BMJ Open

Multilayered and digitally structured presentation formats of trustworthy recommendations: a randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011569
Article Type:	Research
Date Submitted by the Author:	02-Mar-2016
Complete List of Authors:	Brandt, Linn; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Vandvik, Per; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Alonso-Coello, Pablo; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Akl, Elie; American University of Beirut, Department of Internal Medicine Thornton, Judith; National Institute for Health and Care Excellence Rigau, David ; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau Adams, Katie; National Institute for Health and Care Excellence O'Connor, Paul; National Institute for Health and Care Excellence Guyatt, Gordon; Mcmaster University, Department of Clinical Epidemiology and Biostatistics Kristiansen, Annette; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Health informatics, Epidemiology, Communication
Keywords:	Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, EPIDEMIOLOGY

Multilayered and digitally structured presentation formats of trustworthy recommendations: a randomised trial

Linn Brandt[1,2], Per Olav Vandvik Pablo[1,2], Pablo Alonso-Coello[3], Elie A. Akl[4,6], Judith Thornton[5], David Rigau[3], Katie Adams[5], Paul O'Connor[5], Gordon Guyatt[6], and Annette Kristiansen[1,2].

Author Affiliations: [1)] Department of Internal Medicine, Sykehuset Innlandet Hospital Trust, Gjøvik, Norway (Drs. Brandt, Kristiansen and Vandvik), [2)] Institute of Health and Society, Faculty of Medicine, University of Oslo (Drs. Brandt, Kristiansen and Vandvik), [3)] Iberoamerican Cochrane Centre, Biomedical Research Institute (IIB Sant Pau-CIBERESP), Barcelona, Spain (Drs. Alonso-Coello and Rigau), [4)] Department of Internal Medicine, American University of Beirut, Lebanon (Dr. Akl), [5)] National Institute for Health and Care Excellence, Manchester UK (Judith Thornton, Katie Adams, Paul O'Connor), [6)] Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (Drs. Akl and Guyatt).

Corresponding Author: Linn Brandt, Department of Internal Medicine, Sykehuset Innlandet Gjøvik, Kyrre Grepps gate 11, 2819 Gjøvik, Norway. E-mail: brandtlinn@gmail.com

Designers: Rob Fracisco and Frankie Achilles helped designing the novel multilayered format.

Abstract

Objectives: To investigate practicing physicians preferences, perceived usefulness and understanding of a new multilayered guideline presentation format - compared to a standard format - as well as conceptual understanding of trustworthy guideline concepts.

Design: Mixed survey and randomized controlled trial through a standardised lecture for physicians. We presented participants with a clinical scenario and randomised them to view a guideline recommendation in multilayered or standard format. Both groups were presented and asked about guideline concepts. Participants answered multiple-choice questions by use of clickers.

Setting: Mandatory educational lectures in six non-academic and academic hospitals and three primary care centres in Lebanon, Norway, Spain and United Kingdom.

Participants: 181 practicing physicians in internal medicine (156) and general practice (25) attending the lectures.

Interventions: A new digitally structured, multilayered guideline presentation format and a standard narrative presentation format currently in widespread use.

Primary and secondary outcome measures: Our primary outcome was preference for presentation format. Understanding, perceived usefulness and perception of absolute effects were secondary outcomes.

Results: 72% (95% CI 65-79) of participants preferred the multilayered format and 16% (95% CI 10-22) preferred the standard format. A majority agreed that recommendations (multilayered 86% vs. standard 91%, p-value = 0.31) and evidence summaries (79% vs. 77%, p-value = 0.76) were useful in the context of the clinical scenario. 72% of participants randomised to the

multilayered format vs. 58% for standard formats reported correct understanding of the recommendations (p-value = 0.06). Most participants elected an appropriate clinical action after viewing the recommendations (98% vs. 92%, p-value = 0.10). 82% of the participants considered absolute effect estimates in evidence summaries helpful or crucial.

Conclusions: Clinicians clearly preferred a novel multilayered presentation format to the standard format. Whether the preferred format improves decision-making and has an impact on patient important outcomes merits further investigation.

Trial registration: None.

Strengths and limitations of this study

- We conducted a multicentre trial targeting regular educational sessions.
- Both formats were taken from published guidelines, and we used a comparator format that most participants were familiar with.
- To avoid peer pressure, we ensured participants anonymity through use of clickers (audience response technology).
- A weakness of the study is having researchers involved in development of the new presentation format perform most of the educational sessions.
- We did not measure impact of alternative formats on clinical decisions or patient important outcomes.

BACKGROUND

Clinical practice guidelines that provide recommendations addressing diagnosis and treatment can help clinicians optimize their evidence-based practice (EBP) at the point of care.[1] An abundance of guidelines are available, but many have shortcomings with their trustworthiness and dissemination strategies.[2,3] New standards for trustworthy guidelines developed by the Institute of Medicine and the Guideline International Network highlight the need for more rigorous development processes.[4–6]

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (www.gradeworkinggroup.org) represents a systematic, explicit and transparent process for evaluating and reporting quality of research evidence and for moving from evidence to recommendations.[7,8] GRADE facilitates the creation of trustworthy guidelines and has been adopted by more than 90 organisations worldwide.

Implementation of guidelines requires effective dissemination of recommendations.[9] Guidelines should generally answer clinicians' informational needs within two minutes, which implies that recommendations need to be easy to find, understand, apply and share at the point of care.[10] With these challenges in mind, the GRADE working group initiated the Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE) project.[11] DECIDE aimed to improve dissemination of evidence-based recommendations for a range of stakeholders, including health care professionals, policy makers and patients.

Through brainstorming, stakeholder feedback and usability testing with iterative improvements, we developed a multilayered guideline presentation format (figure 1) targeted at health care professionals.[12] During usability testing we confirmed previous research demonstrating limitations in clinicians' conceptual understanding of key standards for trustworthy guidelines. We have since deployed the multilayered format in real-life guidelines.[13] Uncertain of the relative merits of our novel versus existing formats, we conducted a combined survey and

randomised controlled trial to determine clinicians' preferences for the new multilayered presentation format versus a traditional format for guidelines, the perceived usefulness of guideline recommendations and understanding of key concepts of trustworthy guidelines.

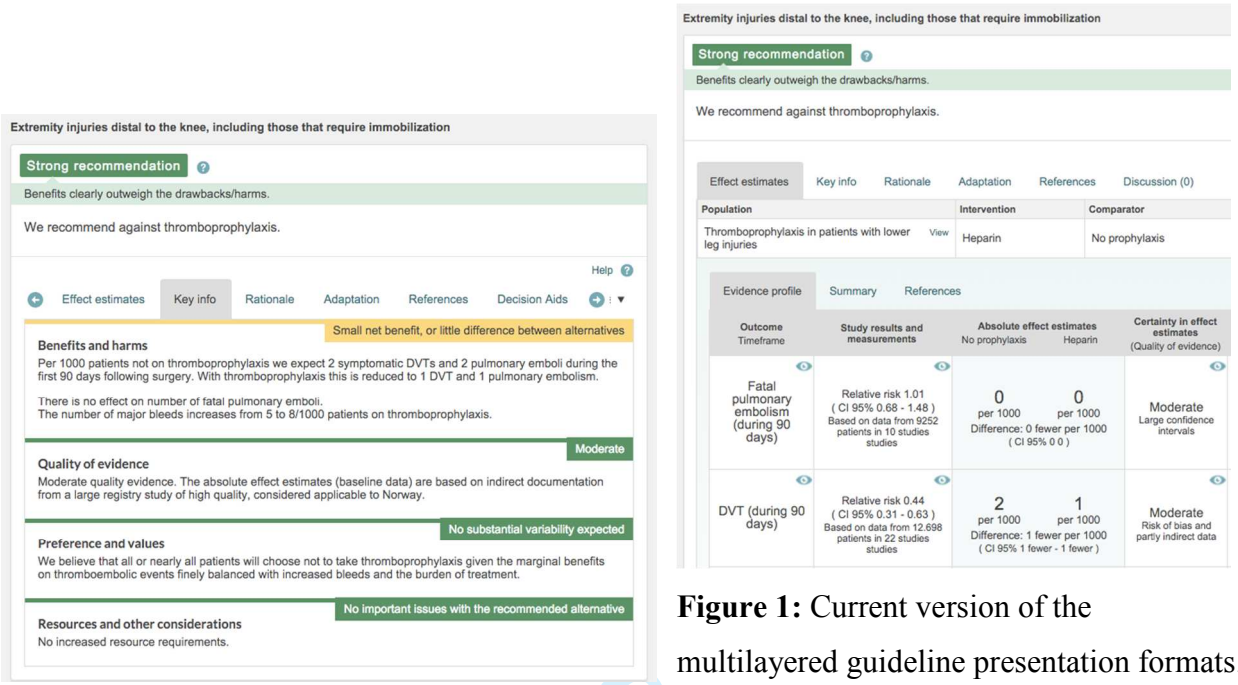


Figure 1: Current version of the multilayered guideline presentation formats.

