ORIGINAL ARTICLE

Do clinicians want recommendations? A multicenter study comparing evidence summaries with and without GRADE recommendations

Ignacio Neumann[a,*], Pablo Alonso-Coello[b], Per Olav Vandvik[c], Thomas Agoritsas[d], Gemma Mas[b], Elie A. Akl[e], Romina Brignardello-Petersen[f], Jose Emparanza[g], Lauren McCullagh[h], Catherine De Sitio[h], Thomas McGinn[i], Hind Almodaimegh[j], Khalid Almodaimegh[k], Solange Rivera[l], Luis Rojas[a], Jérôme Stirnemann[m], Jihad Irani[n], Sani Hlais[o], Reem Mustafa[p], Fadi Bdair[q], Abdelrahman Aly[p], Annette Kristiansen[r,s], Ariel Izcovich[t], Anggie Ramirez[u], Jan Brozek[v], Gordon Guyatt[v], Holger J. Schünemann[v]

[a]*Department of Internal Medicine, Pontificia Universidad Catolica de Chile, Alameda 340, Santiago 8331150, Chile*
[b]*Centro Cochrane Iberoamericano, Instituto de Investigación Biomédica (CIBERESP-IIB Sant Pau), C/ Sant Antoni Maria Claret 167, Barcelona 08025, Spain*
[c]*Department of Medicine, Innlandet Hospital Trust-division Gjøvik, Kyrre Greppsgt.11, Gjøvik 2819, Norway*
[d]*Division of General Internal Medicine & Division of Clinical Epidemiology, University Hospitals of Geneva, Gabrielle-Perret-Gentil 4 Geneva 14 1211, Switzerland*
[e]*Department of Internal Medicine, American University of Beirut Medical Center, P.O. Box: 11-0236, Riad-El-Solh Beirut, Beirut 1107 2020, Lebanon*
[f]*Evidence-Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Av Libertador Bernardo O'Higgins 1058, Santiago, Región Metropolitana 8330111, Chile*
[g]*Clinical Epidemiology Unit, (CASPe-CIBER-ESP), Donostia University Hospital, Paseo Doctor Beriguistain 109, San Sebastian 20014, Spain*
[h]*Department of Medicine, North Shore-LIJ Health System, 600 Community Drive, Suite 300, Manhasset 11030, NY, USA*
[i]*Medicine Service Line Northwell Health, 300 Community Drive, Manhasset 11030, NY, USA*
[j]*College of Pharmacy-Female Branch, King Saud Bin Abdulaziz University for Health Sciences, King Abdulaziz Medical City, PO BOX 22490, Riyadh 11426, Saudi Arabia*
[k]*Family Medicine and Diabetology, Al-Iman Hospital, Riyadh 11544, Saudi Arabia*
[l]*Department of Family Medicine, Pontificia Universidad Catolica de Chile, Alameda 340, Santiago 8331150, Chile*
[m]*Department of General Internal Medicine, Geneva University Hospital, Rue Gabrielle-Perret-Gentil 4, Geneva CH-1211, Switzerland*
[n]*University of Balamand, Faculty of Medicine & Medical Sciences, Beirut, Lebanon*
[o]*Family Medicine Departments, American University of Beirut and Saint Joseph University, Beirut, Lebanon*
[p]*University of Missouri-Kansas City School of Medicine, M4-303 2411 Holmes St., Kansas City 64108-2792, MO, USA*
[q]*Mosaic Life Care, 011 E St Maartens Dr, St Joseph 64506, MO, USA*
[r]*Department of Internal Medicine, Sykehuset Innlandet Hospital Trust, Gjøvik, Norway*
[s]*Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway*
[t]*Department of Internal Medicine, Hospital Alemán, Pueyrredón 1640, Buenos Aires C1118AAT, Argentina*
[u]*International Health Central American Institute Foundation & Cochrane Central America & Spanish Caribbean Branch of the Iberoamerican Cochrane Centre, San Jose, Costa Rica Costa Rica, San José, Santa Ana, Condominio Santa Ana Hills #43, San Jose Zip Code: 10-901, Costa Rica*
[v]*Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University, Main Street West 1200, Hamilton L8S4L8, Ontario, Canada*

Accepted 27 February 2018; Published online 9 March 2018

Abstract

Objectives: Evidence-based clinical practice guidelines provide recommendations to assist clinicians in decision-making and to reduce the gap between best current research evidence and clinical practice. However, some argue that providing preappraised evidence summaries alone, rather than recommendations, is more appropriate. The objective of the study is to evaluate clinicians' preferences, and understanding of the evidence and intended course of action in response to evidence summaries with and without recommendations.

