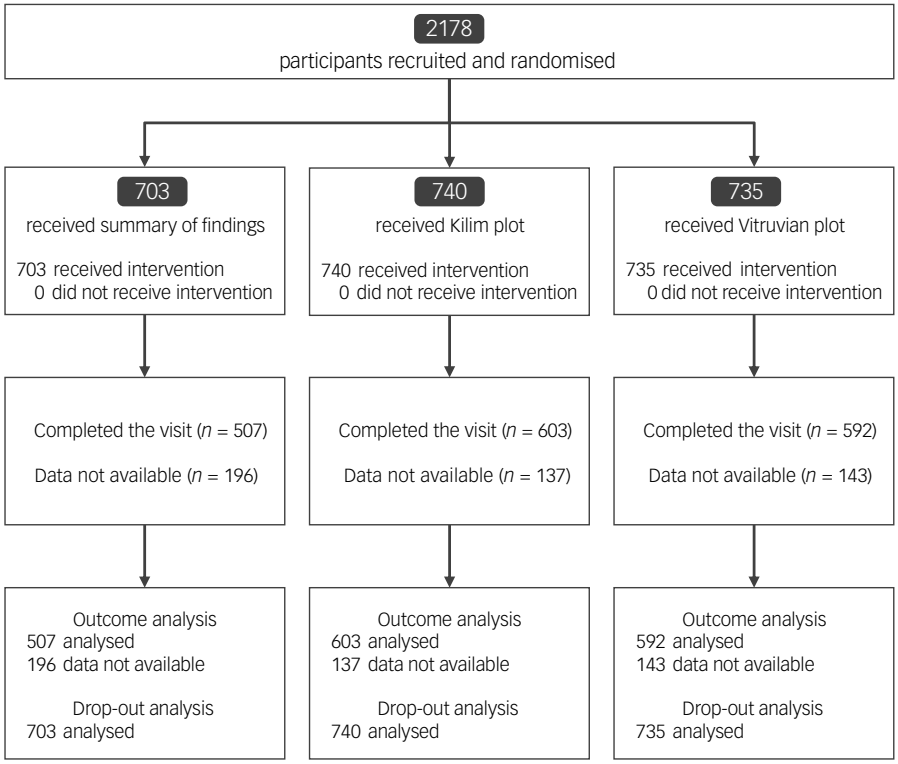


Fig. 1 Flow of participants through the study. Because the automated trial consisted of a single online visit (all or nothing), participants that did not complete the study visit could not contribute data to the outcome analyses.

Table 1 Characteristics of the enroled participants

Variables	Cochrane summary of findings ($n = 507$)	Kilim plot ($n = 603$)	Vitruvian plot ($n = 592$)	Overall ($n = 1702$)
Age (mean ± s.d.)	47.3 ± 16.1	48.3 ± 15.5	49.8 ± 14.6	48.5 ± 15.4
Gender (*n*, %)				
Female	335 (70)	404 (67)	378 (64)	1137 (67)
Male	149 (29)	196 (33)	213 (36)	558 (33)
Prefer not to say	3 (1)	3 (1)	1 (0)	7 (0)
Education (*n*, %)				
Primary school	0 (0)	0 (0)	2 (0)	2 (0)
Secondary school	53 (11)	60 (10)	75 (13)	188 (11)
Upper school	124 (25)	136 (23)	132 (22)	392 (23)
University	208 (41)	233 (39)	225 (38)	666 (39)
Postgraduate	118 (23)	171 (28)	156 (26)	445 (26)
Prefer not to say	4 (1)	3 (1)	2 (0)	9 (1)
Ethnic group (*n*, %)				
Arab	1 (0)	2 (0)	2 (0)	5 (0)
Asian/Asian British	13 (3)	26 (4)	21 (4)	60 (4)
Black	4 (1)	3 (1)	6 (1)	13 (1)
Chinese	6 (1)	4 (1)	0 (0)	10 (1)
Mixed/multiple	11 (2)	5 (1)	13 (2)	29 (2)
White	465 (92)	557 (92)	540 (91)	1562 (92)
Other	7 (1)	6 (1)	10 (2)	23 (1)
Literacy and numeracy (*n*, %)				
Highly likely to be limited	21 (4)	20 (3)	17 (3)	58 (3)
Possibly limited	81 (16)	86 (14)	114 (19)	281 (17)
Adequate	405 (80)	497 (82)	461 (78)	1363 (80)
Lived experience of the medical condition of interest (social anxiety, *n* (%)				
No	404 (80)	501 (83)	463 (78)	1368 (80)
Yes	90 (18)	92 (15)	116 (20)	298 (18)
Prefer not to say	13 (3)	10 (2)	13 (2)	36 (2)

In the analyses, primary and secondary schools were merged into a single category. Percentages do not always total 100, due to rounding.