MATERIALS AND METHODS

This study is the result of a collaboration of DECIDE and the non-profit MAGIC (Making GRADE the Irresistible choice) innovation and research programme (www.magicproject.org), which aims at facilitating the efficient creation, dissemination and dynamic updating of trustworthy guidelines and evidence summaries. As previously reported, we have in MAGIC created a web-based authoring and publication platform (www.magicapp.org) to allow guideline recommendations to be created and published in digitally structured multilayered formats on all devices.[14]

Study design, setting and participants

We applied a mixed survey and randomised controlled trial. We included practicing physicians in internal or family medicine. In order to recruit representative samples of participants, investigators targeted compulsory educational sessions at teaching hospitals. To recruit general

practitioners, we invited them to attend an educational session performed within the context of this study. We performed four pilot sessions and made revisions based on experiences in these sessions.

Participants were considered to provide consent by accepting to answer the survey questions. Research ethics boards in each participating country approved the study.

Standardised lecture and study procedure

At each site, a member of the research team delivered a standardised lecture titled "New standards for trustworthy guidelines " (appendix 1) and conducted the study according to a predefined protocol. Participants provided anonymous answers to questions with predefined response categories using clickers. The questions - as well as screenshots of presentation formats - were embedded in the lecture slides using TurningPoint© and read out loud by the presenter. For the Norwegian and Spanish sites we translated the presentation to their native language; while in Lebanon and the UK the questions were presented in English, which is commonly used in medical education.

The lecture and study procedure included the following components:

1. Collection of the demographics of the participants, their current preference for information resources and understanding of the GRADE system.
2. Presentation of a clinical scenario concerning choice of oral anticoagulation in a patient with atrial fibrillation and high risk of stroke (CHA2DS2-VASc score 2).
3. Presentation of guideline recommendations and evidence summaries relevant to the clinical scenario, sequentially presented in the two alternative guideline presentation formats through randomisation and blinding as outlined below. Both formats were shown side by side to both groups at the end.
4. Short presentation of key conceptual definitions of trustworthy guidelines, specifically explanation of the strength of recommendations and quality of the evidence using the GRADE methodology.

Guideline presentation formats

Figure 2 shows the standard narrative guideline presentation format (format A) and the new multilayered guideline presentation format (format B). We extracted both formats from existing guidelines, but masked the publishers' identity to avoid potentially biasing participants' responses. Both formats were displayed as screenshots in the presentation slides.

Multilayered format: The experimental multilayered format displays recommendations up front with supporting information as collapsible boxes provided by clicking on the recommendation itself. The strength of the recommendation is communicated by use of text and colour coding, and a header describes the population for which the recommendation applies. The example was taken from a Norwegian guideline for antithrombotic therapy applying the multilayered presentation formats published in MAGICapp and translated to English for the purpose of this study.[13,15]

Standard format: The control standard format displays recommendations and an abridged evidence summary from UpToDate.[16] We considered UpToDate's presentations a suitable reference standard for current presentation formats due to its widespread use, commonality with other guidelines, evidence-based approach and use of GRADE.[17,18] UpToDate provides a textual summary of its recommendations in bullet points with links to the supporting information in the main text or in other articles. The recommendations are labelled with numbers (1 & 2) and letters (A-C), depicting the strength of the recommendation and quality of evidence respectively.

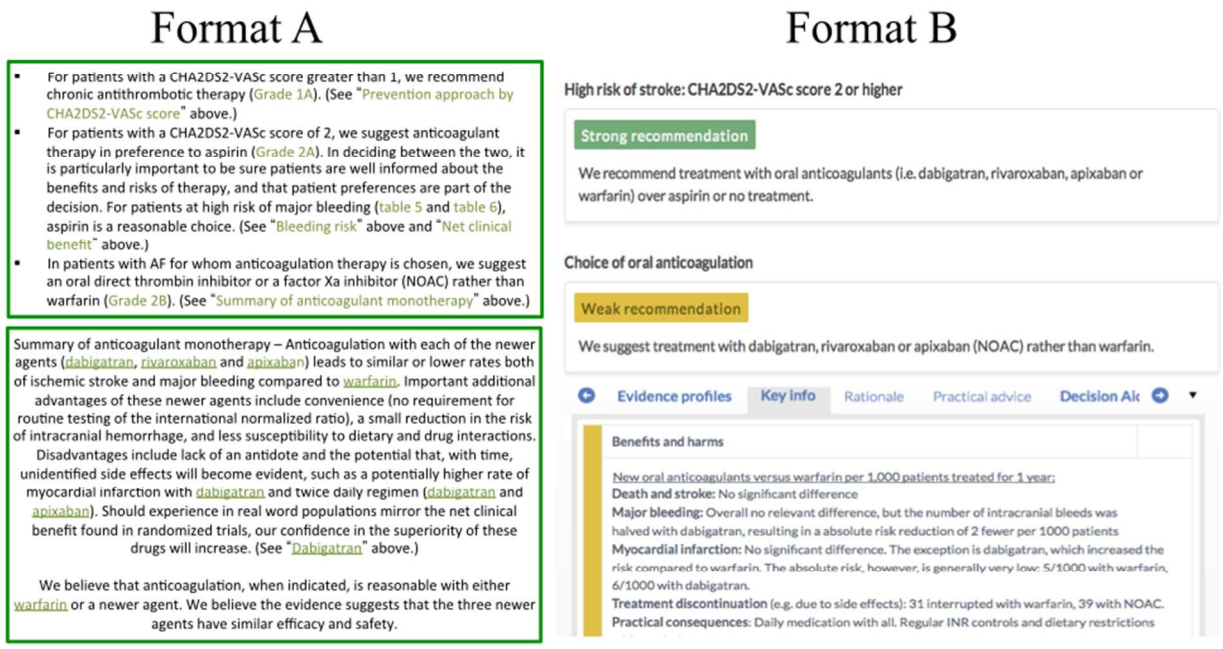


Figure 2: Standard (A) and multilayered guideline (B) presentation formats.

Randomisation and blinding

A research member colour marked the base of half the clickers and rearranged them in their container in a random manner. The presenter handed out the clickers to the participants front-up to conceal the marking. Participants with marked clickers were randomised to group B, being presented the multilayered format. We asked participants in either group to don blindfolds when the other group were shown their allocated presentation format. Questions during randomisation were not read out loud.

Outcome measures

Our primary outcome was preference for either presentation format. We provided three response options: preference for the standard format, no preference and preference for the multilayered format.

Secondary outcomes included:

- Understanding of 1) the evidence summaries recommendations and 2) the recommendation with four potential answers, one alternative being correct.
- Anticipated course of action to the clinical scenario with four potential answers, two alternatives being correct.
- Perceived usefulness of 1) evidence summaries and 2) recommendations. Participants provided answers to the statement "This information/recommendation would help me manage my patient" on a 6 point Likert scale: 1= strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree and 6=strongly agree.
- Correct understanding of the strength of the following example recommendation and the confidence in effect estimates: "We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol). GRADE 2B." We provided participants with four potential answers, one alternative being correct.
- Perceived understanding of the strength of the recommendation. We asked the participants the following question twice, before and after being provided with a short written explanation: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision-making." They provided answers on a 6 point Likert scale with 2 anchors: 1= strongly disagree and 6 = strongly agree.
- Participants' perception of presenting absolute effect estimates. We asked participants "What is your first reaction to being presented with absolute effects?" The answers were collected using a 5-point scale: 1=confusing distraction, leave it out, 2=a little confusing, but not a big problem, 3=doesn't help, but doesn't hurt, 4=not crucial, but helpful and 5=crucial information, should always be included.