Study Design Setting: We included practicing clinicians attending educational sessions across 10 countries. Clinicians were randomized to receive relevant clinical scenarios supported by research evidence of low or very low certainty and accompanied by either strong or

This study was approved by the Hamilton Integrated Research Ethics Board (13-480) and registered in clinicaltrials.gov (NCT02006017).
* Corresponding author. Tel./fax: +562-2354-3030.
E-mail address: ignacio.neumann@gmail.com (I. Neumann).

weak recommendations developed with the GRADE system. Within each group, participants were further randomized to receive the recommendation plus the corresponding evidence summary or the evidence summary alone. We evaluated participants' preferences and understanding for the presentation strategy, as well as their intended course of action.

Results: One hundred eighty-nine of 219 (86%) and 201 of 248 (81%) participants preferred having recommendations accompanying evidence summaries for both strong and weak recommendations, respectively. Across all scenarios, less than half of participants correctly interpreted information provided in the evidences summaries (e.g., estimates of effect, certainty in the research evidence). The presence of a recommendation resulted in a more appropriate intended course of action for two scenarios involving strong recommendations.

Conclusion: Evidence summaries alone are not enough to impact clinicians' course of action. Clinicians clearly prefer having recommendations accompanying evidence summaries in the context of low or very low certainty of evidence (Trial registration NCT02006017).　© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Using best current evidence is an undisputed goal in clinical decision-making [1]. However, this is not often achieved and a considerable gap between evidence and health care exists [2,3]. Key barriers for effective evidence implementation include the overwhelming amount of research literature and the lack of optimal evidence dissemination strategies [4−6].

One way of presenting evidence to clinicians in the course of their practice is through systematic reviews with summaries presenting tabular or narrative descriptions of best current evidence. The Cochrane collaboration and other systematic reviews organizations now provide structured and critically appraised evidence summaries in user-friendly tabulated formats, based on the GRADE system [7−9]. Such evidence summaries—called "summary of findings tables"—are not accompanied with guidance regarding the best course of action; therefore, clinicians need to correctly interpret the evidence and integrate this information with their clinical expertise, patient values and preferences, and resource considerations to decide about the most appropriate course of action.

A key strategy to integrate other information in evidence dissemination is through the development of clinical practice guidelines. Trustworthy guidelines are based on up-to-date systematic reviews of the best evidence about the desirable and undesirable effects of the relevant alternative management strategies, as well as the best available evidence about patients' values and preferences, resource use, and other context [10].

However, some authors argue that recommendations are unnecessary and may even be counterproductive to genuine evidence-based practice [11,12]. Others argue that although recommendations are useful in the context of high certainty evidence, they may be inappropriate in the context of low or very low certainty evidence [13].

In response to this controversy, we conducted an international multicenter randomized controlled trial to determine whether clinicians prefer (1) recommendations accompanying evidence summaries versus (2) structured evidence summaries alone for clinical questions informed by low or very low certainty of evidence.

2. Methods

2.1. Overview of the design

This study is a hybrid between a survey and a randomized trial. We compared evidence summaries plus recommendations versus evidence summaries alone (protocol in Appendix on the journal's web site at www.elsevier.com). First, we randomized participants to receive either (1) two clinical scenarios involving strong recommendations based on low or very low certainty of evidence or (2) two scenarios involving weak recommendations based on low or very low certainty of evidence (Table 1, Fig. 1). Within each group, we provided respondents with an evidence summary for one scenario and with an evidence summary plus a recommendation for the other scenario. Participants were also randomized to the order in which they received the recommendation, that is, the first or second clinical scenario (Fig. 1).

This study was approved by the Hamilton Integrated Research Ethics Board (13-480) and registered in clinicaltrials.gov (NCT02006017).

2.2. Participants

We included clinicians working primarily in general internal medicine or family medicine. To identify and recruit samples of participants as representative as possible, investigators attended grand rounds or clinical meetings of internists or family doctors at the participating institutions in 10 different countries (see Table 2) and distributed the survey to the attendees at the beginning or end of each meeting.