Table 2 Performance of the three communication tools

Variables	Cochrane summary of findings table ($n = 507$)	Kilim plot ($n = 603$)	Vitruvian plot ($n = 592$)
Primary outcome			
DCS (mean ± s.d.)	32.1 ± 24.5	21.9 ± 22.0	20.9 ± 21.6
aMD (95% CI)	Ref.	−9.7 (−12.4 to −7.1)	−10.9 (−13.5 to −8.2)
aMD (95% CI)	9.7 (7.1 to 12.4)	Ref.	−1.1 (−3.7 to 1.4)
Secondary outcomes			
DCS clarity (mean ± s.d.)	20.0 ± 31.0	12.6 ± 26.6	13.0 ± 26.1
DCS knowledge (mean ± s.d.)	21.6 ± 27.7	13.7 ± 24.2	13.8 ± 24.6
DCS support (mean ± s.d.)	36.0 ± 28.9	25.7 ± 27.0	23.8 ± 26.0
DCS uncertainty (mean ± s.d.)	54.2 ± 37.9	37.9 ± 36.8	35.2 ± 34.9
Information comprehension			
Incorrect (*n*, %)	377 (74)	365 (61)	281 (47)
Correct (*n*, %)	130 (26)	238 (39)	311 (53)
DSE (mean ± s.d.)	66.4 ± 25.7	75.1 ± 23.3	74.1 ± 23.7
PDM (mean ± s.d.)	59.2 ± 27.1	71.8 ± 23.9	68.3 ± 25.4
Time, min (median, IQR)	16.9 (12.3 to 24.1)	14.2 (10.3 to 18.9)	15.5 (11.3 to 21.7)

DCS, decisional conflict scale; aMD, adjusted mean difference; Ref., reference; DSE, decision self-efficacy; PDM, preparation for decision-making; IQR, interquartile range. Lower scores on DCS denote less decisional conflict; higher scores on DSE denote higher self-efficacy; higher scores on PDM denote enhanced preparation for decision-making.

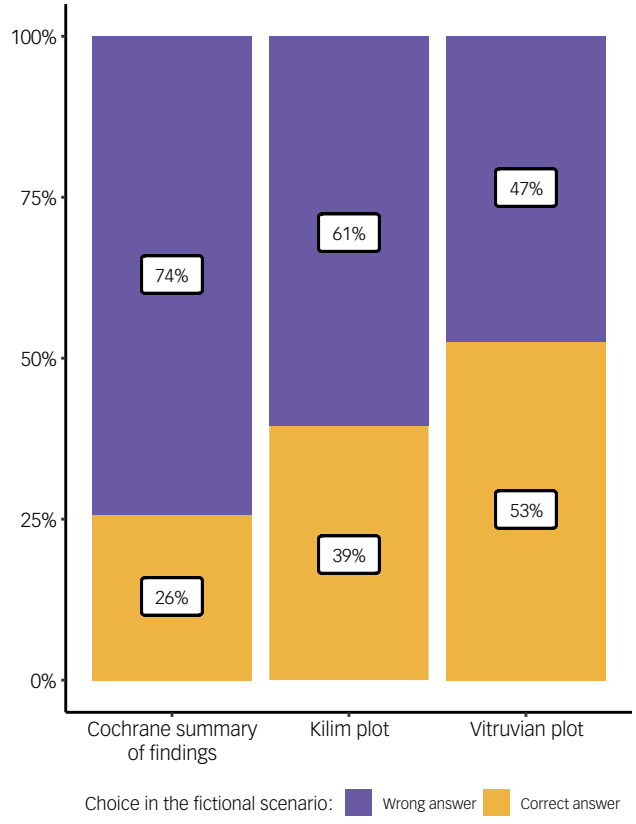


Fig. 2 Scenario-based choice with the given set of preferences and values. Participants were asked to select one of five possible medications, or say that they did not know what to answer. Based on the provided scenario and the given set of preferences and values, only one answer was *a priori* considered correct.