Statistical analysis

We dichotomized all outcomes to either correct/incorrect or agree/disagree, and analysed them by use of Pearson's Chi-square test. We included all randomised participants that answered more than one question in the final analysis. We used SPSS (version 23) for all analyses.

RESULTS

We included 181 practicing physicians across four countries and nine centres (Lebanon 1, Norway 6, Spain 1 and UK 1), 177 were randomised and 174 (96%) answered > 1 question and are included in the final analysis. Their demographics and information resource preferences are provided in table 1. The two groups were fairly similar. One centre (27 participants) in Norway did not include demographic questions, so the demographic presentation includes the remaining eight centres.

Number of participants randomised	**Multilayered format** 92	**Standard format** 85
Country		
(# participants eligible for analysis)	(92)	(83)
Norway (%)	61 (66.3%)	57 (68.7%)
UK (%)	10 (10.9%)	10 (12%)
Lebanon (%)	11 (12.0%)	10 (13.3%)
Spain (%)	10 (10.9%)	5 (6%)
Professional status or Specialty		
(# participants eligible for analysis)	(76)	(72)
Medical student or intern (%)	13 (17.1%)	8 (11.1%)
Internist Resident (%)	21 (27.6%)	27 (37.5%)
Internist Attending/Consultant (%)	23 (30.3%)	21 (29.2)
General practitioner (%)	13 (17.1%)	12 (16.7%)
Unknown (% did not answer that question)	6 (7.9%)	4 (5.6%)
Training in health research methodology		
(# participants eligible for analysis)	(72)	(68)
No training in HRM (%)	34 (47.2%)	35 (51.2%)
≥ 1 HRM course (%)	26 (36.1%)	21 (30.9%)
Degree in HRM (%)	12 (16.7%)	12 (17.6)
Preferred knowledge source		
(# participants eligible for analysis)	(89)	(82)
Local guideline (%)	22 (24.7%)	14 (17.1%)
Systematic review (%)	2 (2.2%)	2 (2.4%)
EBM textbook (%)	17 (19.1%)	13 (15.9%)
National or international guideline (%)	34 (38.2%)	36 (43.9%)
Colleague (%)	14 (15.7%)	17 (20.7%)
Primary study (%)	0 (0%)	0 (0%)

Table 1: Demographics of the groups randomised to different formats. HRM=Health Research Methodology. EBM=Evidence Based Medicine.

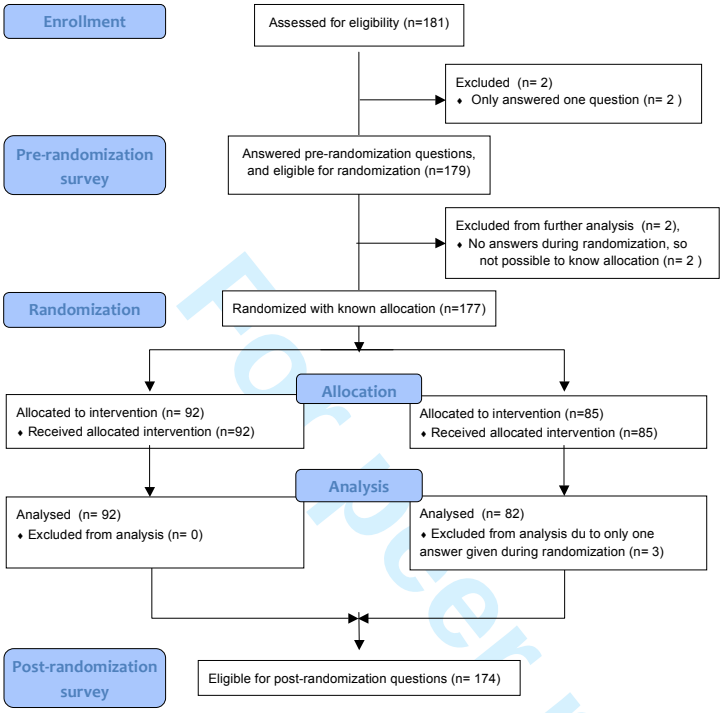


Figure 3: Flow chart of design and enrolment of participants to the multilayered format (n=92) and standard format (n=85)

Preference for alternative presentation formats

When exposed to both formats after completing the randomised part of the study 113 of 156 (72%, 95% CI 65-79) participants preferred the new multilayered format, 25 (16%, 95% CI 10-22) preferred the standard format, and 18 (12%, 95% CI 7-17) reported no preference. Results were very similar in those randomised to the standard vs. the multilayered format (p = 0.66, figure 3).

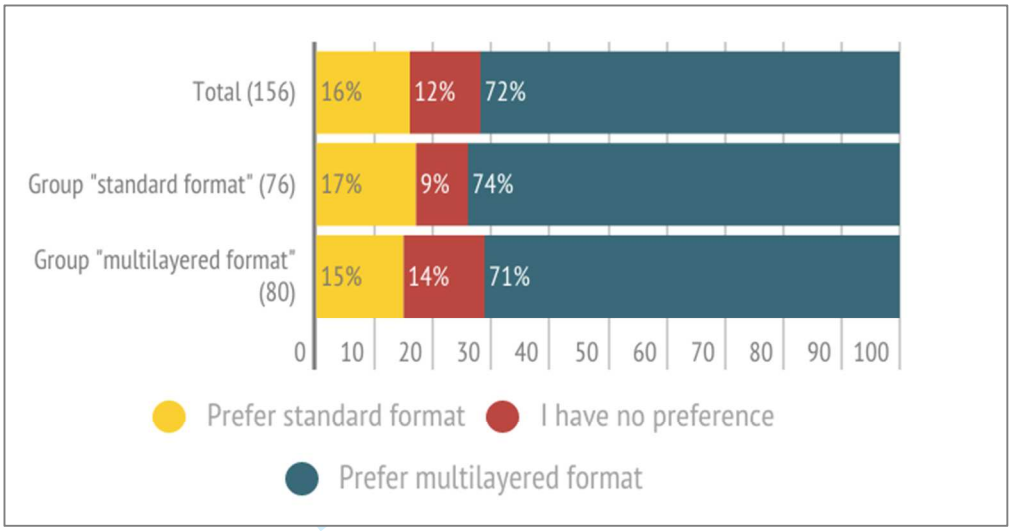


Figure 4: Preferences for standard format versus multilayered presentation format**.**

Understanding and anticipated clinical action

69 of 78 participants (89%) randomised to the standard format, and 57 of 64 participants (89%) randomised to the multilayered format correctly understood the evidence summaries, with no difference between groups (p-value = 0.91). A majority correctly understood the recommendations, with a trend towards favouring the multilayered formats (44/76 (58%) in the standard format, 55/76 (72%) in the multilayered format, p-value = 0.06).

Most participants correctly stated they would consider an oral anticoagulant as an appropriate course of treatment, with a majority preferring novel oral anticoagulants (NOACs) rather than warfarin. We observed a trend towards favouring the multilayered formats (standard format 66/72 (92%) versus multilayered format 79/81 (98%), p-value = 0.10).

Perceived usefulness of evidence summaries and recommendations

Of the 69 participants in the standard format group, 53 (77%) agreed (somewhat agree, agree or strongly agree) that the background evidence summaries were helpful in the context of the clinical scenario, as did 60 of 76 participants in the multilayered format group (79%), with no difference between the randomised groups (p-value = 0.76). 72 of 79 participants (91%)

randomised to the standard format and 74 of 86 participants (86%) to the multilayered format agreed that recommendations were helpful in the context of the clinical scenario (p-value = 0.31). When specifically asked, 84 of 102 (82%) participants considered absolute effect estimates provided in the multilayered format evidence summaries helpful or crucial.