2.3. Interventions

Each survey presented two scenarios for a specific clinical question. We selected two clinical questions from the World Health Organization guidelines [14,15] and two clinical questions from the 9th edition of the American College of Chest Physicians Antithrombotic Guidelines [16] (Table 1). We chose these questions because of the availability of evidence summaries and recommendations, and because they represented clinical situations that, although relevant for internists and family doctors, were not very frequent in their usual practice. Therefore, clinicians

What is new?

Key findings
- We found that recommendations did not change the understanding or the interpretation of the evidence; it did, however, impact on the intended course of action.

- Clinicians overwhelmingly prefer having recommendations accompanying evidence summaries in the context of low or very-low certainty of evidence.

What this adds to what was known?
- Critics of evidence-based guidelines suggest that recommendations may be problematic on the context of uncertainty regarding the effect of the interventions (low or very-low certainty evidence).

- The results of this trials suggest that recommendations are considered useful by clinicians in the context of uncertainty.

What is the implication and what should change now?
- Guideline panels and organizations should consider the value of making recommendation in the context of low or very-low certainty of evidence.

---

probably would not, in advance, have extensive knowledge of the available evidence or of the recommendations.

Evidence summaries were presented in a tabular format following the recommendations of the Cochrane Collaboration and the GRADE working group for summary of findings tables [9]. In addition, when the recommendation was not presented, we developed a narrative summary of the same evidence to both to provide a similar amount of information in both arms of the trial and to ensure that respondents with different preferences for presentation formats received the information in a manner with which they were comfortable with.

Recommendations were presented as they appeared in the original publication, including the text of the recommendation, the rating of the certainty of the evidence, the strength of the recommendation, and any other remarks. The strength of the recommendation was labeled in the published guidelines as "strong" or "weak," following GRADE nomenclature. Strong recommendations reflect treatment alternatives in which the benefits of the intervention clearly outweigh the harms and in which it is anticipated that all or almost all fully informed individuals would choose the treatment. Weak recommendations reflect a closer balance between benefits and harms with anticipated larger variation in treatment choices and courses of

Table 1. Clinical questions used in the trial

Clinical questions related to strong recommendations based on low or very low certainty evidence
Should oral oseltamivir be used for the treatment of avian influenza (H5N1) in otherwise healthy individuals?
Should aspirin be used for long-term thromboprophylaxis in persons with asymptomatic thrombophilia?
Clinical questions related to weak recommendations based on low or very low certainty evidence
Should daily potassium intake in adults be increased to over 3.5 g?
Should compression stockings be used for thromboprophylaxis for long-distance travelers at increased risk of thrombosis?

action, depending on values and preferences of patients or contextual factors.

2.4. Outcomes

The primary outcome was clinicians' preference for evidence summaries plus a recommendation vs. evidence summaries alone. This outcome was measured on the whole group and was not a comparison between groups (survey part).

We asked participants to what extent they agreed with the following statement: "In the specific scenarios presented, I prefer having a recommendation accompanying the evidence," and elicited response on a seven-point Likert scale (ranging from "I strongly disagree" to "I strongly agree").

We also assessed clinicians' preference for evidence summaries plus recommendations or for evidence summaries alone in general, independently of the clinical scenarios in the survey, asking them to what extent they agreed with the following statement: "In general, in the context of my clinical practice, I prefer having recommendations accompanying the evidence" (using the same seven-point Likert scale).

In addition, we evaluated three secondary outcomes. For the following variables, we compared the group that received a recommendation with the group that did not (randomized trial part).