592; aOR versus summary of findings 3.47, 95% CI 2.66 to 4.53, $P < 2 \times 10^{-16}$; aOR versus Kilim plot 1.88, 95% CI 1.48 to 2.38, $P = 2 \times 10^{-7}$) (Fig. 2). Among other prognostic factors, male gender (aOR 1.32, 95% CI 1.05 to 1.63, $P = 0.02$) was associated with higher odds of answering correctly (further details in Supplementary Material pp. 37–38, with an overview of identified prognostic factors on pp. 50–51).

We found similar rates across the 3 interventions in terms of participants selecting 'medication 1' (certainty of effects between multiple outcomes) and 'medication 3' (joint appraisal of magnitude, direction and certainty of effects) as the answer to the fictional scenario. Missing evidence was more commonly overlooked by participants in both the summary of findings (21%, 106 out of 507) and Kilim plot (18%, 109 out of 603) as opposed to those allocated to the Vitruvian plot (4%, 22 out of 592). Similar findings were observed in terms of participants overlooking direction of effects (summary of findings: 22%, 111 out of 507; Kilim plot: 20%, 123 out of 603; Vitruvian plot: 15%, 89 out of 592). Fewer participants allocated to the Kilim plot (1%, 9 out of 603) and the Vitruvian plot (3%, 17 out of 592) were hesitant to make a decision compared with the summary of findings group (8%, 42 out of 507) (Supplementary Material pp 34–38).

DSE scale

Participants allocated to the Kilim plot (mean 75.0, s.d. 23.3, aMD 7.8, 95% CI 5.0 to 10.6, $P = 5 \times 10^{-8}$) and the Vitruvian plot (mean 74.1, s.d. 23.7, aDM 7.4, 95% CI 4.6 to 10.2, $P = 3 \times 10^{-7}$) performed better than those in the summary of findings group (mean 66.4, s.d. 25.7) on DSE, with negligible differences between the first two treatment groups (aMD Vitruvian versus Kilim plot: −0.40, 95% CI −3.1 to 2.3, $P = 0.77$) (Table 2). Further details on prognostic factors are provided in Supplementary Material pp. 40–41.

Preparation for DMS

Participants in the Kilim plot (mean 71.8, s.d. 23.9, aMD 12.3, 95% CI 9.3 to 15.3, $P = 2 \times 10^{-15}$) and the Vitruvian plot (mean 68.3, s.d. 25.4, aMD 8.7, 95% CI 5.7 to 11.7, $P = 2 \times 10^{-8}$) groups performed better than the summary of findings group (mean 59.2, s.d. 27.1) on the PDM, with participants allocated to the Kilim plot group feeling more prepared for decision-making compared with those having accessed the Vitruvian plot (aMD −3.6, 95% CI −6.5 to −0.7, $P = 0.015$) (Table 2). Further details on prognostic factors are provided in Supplementary Material pp. 43–44.

Time to completion

Participants allocated to the summary of findings table required more time (in minutes) to complete the questionnaire (mean 18.4, s.d. 9.1) compared with those in the Kilim plot group (mean 15.2, s.d. 7.5, aMD −16.4, 95% CI −20.8 to −11.9, $P = 7 \times 10^{-13}$) and in the Vitruvian plot group (mean 17.0, s.d. 8.9, aMD −19.4, 95% CI −23.8 to −14.9, $P < 2 \times 10^{-16}$). We did not observe any key differences between the Kilim and Vitruvian plots (aMD −3.0, 95% CI −7.3 to 1.3, $P = 0.2$). Similar findings were recorded when retaining the outliers (Supplementary Material pp. 46–47).