Survey on conceptual understanding

Prior to randomisation, we provided the participants with a recommendation labelled with "2B" and asked them what the number 2 and the letter B meant. Few participants answered correctly that this represented a weak (2) recommendation based on moderate (B) quality evidence (20/154 (13%) stated weak, while 61/144 (42%) stated moderate). We furthermore twice asked to what extent they agreed to the following statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making." Prior to randomisation 63 of 158 participants (40%, 95% CI 32-48) stated that they agreed (agreed or strongly agreed) to this statement. After randomisation, we provided the participants with a one slide explanation of the strength of the recommendation according to the GRADE system, defining the difference between strong and weak recommendations and posed the same question again. 71 of 89 participants (80%, 95% CI 70-88) agreed with the statement. There was a borderline significant difference between participants according to randomised format (standard format 72% vs. multilayered format 88%, p = 0.051).

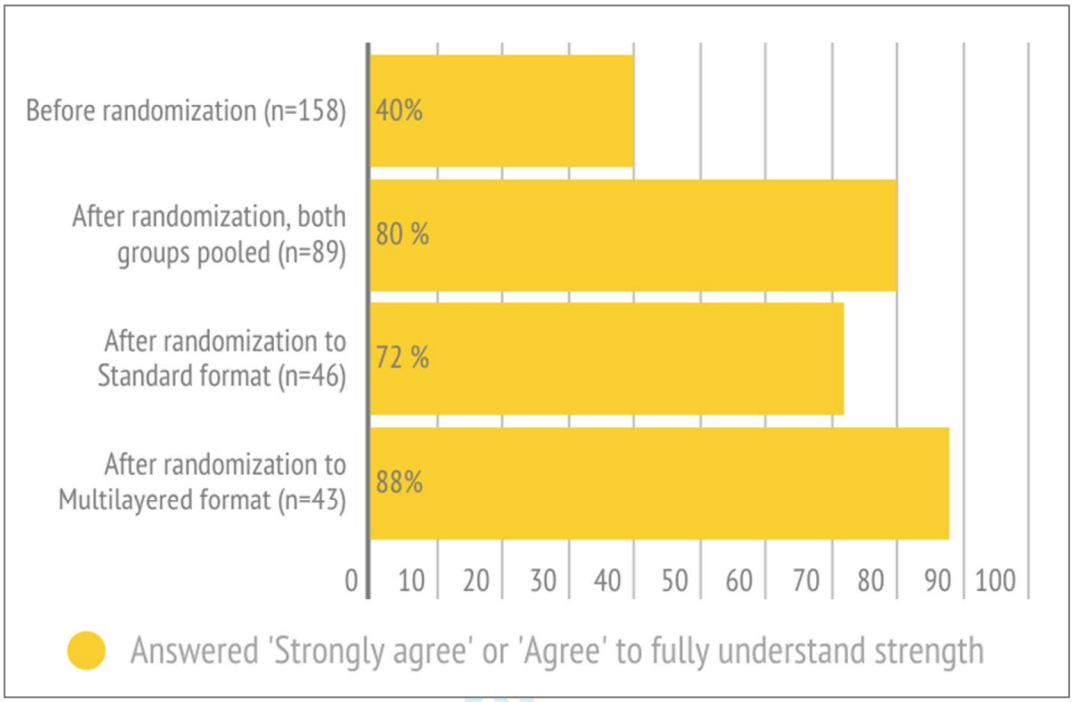


Figure 5: Reported answers to the statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making."

DISCUSSION

Practicing clinicians face an ever-growing amount of research evidence, evidence summaries and guidelines. The GRADE working group has advanced evidence-based practice, by providing comprehensive guidance on evidence synthesis, assessment and development of clinical practice guidelines. In 2011 the working group set out to systematically explore strategies to increase the uptake of guidelines at the point of care.

We devised a guideline presentation format tailored to the GRADE framework, aiming at facilitating a trustworthy guideline authoring process and enhancing dissemination. Results from usability testing uncovered challenges with ensuring an efficient communication to the end user clinician of guidelines without compromising key elements of information needed for optimal decision-making. We hypothesised that clinicians' apparent discomfort with more complex methodological concepts could be alleviated with proper education and an optimal user interface.

In this study of practising physicians we demonstrated a clear preference for the new guideline multilayered presentation format rather than a traditional narrative format, with a majority agreeing that recommendations and underlying evidence summaries - regardless of presentation format - were useful in the context of the clinical scenario. A majority of the physicians (82%) also reported absolute effect estimates provided in the multilayered format evidence summaries to be helpful or crucial for decision-making. Conceptual understanding of strength of recommendations and quality of evidence was limited prior to being presented with an explanation and when expressed through numbers and letters - as done in the standard format example from UpToDate. A short explanation of these concepts according to the GRADE system vastly increased reported understanding.

Strengths and weaknesses

We were able to test a substantial and diverse set of practising physicians across six different countries, targeting mandatory educational sessions, making our participant sample representative of the everyday clinician and our results more generalizable. We were able to map both apparent understanding of key methodological concepts and preference for entire guideline formats, as opposed to single elements.

There are certain limitations to the chosen study design. Firstly, devising multiple-choice questions that accurately test key outcomes such as understanding is challenging. Our approaches, including subjectively perceived understanding and anticipated clinical action are less satisfactory than detailed testing of understanding. We thus face the possibility that actual understanding is less than demonstrated within this study, or that the correct clinical choice of action is highly influenced by previous knowledge on the subject. A substantial proportion of physicians reported to "somewhat agree" rather than agree or strongly agree with statements concerning usefulness of recommendations and evidence summaries in clinical practice guidelines. This finding highlights some limitations of our study related to using a hypothetical

scenario - rather than assessing real life decision making - and the challenge of phrasing questions and devising appropriate response categories within this field of research.

Regardless of these limitations, we have demonstrated that the majority of physicians expressed a preference for having recommendations available, and that educating clinicians in key methodological concepts seems to be possible with little effort.

Implications for clinicians and policymakers

The finding that the majority of clinicians preferred to be informed about the absolute effects on patient-important outcomes in evidence summaries is encouraging, as numeracy among health care professionals is highly variable and many guidelines omit contextualised effect estimates necessary for shared decision making.[19–21] GRADE Summary of findings tables - as displayed in the multilayered guideline formats - have emerged as user-friendly and well accepted formats in the context of systematic reviews, with recent advances in formatting resulting in further increasing understanding.[22,23]

40% of the participants stated that they fully understood the difference between strong and weak recommendations, but only 13% could correctly recognize a weak recommendation applying the commonly used numbered labelling (1=strong, 2=weak). However, when given recommendation and evidence summary linked to a clinical scenario 95% of participants would treat according to current guidelines. There are thus two reasons for optimism: Most participants correctly interpreted the intent and meaning of the recommendations when communicated within a larger context. Furthermore, perceived understanding improved after being presented a one slide explanation on the meaning of key concepts.

We did not explore in detail why participants preferred the new DECIDE multilayered format. However, during informal discussions following the survey, and feedback throughout the design process, clinicians have expressed a need for short and clear advice, provision of strength that is easy to interpret and details around the key factors that drove the recommendations, which is

provided within the multilayered format.

Optimized guideline presentation formats and sufficient conceptual understanding, as researched in this trial, can potentially facilitate the uptake of trustworthy guidelines and application of research evidence in practice.[24,25] The multilayered format serves several purposes: It is devised around the GRADE framework and thus guides guideline authors through the appropriate methodological steps. It provides end-user clinicians with actionable, graded recommendations, as advocated by the Institute of Medicine.[26] Lastly, having the guideline digitally structured entails easy translation into several outputs; such as decision aids, clinical decision support systems within the electronic medical record, tablets, smartphones, online and for those with an analogous inclination - as PDFs. It also facilitates continuous updating and adaptation to local settings.[14] This can potentially minimize the workload for guideline developers. Further research into different ways of communicating key guideline concepts to health care professionals to improve understanding and adoption is still necessary.

CONCLUSION

The multilayered guideline presentation formats are currently implemented in a handful of published guidelines from multiple organisations (www.magicapp.org), and more are under development. Through ongoing research projects we will increase knowledge on the usability and experiences of the format, both from authors and end-users, informing continuing improvements.