- Understanding: Following the methods used in previous trials [7,8], we developed four multiple choice questions assessing the understanding of four concepts: baseline risk, risk difference, 95% confidence interval for the risk difference, and certainty of the evidence. To avoid problems with multiple testing, we calculated the proportion of participants with correct answers to the four questions per group rather than the percentage of correct answers per question.
- Interpretation of benefits and harms: We asked participants to judge the magnitude of the balance between the benefits and the harms on a four-point scale. Participant had to choose one of the following options: "the benefits of the intervention clearly outweigh the harms"; "the benefits of the intervention slightly outweigh the harms"; "the harms of the intervention
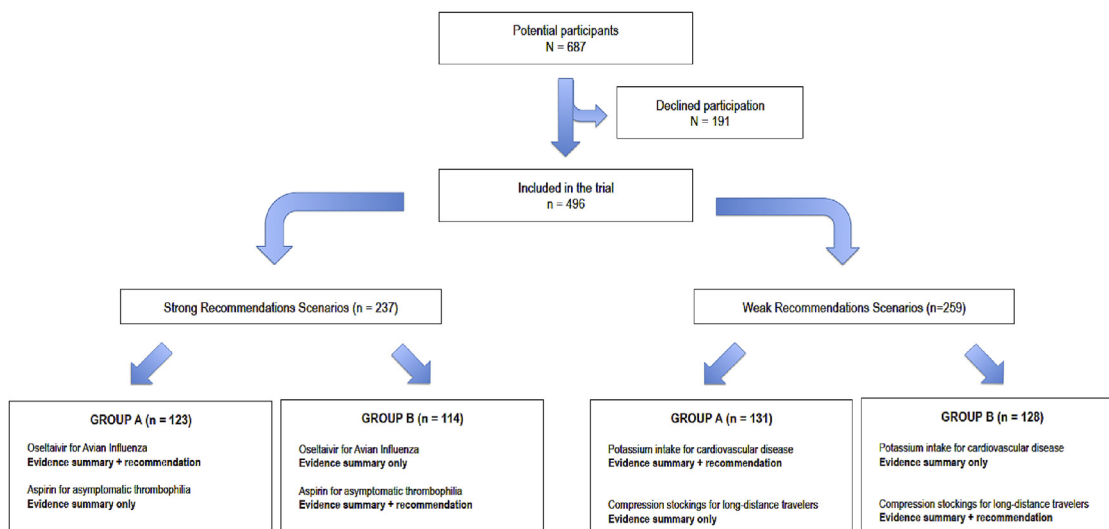
Fig. 1. Flowchart of randomization scheme.

slightly outweigh its benefits''; or ''the harms of the intervention clearly outweigh its benefits''.

- Intended course of action: For the scenarios for which a real panel had issued a strong recommendation despite low certainty in the evidence, we measured the intended course of action by asking participants whether they would recommend the intervention to the patient presented in the clinical scenario. The answers were again collected using a four-point scale that included the following options: ''yes''; ''probably yes''; ''probably no''; or ''no''. We did not analyze this outcome for the scenarios related with weak recommendations, given the inherent multiple valid courses of action for situations in which there is no clearly superior alternative.

2.5. Sample size

We specified in our protocol a sample of 194 participants to detect a difference of 10% in preferences with a power of 80% and alpha level of 0.05. Such sample size calculation, however, assumed no interaction between the type of scenario (related to strong or weak recommendations) and the outcomes of the trial. In a planned exploratory analysis, we detected a statistically significant interaction between strength of the recommendation and the primary outcome (nonparametric Mann-Whitney U test, $P = 0.01$). We thus doubled our sample size ($n = 388$) to have enough power to test the outcomes within each scenario related to strong and weak recommendations.

2.6. Randomization

We arranged the surveys following a computer-generated randomization scheme with varying block sizes and distributed them consecutively. This allowed participants to receive at random either a pair of scenarios

associated with a strong recommendation or a pair of scenarios associated with a weak recommendation (Fig. 1). Of each pair of scenarios one included only an evidence summary and the other an evidence summary plus a recommendation. The order in which participants received the recommendation (i.e., first or second clinical scenario) was also randomized.

2.7. Blinding

The data analysis was conducted in a blinded fashion by one investigator (R.B.P.) who was unaware of the purpose of the study.

2.8. Analysis

We evaluated clinicians' preferences using the nonparametric Wilcoxon signed-rank test, considering the seven-point Likert scale as a continuous variable and assuming that no preference would result in a median of 4 (labeled as ''no preference''). In addition, we dichotomized the primary outcome considering answer options 5, 6, and 7 (from somewhat agree to strongly agree) as preference for having a recommendation.

We analyzed the understanding of the evidence, the interpretation of the balance between benefits and harms, and the intended course of action with the exact Wilcoxon Mann-Whitney rank-sum test. We also calculated risk ratios and 95% confidence intervals for dichotomous outcomes.