Choices with participants' own set of preferences

When looking at participants' choices based on their own set of preferences and values, those in the Kilim (aOR 4.25, 95% CI 2.63 to 7.10, $P = 1 \times 10^{-8}$) and Vitruvian plot groups (aOR 4.35, 95% CI 2.69 to 7.27, $P = 6 \times 10^{-9}$) had lower odds of being unsure compared with those allocated to the summary of findings table, with the comparison between Kilim and Vitruvian plots showing no evidence of difference (aOR 1.02, 95% CI 0.56 to 1.87, $P = 0.9$). Further details on prognostic factors are provided in Supplementary Material p. 48.

Drop-out rate (additional analysis)

In a *post hoc* analysis, participants allocated to the Kilim (aOR 0.59, 95% CI 0.46 to 0.75, $P = 3 \times 10^{-5}$) and Vitruvian plots (aOR 0.62, 95% CI 0.49 to 0.80, $P = 0.0002$) had lower odds of discontinuing the intervention compared with those in the summary of findings group, with the comparison between Kilim and Vitruvian plots showing no evidence of an effect (aOR 1.06, 95% CI 0.82 to 1.38, $P = 0.7$) (Supplementary Material p. 49).

Associations between outcomes

Overall, better self-efficacy scores positively correlated with improved preparedness for decision-making ($\rho = 0.63$, 95% CI 0.60 to 0.66, $P < 0.001$) and lower decisional conflict ($\rho = -0.60$, 95% CI −0.63 to −0.57, $P < 0.001$); similarly, scores on higher preparedness for decision-making correlated with those on lower decisional conflict ($\rho = -0.61$, 95% CI −0.64 to −0.58, $P < 0.001$). We observed similar findings across the three interventions (Supplementary Material pp. 53–54). Following removal of outliers, we found that time positively correlated with improved performance in terms of self-efficacy, preparedness for decision-making and decisional conflict for participants allocated to summary of findings, but not for the Kilim and Vitruvian plot groups (Supplementary Material pp. 55–56).

Participants who correctly understood the available information reported higher scores for decision self-efficacy (aMD 4.3, 95% CI 2.0 to 6.7, $P = 0.0004$), preparedness for decision-making (aMD 4.7, 95% CI 2.2 to 7.2, $P = 0.0002$) and lower decisional conflict (aMD −3.2, 95% CI −5.4 to −0.9, $P = 0.006$) (Supplementary Material pp. 57–60).

Discussion

The findings from this online randomised trial show that, having compared three communication tools of benefits and harms of multiple medical interventions in 2178 adults from the general population, combined graphical and textual communication (Kilim and Vitruvian plots) was associated with lower decisional conflict in making health-related decisions compared with textual-only information (summary of findings table). This was particularly evident in the first step of communication, with the Vitruvian plot leading to better information comprehension than the Kilim plot and both outperforming the summary of findings table. The same effect was observed across subsequent steps, with people receiving a combination of graphical and textual information feeling more confident about it and more prepared for taking an active role in SDM. Individual literacy and numeracy skills, prior exposure to the medical condition and treatment of interest, ethnic group, age and gender assigned at birth emerged as the key prognostic factors materially impacting individuals' potential engagement with SDM.

Implications

Our findings suggest that well-designed communication strategies can support people in making an informed decision on multiple medical interventions while allowing them to weigh the corresponding benefits and harms.[24] This is in line with earlier evidence on the efficacy of tabular data for single medical treatments among the general population.[20,21] In several medical areas, multiple medical treatments with competing performances are available:[25,26] the decision on whether to start treatment or not, and which treatment to opt for, rely on subjective preferences and values. In our study, combined graphical and textual information provided advantages over text-only information, with the Kilim plot (39%) performing better than the summary of findings table (26%) in terms of information comprehension, and the Vitruvian plot (53%) outperforming both. Although these figures may appear relatively modest in absolute terms, they should be interpreted in the context of our study design: participants were asked to engage with one hypothetical scenario, without the personal relevance or emotional investment typically present in real clinical encounters or if dealing with treatment options that directly affect their own health/condition. Moreover, we intentionally designed the task to be cognitively demanding, by including five treatment alternatives to test the limits of comprehension in a complex decision-making context. The entire online task (from providing baseline information to completion of the final questionnaire) required an active engagement of about 15–20 min, on average, which is considerably higher than what happens when it involves several research steps (e.g. collection of baseline and outcome information). These results underscore the challenges of promoting patient understanding through stand-alone materials, and highlight the importance of simplifying information, reducing cognitive load and integrating clinician support to facilitate effective SDM when developing PDAs.