CONTRIBUTORSHIP STATEMENT

LB, POV and AK had access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. LB was principal investigator on the study. LB, AK, POV, PAC, JT and AE contributed to the conception, design, and ethical approval of the study. LB, POV, EA, DR, KA and POC collected the data. LB, POV and AK contributed to writing the first draft of the article; and LB, POV, PAC, EA, JT, DR, KA, POC,

GG and AK contributed to editing and approval of the final manuscript.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form. LB, AK, GG and POV are members of a non-profit research and innovation project MAGIC: www.magicproject.org, which has an open technical platform where the new DECIDE multilayered formats were prototyped. LB, AK, GG, POV, PAC, EA, JT and DR are members of the Grade Working group. The strategy evaluated in the study is based on the GRADE approach.

No other competing interests were declared.

FUNDING

This work was supported by Sykehuset Innlandet Hospital Trust under grant number 150194 (PLUGGED IN), The Research Council of Norway (VERDIKT) under grant 193022 (EVICARE), and European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 258583 (DECIDE project).

DATA SHARING MANAGEMENT

Technical appendix and full dataset available from the Dryad repository, DOI: 10.5061/dryad.2qv30

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):427–433. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-8544219762&partnerID=40&md5=0a2c9df801f3d109f2c901d5206b98b7.
2. Kung J, Miller RR, Mackowiak P a. Failure of clinical practice guidelines to meet institute of medicine standards: Two more decades of little, if any, progress. *Arch Intern Med*. 2012;172(21):1628–33. doi:10.1001/2013.jamainternmed.56.
3. Alonso-Coello P, Irfan A, Solà I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58. doi:10.1136/qshc.2010.042077.
4. Institute of Medicine, Graham R, Mancher M, Wolman DM. *Clinical Practice Guidelines We Can Trust*. Washington, DC: The National Academies Press.; 2011. Available at: http://www.iom.edu/~/media/Files/Report Files/2011/Clinical-Practice-Guidelines-We-Can-Trust/Clinical Practice Guidelines 2011 Report Brief.pdf.
5. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med*. 2012;156(7):525–31. doi:10.1059/0003-4819-156-7-201204030-00009.
6. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. *Ann Intern Med*. 2011;154(11):774–5. doi:10.7326/0003-4819-154-11-201106070-00011.
7. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6. doi:10.1136/bmj.39489.470347.AD.
8. Goldet G, Howick J. Understanding GRADE : an introduction. *J Evid Based Med*. 2013;6:50–54. doi:10.1111/jebm.12018.
9. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: Time for a map? *J Contin Educ Health Prof*. 2006;26(1):13–24. doi:10.1002/chp.47.
10. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348(jun13 4):g3725–g3725. doi:10.1136/bmj.g3725.
11. Treweek S, Oxman AD, Alderson P, et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. *Implement Sci*. 2013;8:6. doi:10.1186/1748-5908-8-6.
12. Kristiansen A, Brandt L, Alonso-Coello P, et al. Development of a novel, multilayered presentation format for clinical practice guidelines. *Chest*. 2015;147(3):754–763. doi:10.1378/chest.14-1366.
13. Kristiansen A, Brandt L, Agoritsas T, et al. Applying new strategies for the national adaptation, updating and dissemination of trustworthy guidelines: Results from the Norwegian adaptation of the American College of Chest Physicians Evidence-based Guidelines on Antithrombotic Therapy and the Preven. *Chest*. 2014;146(3):735–61. doi:10.1378/chest.13-2993.
14. Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we

can trust, use, and share: a new era is imminent. *Chest*. 2013;144(2):381–9. doi:10.1378/chest.13-0746.

15. www.nsth.no. :Accessed February 2016.
16. 2013 AJ. http://www.uptodate.com/contents/atrial-fibrillation-anticoagulant-therapy-to-prevent-embolization?source=machineLearning&search=atrial+fibrillation+anticoagulation&selectedTitle=1~150&sectionRank=1&anchor=H31#H31.
17. Isaac T, Zheng J, Jha A. Use of UpToDate and outcomes in US hospitals. *J Hosp Med*. 2012;7(2):85–90. doi:10.1002/jhm.944.
18. Kwag KH, González-Lorenzo M, Banzi R, Bonovas S, Moja L. Providing Doctors With High-Quality Information: An Updated Evaluation of Web-Based Point-of-Care Information Summaries. *J Med Internet Res*. 2016;18(1):e15. doi:10.2196/jmir.5234.
19. Manrai AK, Bhatia G, Strymish J, Kohane IS, Jain SH. Medicine's UncomfortableRelationship With Math: CalculatingPositive PredictiveValue. *JAMA Intern Med*. 2014;174(6):991–993. doi:10.1001/jamainternmed.2014.1001.Author.
20. Johnston B, Alonso-Coello P, Friedrich JO, Reem MA. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ*. 2015;(November):1–8. doi:10.1503/cmaj.150430.
21. Agoritsas T, Heen AF, Brandt L, et al. Decision aids that really promote shared decision making: the pace quickens. *BMJ*. 2015;350(Feb10):g7624. doi:10.1136/bmj.g7624.
22. 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(6):620–6. doi:10.1016/j.jclinepi.2009.12.014.
23. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, 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;(In press). doi:10.1016/j.jclinepi.2015.12.007.
24. Hsiao J, Chen R. Critical factors influencing physicians ' intention to use computerized clinical practice guidelines : an integrative model of activity theory and the technology acceptance model. *BMC Med Inform Decis Mak*. 2016:1–15. doi:10.1186/s12911-016-0241-3.
25. Cuello García C a, Pacheco Alvarado KP, Pérez Gaxiola G. Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems. *Arch Dis Child*. 2011;96(8):723–728. doi:10.1136/adc.2010.199307.
26. Neumann I, Santesso N, Akl EA, et al. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol*. 2015;(c). doi:10.1016/j.jclinepi.2015.11.017.

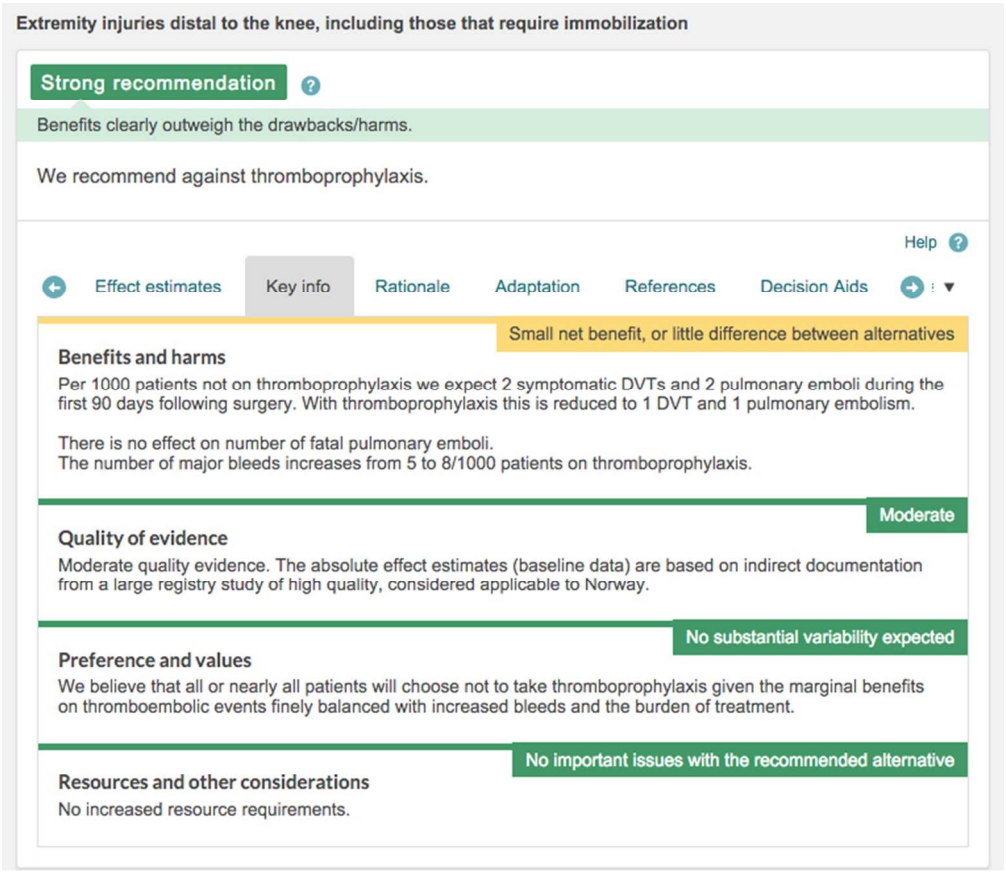


Figure 1: Current version of the multilayered guideline presentation formats.
Figure 1
268x232mm (72 x 72 DPI)