We explored the influence of geographical location, professional status, training in health research methodology, and previous use of systematic reviews or evidence summaries on the primary outcome. We hypothesized that participants with training in health research methodology or more familiar with systematic reviews or evidence summaries would report a smaller difference in preferences between evidence summaries plus recommendations versus evidence summaries alone. We hypothesized no important

Table 2. Characteristics of included participants

	Strong recommendation scenarios		Weak recommendation scenarios	
	Group A (*n* = 123)	Group B (*n* = 114)	Group A (*n* = 131)	Group B (*n* = 128)
Females (%)	49 (39.8)	49 (43.0)	69 (52.7)	59 (46.1)
Mean age (SD)	35.8 (11.1)	36.0 (11.1)	35.0 (10.2)	34.8 (9.5)
Country				
Argentina (%)	4 (3.3)	3 (2.6)	3 (2.3)	3 (2.3)
Canada (%)	6 (4.9)	6 (5.3)	8 (6.1)	6 (4.7)
Chile (%)	13 (10.6)	12 (10.5)	15 (11.5)	15 (11.7)
Costa Rica (%)	2 (1.6)	2 (1.8)	1 (0.8)	2 (1.6)
Lebanon (%)	13 (10.6)	12 (10.5)	12 (9.2)	13 (10.2)
Norway (%)	6 (4.9)	4 (3.5)	5 (3.8)	5 (3.9)
Saudi Arabia (%)	18 (14.6)	16 (14.0)	18 (13.7)	18 (14.1)
Spain (%)	18 (18.7)	21 (18.4)	24 (18.3)	23 (18.0)
Switzerland (%)	13 (10.6)	13 (11.4)	15 (11.5)	14 (10.9)
United States (%)	25 (20.3)	25 (21.9)	30 (22.9)	29 (22.7)
Specialty				
Internal medicine (%)	74 (60.2)	63 (55.3)	79 (60.3)	70 (54.7)
Family medicine (%)	34 (27.6)	35 (30.7)	42 (32.1)	38 (29.7)
Other (%)	15 (12.2)	16 (14.0)	10 (7.6)	20 (15.6)
Professional status				
Attending physician (%)	49 (39.8)	47 (41.2)	53 (40.5)	55 (43.0)
Clinical fellow (%)	4 (3.3)	3 (2.6)	2 (1.5)	3 (2.3)
Resident (%)	66 (53.7)	62 (54.4)	73 (55.7)	64 (50.0)
Other (%)	4 (3.3)	2 (1.8)	3 (2.3)	6 (4.7)
Training in health research methodology				
No formal training (%)	51 (41.5)	52 (45.6)	65 (49.6)	54 (42.2)
One or more formal courses (%)	67 (54.5)	57 (50.0)	62 (47.3)	68 (53.1)
MSc or PhD degree (%)	5 (4.1)	5 (4.4)	4 (3.1)	6 (4.7)
Use of evidence summaries				
Once a week or more often (%)	69 (56.1)	56 (49.1)	55 (42.0)	67 (52.3)
Once per month (%)	31 (25.2)	30 (26.3)	44 (33.6)	35 (27.3)
Once per 3 months or less (%)	23 (18.7)	28 (24.6)	32 (24.4)	26 (20.3)
Use of evidence-based recommendations				
Once a week or more often (%)	75 (61.0)	65 (57.0)	77 (58.8)	76 (59.4)
Once per month (%)	33 (26.8)	29 (25.4)	28 (21.4)	31 (24.2)
Once per 3 months or less (%)	15 (12.2)	20 (17.5)	26 (19.9)	21 (16.4)

SD, standard deviation.

differences in preferences across different geographical locations and professional status. We evaluated this interaction using the nonparametric Kruskal-Wallis test.

2.9. Sensitivity analysis

To assess the impact of missing responses on the primary outcome, we tested the following assumptions: (1) all the missing answers would have shown no preference for evidence summaries plus recommendations or for evidence summaries alone (conservative scenario); (2) all the missing answers would have shown a strong preference for not having recommendations accompanying evidence summaries (worst-case scenario).