Correctly understanding medical information was associated with feeling more confident about it, feeling more prepared to engage with SDM and an overall lower decisional conflict. These domains represent key milestones in the sequence of events that lead to informed decision-making. Challenges in the process of interpreting medical information may translate to barriers in actively making decisions about one's own health and engaging with healthcare providers.[2] For instance, not feeling confident may result in disengagement from SDM and, subsequently, in missed opportunities to access better healthcare. Feeling confident and ready to engage in SDM might be jeopardised by an underlying misunderstanding of the available information. Combined graphical and textual communication, and the inclusion of uncertainty and availability regarding the relevant evidence, can reduce these challenges and support patients and clinicians throughout the SDM process.[2] Development of future PDAs should allow for interactive exploration of medical information tailored to individual needs and preferences,[22] with digital interventions representing only one of the ingredients for productive and informed SDM.[2,5]

This study found that increasing levels of literacy and numeracy skills led to higher chances of correctly understanding the provided information, feeling confident about it and experiencing overall lower decisional conflict. This adds to the previously reported positive impact of literacy and numeracy skills on information comprehension.[27] Visualisation relying on graphical communication may help in accommodating variable levels of numeracy skills from individuals accessing medical information. Increasing attention has been advocated to initiatives promoting and supporting health literacy at the individual level,[28] but these changes should be accompanied by reforms of the wider societal and structural architecture.[29,30]

An interesting additional finding is the identification of prior exposure to the medical condition and treatment of interest, ethnic group, age and gender assigned at birth as non-modifiable characteristics that impacted communication of the benefits and harms of medical interventions. Our results support recent calls

voicing concerns on the current state of diversity and inclusivity related to accessing digital health, particularly with regard to women from minority ethnic groups.[31,32] The surge in ethnic minority health apps mirrors the need to culturally tailor PDAs and other e-products to specific ethnic communities.[33] Teams involved in digital healthcare should include people with lived experience and be ethnically, gender- and age-diverse to dismantle inequalities, promote the use of digital interventions and offer trust-building equitable approaches to the community.[2,34,35] One such example is the 'Innovating for all' workstream from the UK National Health Service (NHS) Race and Health Observatory, focusing on the development and deployment of digital interventions with a particular focus on genomics and precision medicine.[35] At the same time, it is important to note how these initiatives should be streamlined and become part of comprehensive and systemic responses to health inequalities.[35]

Our findings have important implications for various entities. First, organisations involved in the production and dissemination of scientific evidence, such as journals and working groups like Cochrane and GRADE, can leverage these findings to improve their communication of evidence syntheses. Second, researchers and practitioners who are developing PDAs can utilise a combination of textual and graphical presentation to promote more effective participation in SDM. Last, committees tasked with formulating research-reporting guidelines can use our findings to improve their standards.

Limitations of the study

Our findings may not apply to everyone, because access to digital devices was required given the online nature of the CICERO trial. Although digital access has sensibly improved for up to 93% of UK households following the COVID-19 pandemic, people aged >65 years and those from lower-income households stand at higher risk of digital divide.[36] When generalising our findings, it is essential to recognise that not all people may want an active role in decision-making, nor would like to take part in it.[2] Our recruited sample showed that one in five had inadequate literacy or numeracy, a finding aligned with the available data on proficiency and education in the UK.[37,38] However, it is important to mention that people identifying with a White ethnic group were probably over-represented, highlighting the need to tackle ethnic inequality in digital health.[34,35]

Furthermore, we did not evaluate all recently developed communication tools for multiple interventions and outcomes, which should be explored in future studies.[39,40] We chose these interventions based on their widespread use (Cochrane's summary of findings table), and on the different approaches taken in using textual and graphical information. The analogies and differences across these interventions allowed us to infer about the effect of combining different sources of information on the communication of benefits and harms.

We acknowledge that the measured duration (15–20 min on average) is hardly compatible with the time constraints of routine clinical practice. However, the measured duration reflects provision of baseline data, familiarisation with the fictional scenario and collection of outcome data. These additional steps would not occur in real-world practice, and we believe that similar communication tools could be used in a more streamlined and targeted manner.