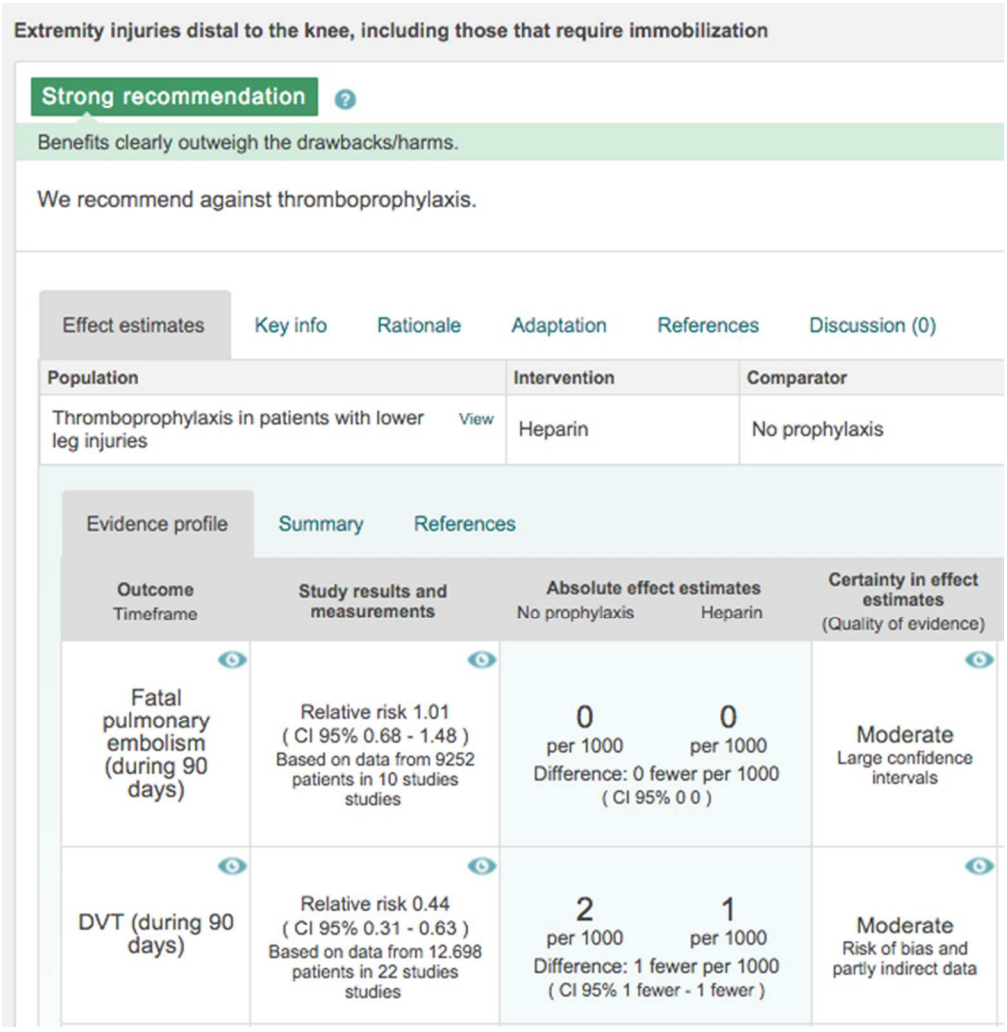


Figure 1: Current version of the multilayered guideline presentation formats.
Figure 1
215x219mm (72 x 72 DPI)

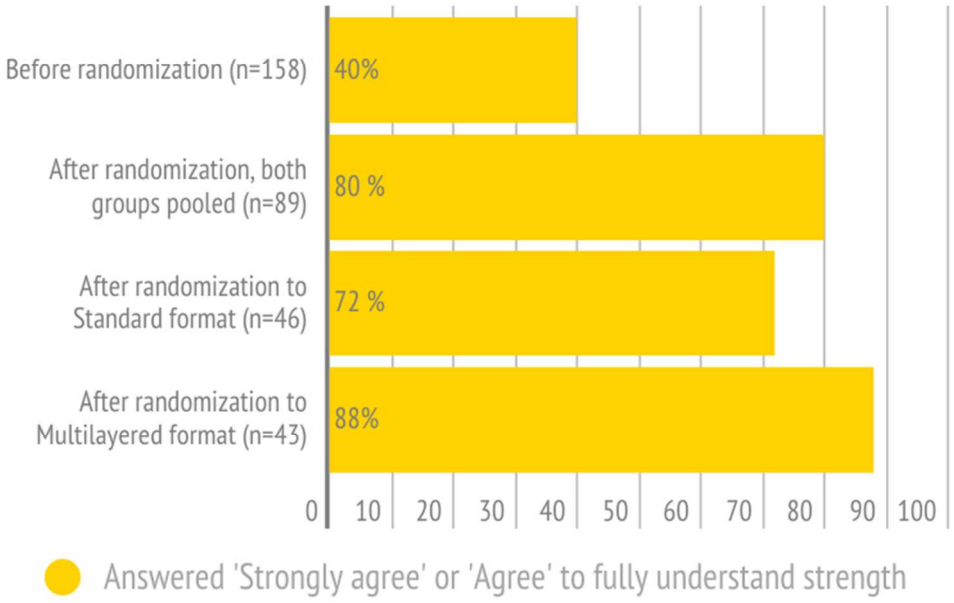


Figure 5: Reported answers to the statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making."
Figure 5
274x178mm (72 x 72 DPI)

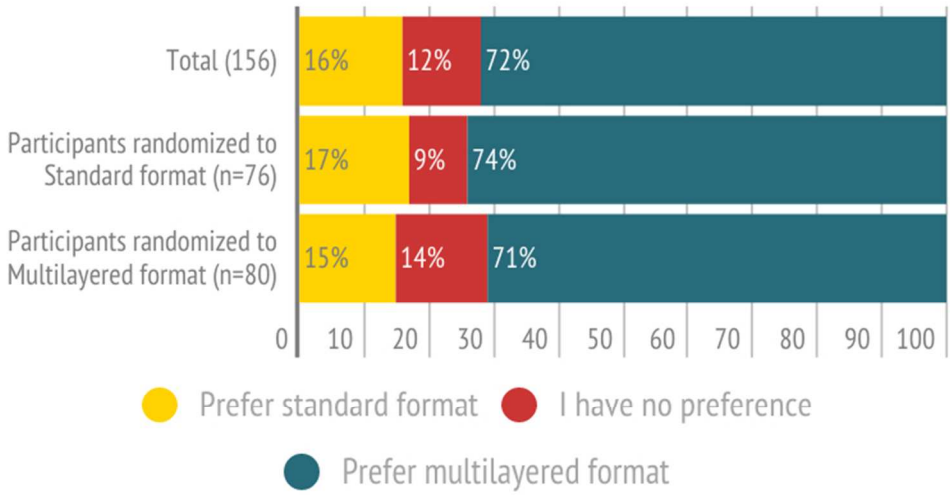


Figure 4: Preferences for standard format versus multilayered presentation format.
Figure 4
276x145mm (72 x 72 DPI)