3. Results

Of the 687 individuals who attended the meetings in which the survey was administrated, 496 (72%) completed the survey, of whom 237 were allocated to the scenarios related to strong recommendations and 259 to the scenarios related to weak recommendations (Fig. 1). Table 1 presents the characteristics of participants.

3.1. Primary outcome

For the scenarios addressing strong recommendations, 189/219 (86%) participants preferred having recommendations in addition to evidence summaries (options 5, 6, and 7 on the Likert scale; most chose 6 or 7) and 193/219 (88%)
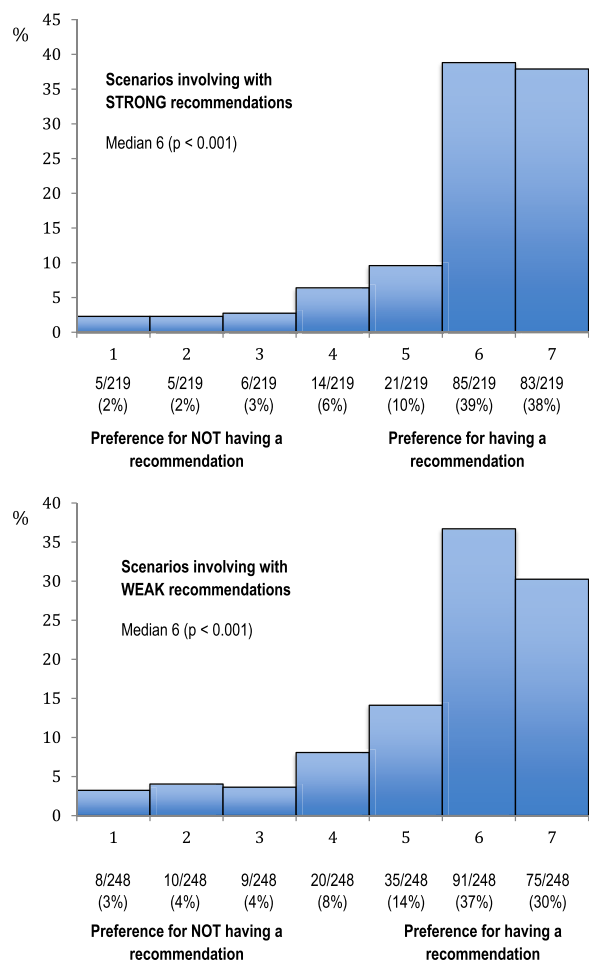
Fig. 2. Clinicians' preferences for evidence summaries plus recommendations versus evidence summaries alone in relation with the clinical scenarios presented in a seven-point Likert scale. The numbers under the bars correspond to the proportion and percentage of participants that answered each option. The *P* values were calculated using the Wilcoxon signed-rank test assuming "4" (no preference) as the hypothetic median.

preferred having recommendations in the context of typical clinical practice (*P* values for both questions <0.001). On the scenarios related to weak recommendations, 201/248 (81%) preferred the addition of recommendation for the scenarios presented and 215/248 (87%) preferred having recommendations for clinical practice (*P* values for both comparisons <0.001) (Fig. 2).

3.2. Secondary outcomes

In most instances, only a minority of respondents answered all questions testing understanding correctly (Table 3). For the questions about oseltamivir, aspirin, and potassium intake, the proportion of participants answering correctly was not influenced by the presence of recommendations (relative risks and 95% confidence intervals (CIs) 0.99 (0.71—1.37), 0.79 (0.52—1.20), and 0.95 (0.70—1.27), respectively).

On the question regarding compression stockings, however, fewer participants answered correctly when there was a recommendation: relative risk, 0.77 95% CI, 0.60—0.99 (Table 3).

For three of the four scenarios, the presence of a recommendation did not change the interpretation of the balance between benefits and harms. In the scenario regarding the use of oseltamivir for avian influenza, the presence of the recommendation significantly increased the number of participants considering that the benefits clearly outweighed harms ($P < 0.001$; Table 3).

The presence of recommendations significantly influenced clinicians' intended course of action. In the scenario addressing the use of oseltamivir for avian influenza, 94/119 (79%) participants would recommend oseltamivir after seeing the recommendation, while 60/112 (54%) would have suggested it after seeing the evidence summary alone ($P < 0.01$). In the scenario regarding the use of aspirin for patients with asymptomatic thrombophilia, 101/111 (91%) participants would not recommend aspirin after seeing the recommendation compared with 102/120 (85%) with the evidence summary alone ($P = 0.01$).