Finally, participants in the CICERO trial were asked to become involved in a scenario involving fictional medications for a specific disorder (social anxiety), and to make a choice according to a given set of preferences. Such an experimental framework may explain the overall performance in understanding medical information, although it is not possible to exclude the possibility that these findings are representative of existing challenges in understanding and trusting science. We believe that our findings are generalisable to medical conditions beyond mental health; however, the effect of these communication strategies when people are effectively making decisions for a condition that is relevant to them, using their own set of values and preferences, should be further explored. For instance, including a human-to-human interaction may allow further insights on the steps leading from information comprehension to becoming engaged with SDM.

What does it mean for the future?

The rapid rise of PDAs that use prognostication has the potential to transform clinical practice. For these innovations to translate into meaningful changes, people must be supported in accessing and interpreting the available evidence – its complexity, its uncertainties and its clinical relevance. Oversimplifying findings ('treatment A is the best') or defaulting to the top-ranked option in a prediction model (digital paternalism) does not constitute genuine SDM. The communication of how competing medical treatments perform requires sensitivity to different levels of health literacy and numeracy skills, using a combination of graphical and textual information.

Edoardo G. Ostinelli, MD, DPhil, Department of Psychiatry, Oxford University, Oxford, UK; Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, Oxford, UK; and Oxford Health NHS Foundation Trust, Oxford, UK; **Orestis Efthimiou**, PhD, Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland; **Adeola Agunbiade**, Independent Scholar; **Huseyin Naci**, PhD, Department of Health Policy, The London School of Economics and Political Science, London, UK; **Andrea Cipriani**, MD, PhD, Department of Psychiatry, Oxford University, Oxford, UK; Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, Oxford, UK; and Oxford Health NHS Foundation Trust, Oxford, UK

Correspondence: Edoardo G. Ostinelli. Email: edoardo.ostinelli@psych.ox.ac.uk

First received 26 Mar 2025, final revision 16 Dec 2025, accepted 17 Jan 2026

Supplementary material

The supplementary material is available online at https://doi.org/10.1192/bjp.2026.10555

Data availability

Pseudonymised patient-level data are available on reasonable request from the corresponding author, following approval by the Medical and Health Research Ethics Committee in the UK.

Acknowledgements

We acknowledge the role of the following user-led or focused charities and initiatives in helping with recruitment for the CICERO trial (in alphabetical order): Affa Sair, Bipolar UK, British Menopause Society, Crohn's & Colitis UK, Diabetes UK, McPin Foundation, MQ Mental Health, NIHR Be Part of Research, Pain Concern, Pain UK, Parkinson's UK, Psoriasis Association, Versus Arthritis, Women's Health Concern. We thank all the participants for their non-remunerated involvement and feedback received throughout the trial.

Author contributions

E.G.O. developed the concept and designed the study. E.G.O. and A.C. implemented the trial. E.G.O. developed the randomisation and data collection platform. E.G.O. and O.E. developed the analysis plan. E.G.O. performed statistical analysis of the data. E.G.O., O.E., A.A., H.N. and A.C. drafted the manuscript. All authors critically revised the manuscript for important intellectual content and gave the final approval for this version to be published. E.G.O. is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding

E.G.O. and A.C. were supported by the National Institute for Health and Care Research (NIHR) Research Professorship (grant no. RP-2017-08-ST2-006); the NIHR Applied Research Collaboration Oxford and Thames Valley (no. ARC OxTV) at Oxford Health NHS Foundation Trust; the NIHR Oxford Health Clinical Research Facility; the NIHR Oxford Health Biomedical Research Centre (grant no. NIHR203316); and the Wellcome Trust (GALENOS project). E.G.O. was supported by the Brasenose College Senior Hulme scholarship. The views expressed are those of the authors and are not necessarily those of the UK NHS, NIHR or the Department of Health and Social Care.