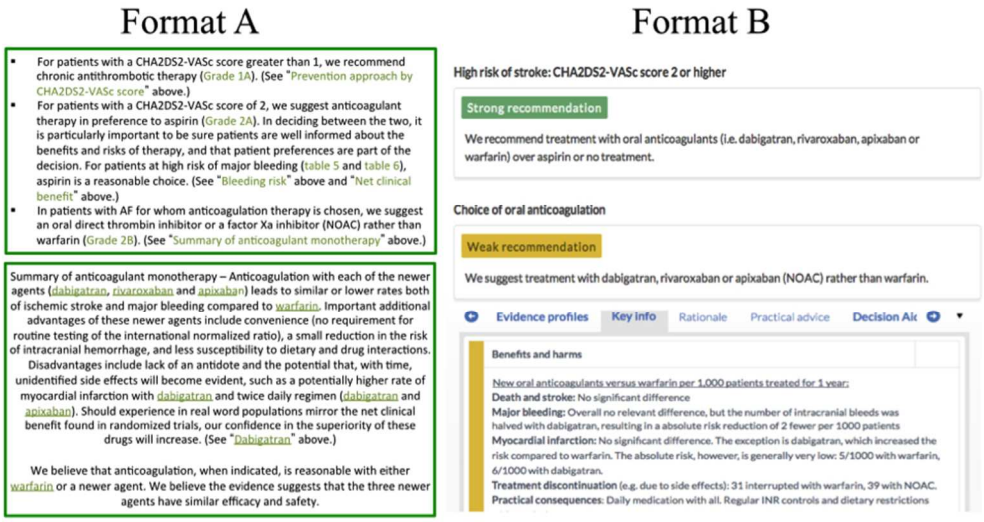
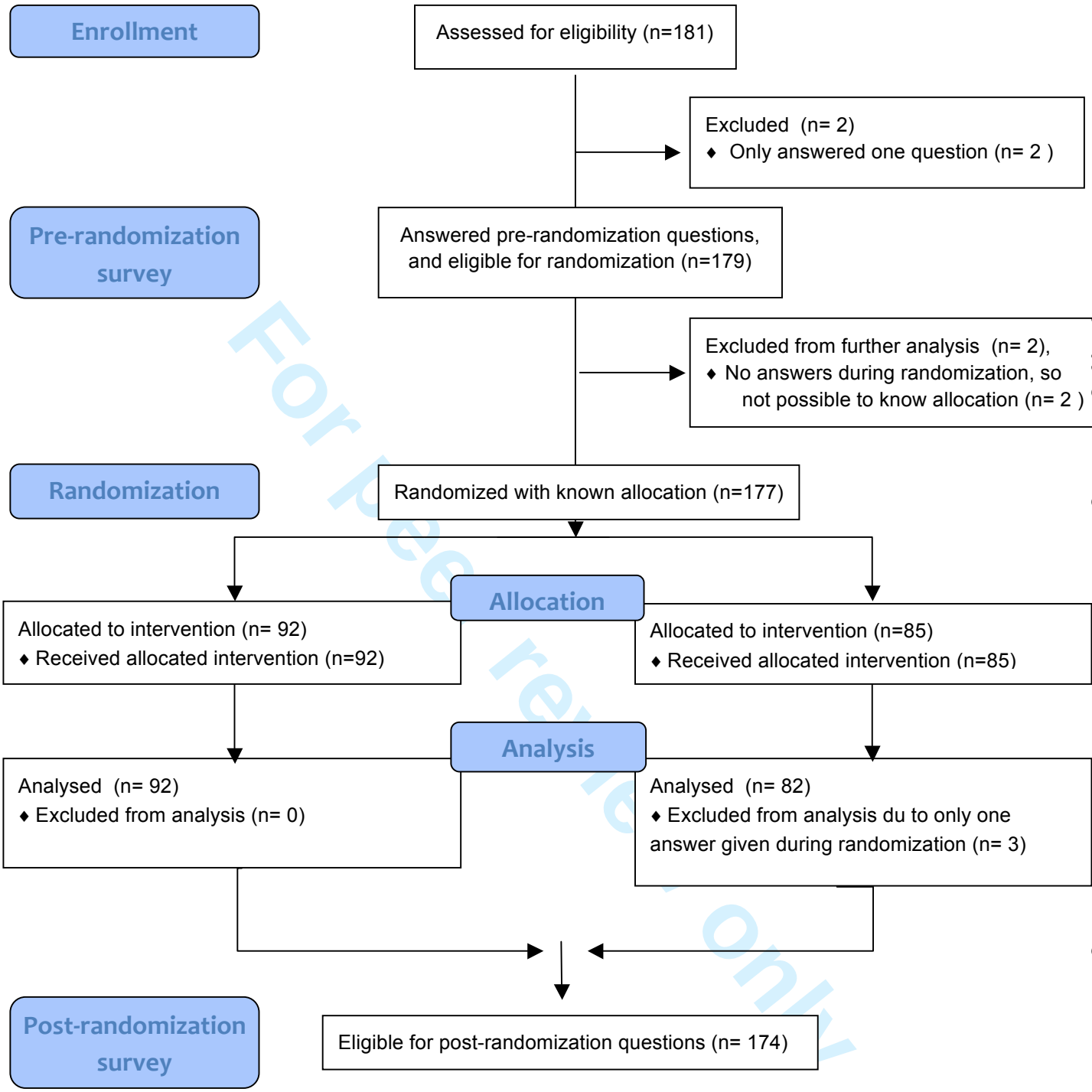


Figure 2: Standard (A) and multilayered guideline (B) presentation formats.
Figure 2



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4
	2b	Specific objectives or hypotheses	4-5
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	5-9
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	9-10
Participants	4a	Eligibility criteria for participants	5
	4b	Settings and locations where the data were collected	9
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	5
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8-9
	6b	Any changes to trial outcomes after the trial commenced, with reasons	-
Sample size	7a	How sample size was determined	protocol
	7b	When applicable, explanation of any interim analyses and stopping guidelines	-
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	8
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	8
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	8

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	6-7
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	9
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	-
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	10
	13b	For each group, losses and exclusions after randomisation, together with reasons	10
Recruitment	14a	Dates defining the periods of recruitment and follow-up	5
	14b	Why the trial ended or was stopped	-
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	9
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	9-13
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	9-13
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	9-13
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	9-13
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	-
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	14-16
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	14-16
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	14-16
Other information			
Registration	23	Registration number and name of trial registry	-
Protocol	24	Where the full trial protocol can be accessed, if available	3
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	3

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

BMJ Open

Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011569.R1
Article Type:	Research
Date Submitted by the Author:	06-Sep-2016
Complete List of Authors:	Brandt, Linn; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Vandvik, Per; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Alonso-Coello, Pablo; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Akl, Elie; American University of Beirut, Department of Internal Medicine Thornton, Judith; National Institute for Health and Care Excellence Rigau, David ; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau Adams, Katie; National Institute for Health and Care Excellence O'Connor, Paul; National Institute for Health and Care Excellence Guyatt, Gordon; Mcmaster University, Department of Clinical Epidemiology and Biostatistics Kristiansen, Annette; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Health informatics, Epidemiology, Communication
Keywords:	Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, EPIDEMIOLOGY

Target journal BMJ Open

Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial

Linn Brandt[1,2], Per Olav Vandvik Pablo[1,2], Pablo Alonso-Coello[3,6], Elie A. Akl[4,6], Judith Thornton[5], David Rigau[3], Katie Adams[5], Paul O'Connor[5], Gordon Guyatt[6], and Annette Kristiansen[1,2].

Author Affiliations: [1)] Department of Internal Medicine, Sykehuset Innlandet Hospital Trust, Gjøvik, Norway (Drs. Brandt, Kristiansen and Vandvik), [2)]Institute of Health and Society, Faculty of Medicine, University of Oslo (Drs. Brandt, Kristiansen and Vandvik), [3)] Iberoamerican Cochrane Centre, Biomedical Research Institute (IIB Sant Pau-CIBERESP), Barcelona, Spain (Drs. Alonso-Coello and Rigau), [4)] Department of Internal Medicine, American University of Beirut, Lebanon (Dr. Akl), [5)] National Institute for Health and Care Excellence, Manchester UK (Dr. Judith Thornton, Katie Adams, Paul O'Connor), [6)] Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (Drs. Alonso-Coello, Akl and Guyatt).

Corresponding Author: Linn Brandt, Department of Internal Medicine, Sykehuset Innlandet Gjøvik, Kyrre Grepps gate 11, 2819 Gjøvik, Norway. E-mail: brandtlinn@gmail.com

Designers: Rob Fracisco, Frankie Achille and Sarah Rosenbaum helped designing the novel multilayered format.