3.3. Subgroup and sensitivity analysis

We observed a significant interaction between geographical location and clinicians' preferences. Participants recruited from Saudi Arabia showed a stronger preference for recommendations (median on the Likert scale = 7) than the rest of the participants (median on the Likert scale = 6, $P < 0.01$). We found no interaction with professional status ($P = 0.8$), training in health research methodology ($P = 0.9$), or previous exposure to systematic reviews or evidence summaries ($P = 0.2$).

There were 29/496 (5.8%) participants who omitted the question evaluating the preference between evidence summaries with or without recommendations. The sensitivity analysis with the conservative assumption (missing answers: no preference) and worst-case scenario assumption (missing answers: strong preference for not having a recommendation) did not appreciably change the results.

4. Discussion

Clinicians participating in this study preferred having recommendations accompanying evidence summaries. In most cases, the presence of a recommendation did not change the understanding or the interpretation of the evidence; it did, however, impact on the intended course of action. Understanding of the information included in the evidence summaries was, for many clinicians, limited.

We found a stronger preference for recommendations in those exposed to strong rather than weak recommendations, but the difference was small (86% and 81%);

Table 3. Understanding and evidence interpretation with and without recommendations

	With the recommendation proportion (%)	Without the recommendation proportion (%)	P Value
Should oral oseltamivir be used for the treatment of avian influenza (H5N1) in otherwise healthy individuals?			
Understanding			
Correct answers (4/4)	45/115 (39)	43/109 (39)	0.62
Interpretation			
Benefits >> harms	34/118 (29)	11/109 (10)	0.03
Benefits > harms	64/118 (54)	81/109 (74)	
Harms > benefits	13/118 (11)	12/109 (11)	
Harms >> benefits	7/118 (7)	5/109 (5)	
Should aspirin be used for long-term thromboprophylaxis in persons with asymptomatic thrombophilia?			
Understanding			
Correct answers (4/4)	27/107 (25)	38/119 (32)	0.82
Interpretation			
Benefits >> harms	5/109 (5)	6/118 (5)	0.31
Benefits > harms	19/109 (17)	26/118 (22)	
Harms > benefits	41/109 (38)	45/118 (38)	
Harms >> benefits	44/109 (40)	41/118 (35)	
Should daily potassium intake in adults be increased to over 3.5 g?			
Understanding			
Correct answers (4/4)	51/126 (40)	52/122 (43)	0.70
Interpretation			
Benefits >> harms	9/125 (7)	3/120 (3)	0.08
Benefits > harms	87/125 (70)	80/120 (67)	
Harms > benefits	23/125 (18)	29/120 (24)	
Harms >> benefits	6/125 (5)	8/120 (7)	
Should compression stockings be used for thromboprophylaxis for long-distance travelers at increased risk of thrombosis?			
Understanding			
Correct answers (4/4)	55/125 (44)	69/121 (57)	0.02
Interpretation			
Benefits >> harms	61/125 (49)	69/124 (56)	0.30
Benefits > harms	56/125 (45)	48/124 (39)	
Harms > benefits	5/125 (4)	4/124 (3)	
Harms >> benefits	3/125 (2)	3/124 (2)	

The *P* values were calculated using the exact Wilcoxon Mann-Whitney rank-sum test.

furthermore, there was no difference when the question referred to recommendations in the context of usual clinical practice (86% and 87%).

Our study has a number of strengths, including the use of real recommendations and evidence summaries, a randomization process to allocate the intervention and its order, and a large international sample from 10 countries of frontline clinicians. The main limitation of our study is the use of a convenience sample of teaching hospitals. However, there was no significant difference in preference between residents and attending clinicians. Also, in academic environments, clinicians are likely more familiar and trained to interpret evidence summaries. Plausibly, the preference for having a recommendation might be even stronger in less academic clinical settings.

Our findings are concordant with the results of the study by Nasser et al. [17] where the uptake on national policy of strong and weak recommendations developed by the World Health Organization was relatively high for both, showing that clinicians, in general, value strong and weak recommendations. Also, the observation of a limited understanding of the information presented on evidence summaries is concordant with the study by Johnston et al. [18], where different formats for continuous outcomes were not well understood or perceived as useful by frontline clinicians.