Declaration of interest

All authors have completed the ICMJE uniform disclosure form, at http://www.icmje.org/disclosure-of-interest/. E.G.O. received research and consultancy fees from Angelini Pharma. A.C. received research, educational and consultancy fees from the Italian Network for Paediatric Trials, CARIPLO Foundation, Lundbeck and Angelini Pharma. H.N. reports receiving grants from the Commonwealth Fund, Health Foundation, UK Research and Innovation and NIHR, and personal fees from the World Health Organization, *BMJ* and the Pharmaceutical Group of the European Union, outside the submitted work. All other authors declare no conflicts of interest. E.G.O. and A.C. are members of the *BJPsych* editorial board and confirm they did not take part in the review or decision-making process of this manuscript. The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained. Dissemination to participants and related patient and public communities: a link to the publication will be published on the departmental website and shared with the patient-led initiatives that contributed with recruitment. The main findings of the trial will be presented to patients and the public on social networking sites. A lay English-language summary of the results will be shared with the patient-led initiatives that contributed to recruitment.

Ethical standards

This study was conducted in accordance with the International Council for Harmonisation guidelines for Good Clinical Practice and the Declaration of Helsinki, and was approved by the local institutional review boards or ethics committees of all participating centres (University of Oxford, Central University Research Ethics Committee, no. R86270/RE003).

References

1 Armstrong KA, Metlay JP. Annals clinical decision making: communicating risk and engaging patients in shared decision making. *Ann Intern Med* 2020; **172**: 688–92.

2 National Institute for Health and Care Excellence. *Shared Decision Making (NG197)*. NICE, 2021.

3 Stacey D, Lewis KB, Smith M, Carley M, Volk R, Douglas EE, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2024; **1**: CD001431.

4 The Lancet. Pasteur's legacy in 21st century medicine. *Lancet* 2022; **400**: 2157.

5 Ostinelli EG, Jaquiery M, Liu Q, Elgarf R, Haque N, Potts J, et al. Personalising antidepressant treatment for unipolar depression combining individual choices, risks and big data: the PETRUSHKA tool: personnalisation du traitement antidépresseur de la dépression unipolaire associant choix individuels, risques et mégadonnées: l'outil PETRUSHKA. *Can J Psychiatry* 2025; **70**: 768–81.

6 Grotlüschen A, Desjardins R, Liu H. Literacy and numeracy: global and comparative perspectives. *Int Rev Educ* 2020; **66**: 127–37.

7 McCaffery KJ, Dixon A, Hayen A, Jansen J, Smith S, Simpson JM. The influence of graphic display format on the interpretations of quantitative risk information among adults with lower education and literacy: a randomized experimental study. *Med Decis Making* 2012; **32**: 532–44.

8 Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PK, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. *BMC Med Inform Decis Making* 2013; **13**: S7.

9 Hoffmann TC, Del Mar C. Clinicians' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. *JAMA Intern Med* 2017; **177**: 407–19.

10 Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; **64**: 383–94.

11 Guidelines International Network. *Presenting Treatment Options and Communicating Their Risks and Harms in Patient-directed Knowledge Tools*. Guidelines International Network, 2016 (https://g-i-n.net/toolkit/presenting-treatment-options-and-communicating-their-risks-and-harms-in-patient-directed-knowledge-tools [cited 18 Jul 2025]).

12 Seo M, Furukawa TA, Veroniki AA, Pillinger T, Tomlinson A, Salanti G, et al. The Kilim plot: a tool for visualizing network meta-analysis results for multiple outcomes. *Res Synth Methods* 2021; **12**: 86–95.

13 Ostinelli EG, Efthimiou O, Naci H, Furukawa TA, Leucht S, Salanti G, et al. Vitruvian plot: a visualisation tool for multiple outcomes in network meta-analysis. *Evid Based Ment Health* 2022; **25**: e65–70.

14 SPIRES. *Welcome to SPIRES*. SPIRES, 2025. (https://spires-platform.com [cited 18 Jul 2025]).

15 Powers BJ, Trinh JV, Bosworth HB. Can this patient read and understand written health information? *JAMA* 2010; **304**: 76–84.

16 Rowlands G, Khazaezadeh N, Oteng-Ntim E, Seed P, Barr S, Weiss BD. Development and validation of a measure of health literacy in the UK: the newest vital sign. *BMC Public Health* 2013; **13**: 116.

17 NHS. *Social Anxiety (Social Phobia)*. NHS, 2024 (https://www.nhs.uk/mental-health/conditions/social-anxiety/ [cited 20 Sep 2024]).

18 Meister R, von Wolff A, Mohr H, Nestoriuc Y, Härter M, Hölzel L, et al. Adverse event methods were heterogeneous and insufficiently reported in randomized trials on persistent depressive disorder. *J Clin Epidemiol* 2016; **71**: 97–108.

19 Phillips R, Cro S, Wheeler G, Bond S, Morris TP, Creanor S, et al. Visualising harms in publications of randomised controlled trials: consensus and recommendations. *BMJ* 2022; **377**: e068983.

20 Schwartz LM, Woloshin S, Welch HG. The drug facts box: providing consumers with simple tabular data on drug benefit and harm. *Med Decis Making* 2007; **27**: 655–62.

21 Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. *Ann Intern Med* 2009; **150**: 516–27.

22 Hopkin G, Au A, Collier VJ, Yudkin JS, Basu S, Naci H. Combining multiple treatment comparisons with personalized patient preferences: a randomized trial of an interactive platform for statin treatment selection. *Med Decis Making* 2019; **39**: 264–77.

23 Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, I Boutron, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366**: l4898.

24 Abukmail E, Bakhit M, Del Mar C, Hoffmann T. Effect of different visual presentations on the comprehension of prognostic information: a systematic review. *BMC Med Inform Decis Mak* 2021; **21**: 249.

25 De Crescenzo F, D'Alò GL, Ostinelli EG, Ciabattini M, Di Franco V, N Watanabe, et al. Comparative effects of pharmacological interventions for the acute and long-term management of insomnia disorder in adults: a systematic review and network meta-analysis. *Lancet* 2022; **400**: 170–84.

26 Ostinelli EG, Schulze M, Zangani C, Farhat LC, Tomlinson A, Del Giovane C. Comparative efficacy and acceptability of pharmacological, psychological, and neurostimulatory interventions for ADHD in adults: a systematic review and component network meta-analysis. *Lancet Psychiatry* 2025; **12**: 32–43.

27 Schwartz LM, Woloshin S, Welch HG. Can patients interpret health information? An assessment of the medical data interpretation test. *Med Decis Mak* 2005; **25**: 290–300.

28 NHS England. *Enabling People to Make Informed Health Decisions*. NHS England, 2023 (https://www.england.nhs.uk/personalisedcare/health-literacy/ [cited 18 Jul 2025]).

29 The Lancet. Why is health literacy failing so many? *Lancet* 2022; **400**: 1655.

30 Paakkari L, Balch-Crystal E, Manu M, Ruotsalainen J, Salminen J, Ulvinen E, et al. Health-literacy education drives empowerment and agency. *Lancet* 2023; **401**: 343–4.

31 Chok S. Racism and misogyny persist in digital health. *BMJ* 2024; **384**: q761.

32 The Lancet Digital Health. Empowering women in health technology. *Lancet Digit Health* 2022; **4**: e149.

33 Samarasekera U. The rise of racial minority health apps. *Lancet Digit Health* 2022; **4**: e218–9.

34 Samarasekera U. UK digital health initiatives tackle racial inequality. *Lancet Digit Health* 2022; **4**: e775–6.

35 NHS Race & Health Observatory. *Tackling Race Inequities in Health and Social Care*. NHS Race & Health Observatory, 2024 (https://www.nhsrho.org/ [cited 18 Jul 2025]).

36 OFCOM. *Adults' Media and Attitudes Report 2023*. OFCOM, 2023 (https://www.ofcom.org.uk/__data/assets/pdf_file/0028/255844/adults-media-use-and-attitudes-report-2023.pdf [cited 18 Jul 2025]).

37 National Literacy Trust. *Adult Literacy*. National Literacy Trust, 2025 (https://literacytrust.org.uk/parents-and-families/adult-literacy/ [cited 18 Jul 2025]).

38 Office for National Statistics. *Education, England and Wales: Census 2021*. Office for National Statistics, 2021 (https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/bulletins/educationenglandandwales/census2021 [cited 18 Jul 2025]).

39 Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol* 2016; **76**: 193–9.

40 Daly CH, Mbuagbaw L, Thabane L, Straus SE, Hamid JS. Spie charts for quantifying treatment effectiveness and safety in multiple outcome network meta-analysis: a proof-of-concept study. *BMC Med Res Methodol* 2020; **20**: 266.