As our findings suggest, the inclusion of evidence-based recommendations probably will not help clinicians to better understand or interpret evidence summaries. Therefore, strategies to increase clinicians' competencies in evidence-based medicine and to make evidence summaries more accessible and understandable to practitioners are still needed. The GRADE working group, as part of the DECIDE project [6], has developed communication strategies to improve the dissemination of research findings to different target audiences and in particular to clinicians (http://www.decide-collaboration.eu).

We observed that recommendations could influence clinicians' intended behavior and therefore make optimal practice more likely. In our study, despite an appropriate assessment of the benefits and harms of the interventions, some clinicians failed to act accordingly when presented with the evidence summary without the recommendation. In the case of the use of oseltamivir for avian influenza, most clinicians agreed that the potential benefits of the intervention outweighed the harms. However, only 54% would have suggested oseltamivir to the hypothetical patient after reviewing the evidence summary alone versus 79% who had access to the recommendation. The same was true for the use of aspirin in patients with asymptomatic thrombophilia; most clinicians agreed that potential harms of aspirin outweighed the benefits in this situation. In this case, 15% of those who saw only the evidence summary would have suggested using aspirin versus 9% exposed to the recommendation.

In conclusion, our study showed that, in the context of low or very low certainty of evidence, clinicians clearly prefer to have recommendations accompanying evidence summaries.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2018.02.026.

References

[1] Bahtsevani C, Udén G, Willman A. Outcomes of evidence-based clinical practice guidelines: a systematic review. Int J Technol Assess Health Care 2004;20(4):427—33.

[2] McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635—45.

[3] Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med 2014;174(5):710—8.

[4] Haynes B, Haines A. Barriers and bridges to evidence based clinical practice. BMJ 1998;317:273—6.

[5] Glenton C, Underland V, Kho M, Pennick V, Oxman AD. Summaries of findings, descriptions of interventions, and information about adverse effects would make reviews more informative. J Clin Epidemiol 2006;59:770—8.

[6] Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brożek J, et al, DECIDE Consortium. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (decide): protocol and preliminary results. Implement Sci 2013;8:6.

[7] Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. J Clin Epidemiol 2010;63:620—6.

[8] Carrasco-Labra A, Brignardello-Petersen R, Santesso N, Neumann I, Mustafa RA, Mbuagbaw L, et al. . Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. J Clin Epidemiol 2016;74:7—18.

[9] 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.

[10] Graham R, Mancher M, Dianne MW, Greenfeld S, Steinberg E. Clinical practice guidelines we can trust. 1st ed. Washington (DC): Institute of Medicine of the National Academies; 2011.

[11] Fava GA. Guidelines and conflicts of interest. Ann Intern Med 2010; 153:421—2.

[12] Parmley WW. Clinical practice guidelines. Does the cookbook have enough recipes? JAMA 1994;272:1374—5.

[13] U.S. Preventive Services Task Force. Grade definitions 2012. Available at: http://www.uspreventiveservicestaskforce.org/. Accessed October 10, 2017.

[14] Schünemann HJ, Hill SR, Kakad M, Bellamy R, Uyeki TM, Hayden FG, et al. WHO rapid advice guideline panel on avian influenza. WHO rapid advice guidelines for pharmacological management of sporadic human infection with avian influenza A (H5N1) virus. Lancet Infect Dis 2007;7:21—31.

[15] WHO. Guideline: Potassium intake for adults and children. Geneva: World Health Organization (WHO); 2012.

[16] Kahn SR, Lim W, Dunn AS, Cushman M, Dentali F, Akl EA, et al. Prevention of VTE in nonsurgical patients: antithrombotic therapy and prevention of thrombosis, 9th ed.: American College of chest Physicians evidence-based clinical practice guidelines. Chest 2012; 141:e195S—226S.

[17] Nasser SM, Cooke G, Kranzer K, Norris SL, Olliaro P, Ford N. Strength of recommendations in WHO guidelines using GRADE was associated with uptake in national policy. J Clin Epidemiol 2015;68:703—7.

[18] Johnston BC, Alonso-Coello P, Friedrich JO, Mustafa RA, Tikkinen KA, Neumann I, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. CMAJ 2016;188:25—32.