Abstract

Objectives: To investigate practicing physicians preferences, perceived usefulness and understanding of a new multilayered guideline presentation format - compared to a standard format - as well as conceptual understanding of trustworthy guideline concepts.

Design: Participants attended a standardized lecture in which they were presented with a clinical scenario and randomised to view a guideline recommendation in a multilayered format or standard format after which they answered multiple-choice questions using clickers. Both groups were also presented and asked about guideline concepts.

Setting: Mandatory educational lectures in seven non-academic and academic hospitals, and two settings involving primary care in Lebanon, Norway, Spain and United Kingdom.

Participants: 181 practicing physicians in internal medicine (156) and general practice (25).

Interventions: A new digitally structured, multilayered guideline presentation format and a standard narrative presentation format currently in widespread use.

Primary and secondary outcome measures: Our primary outcome was preference for presentation format. Understanding, perceived usefulness and perception of absolute effects were secondary outcomes.

Results: 72% (95% CI 65-79) of participants preferred the multilayered format and 16% (95% CI 10-22) preferred the standard format. A majority agreed that recommendations (multilayered 86% vs. standard 91%, p-value = 0.31) and evidence summaries (79% vs. 77%, p-value = 0.76) were useful in the context of the clinical scenario. 72% of participants randomised to the multilayered format vs. 58% for standard formats reported correct understanding of the recommendations (p-value = 0.06). Most participants elected an appropriate clinical action after

viewing the recommendations (98% vs. 92%, p-value = 0.10). 82% of the participants considered absolute effect estimates in evidence summaries helpful or crucial.

Conclusions: Clinicians clearly preferred a novel multilayered presentation format to the standard format. Whether the preferred format improves decision-making and has an impact on patient important outcomes merits further investigation.

Trial registration: None.

Strengths and limitations of this study

- We conducted a multicentre trial targeting regular educational sessions.
- Both formats were taken from published guidelines, and we used a comparator format that most participants were familiar with.
- To avoid peer pressure, we ensured participants anonymity through use of clickers (audience response technology).
- A weakness of the study is having researchers involved in development of the new presentation format perform most of the educational sessions.
- We did not measure impact of alternative formats on clinical decisions or patient important outcomes.

BACKGROUND

Clinical practice guidelines that provide recommendations addressing diagnosis and treatment can help clinicians optimize their evidence-based practice (EBP) at the point of care.[1] An abundance of guidelines are available, but many have shortcomings with their trustworthiness and dissemination strategies.[2,3] New standards for trustworthy guidelines developed by the Institute of Medicine and the Guideline International Network highlight the need for more rigorous development processes.[4–6]

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (www.gradeworkinggroup.org) represents a systematic, explicit and transparent process for evaluating and reporting quality of research evidence and for moving from evidence to recommendations.[7,8] GRADE facilitates the creation of trustworthy guidelines and has been adopted by more than 100 organisations worldwide.

Implementation of guidelines requires effective dissemination of recommendations.[9] Guidelines should generally answer clinicians' informational needs within two minutes, which implies that recommendations need to be easy to find, understand, apply and share.[10] Traditionally, guidelines have often been distributed as comprehensive PDFs, impeding efficient use at the point of care. With these challenges in mind, the GRADE working group initiated the Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE) project.[11] DECIDE aimed to improve dissemination of evidence-based recommendations for a range of stakeholders, including health care professionals, policy makers and patients, as well as to ensure and facilitate adherence to trustworthy guideline standards.[4–6]

As detailed in articles by Treweek et al.[11] and Kristiansen et al.,[12] a multidisciplinary group of clinicians, guideline developers, methodologists and graphical designers developed a multilayered guideline presentation format (figure 1a and 1b) targeted at health care professionals. We hypothesized that clinicians' apparent discomfort with more complex

methodological concepts could be alleviated with proper education and an optimal user interface. Based on the groups´ extensive experience from clinical practice and guideline development and informed by a narrative review of guideline formats, we designed a prototype presentation format through brainstorming sessions. This prototype format was iteratively improved based on results from stakeholder feedback and usability testing with clinicians. [12]

During usability testing we confirmed previous research demonstrating limitations in clinicians' conceptual understanding of key standards for trustworthy guidelines. We have since deployed the multilayered format in real-life guidelines.[13] Uncertain of the relative merits of our novel versus existing formats, we conducted a combined survey and randomised controlled trial to determine clinicians' preferences for the new multilayered presentation format versus a traditional format for guidelines, the perceived usefulness of guideline recommendations and understanding of key concepts of trustworthy guidelines.

Figure 1a and 1b: Current version of the multilayered guideline presentation formats.

MATERIALS AND METHODS

This study is the result of a collaboration of DECIDE and MAGIC (Making GRADE the Irresistible choice), a non-profit innovation and research program (www.magicproject.org), which aims at facilitating the efficient creation, dissemination and dynamic updating of trustworthy guidelines and evidence summaries. As previously reported, MAGIC has created a web-based guideline authoring and publication platform (www.magicapp.org), incorporating the digitally structured multilayered presentation format used in this study.[14] Through DECIDE these novel formats have also been incorporated in GRADEpro (http://gradepro.org/). Organizations can use these platforms or freely adopt research outputs from the DECIDE project, including but not limited to the multilayered presentation format, into their own workflow, tools or platforms.

Study design, setting and participants

We applied a combined survey and randomised controlled trial. We included practicing

physicians in internal or family medicine. In order to recruit representative samples of internal medicine physicians, investigators targeted compulsory educational sessions at teaching hospitals. Two facilities recruited general practitioners in family medicine. One by targeting a compulsory educational session at a larger family practice centre (Spain), the other by inviting individual physicians, including general practitioners, to a specific CLICK-IT study session (NICE, UK). All participants attended a standardized educational session on key guideline standards performed within the context of this study. We performed four pilot sessions and made revisions to the survey questions based on experiences in these sessions. The revisions were minor and concerned mainly phrasing of the questions asked. Our reported primary and secondary outcomes are in accordance with the original study protocol.

Participants were considered to provide consent by accepting to answer the survey questions. Research ethics boards in each participating country approved the study.

Standardized lecture and study procedure

At each site, a member of the research team delivered a standardized lecture according to a predefined protocol. The lecture was titled "New standards for trustworthy guidelines" (appendix 1) and was given in power point on a standard projector screen. Participants provided anonymous answers to questions with predefined response categories using clickers. The questions (both for the survey and the randomised part of the trial) - as well as screenshots of presentation formats - were embedded in the lecture slides using an audience response software, TurningPoint©, and read out loud by the presenter. For the Norwegian and Spanish sites we translated the presentation to their native language; while in Lebanon and the UK the questions were presented in English, which is commonly used in medical education.

The lecture and study procedure were developed by the investigators and included the following components:

1. Collection of the demographics of the participants, their current preference for information resources and understanding of the GRADE system.

2. Presentation of a clinical scenario concerning choice of oral anticoagulation in a patient with atrial fibrillation and high risk of stroke (CHA2DS2-VASc score 2).
3. Presentation of guideline recommendations and evidence summaries relevant to the clinical scenario, sequentially presented in the two alternative guideline presentation formats through randomisation and blinding as outlined below. Both formats were shown side by side to both groups at the end.
4. Short presentation of key conceptual definitions of trustworthy guidelines, specifically explanation of the strength of recommendations and quality of the evidence using the GRADE methodology.

Guideline presentation formats

Figure 2 shows the standard narrative guideline presentation format (format A) and the new multilayered guideline presentation format (format B). We extracted both formats from existing guidelines, but masked the publishers' identity to avoid potentially biasing participants' responses. Both formats were displayed as screenshots in the presentation slides.

Multilayered format: The experimental multilayered format displays recommendations up front with supporting information as collapsible boxes provided by clicking on the recommendation itself. The strength of the recommendation is communicated by use of text and colour coding, and a header describes the population for which the recommendation applies. The example was taken from a Norwegian guideline for antithrombotic therapy applying the multilayered presentation formats published in MAGICapp and translated to English for the purpose of this study.[13,15]

Standard format: The control standard format displays recommendations and an abridged evidence summary from UpToDate.[16] We considered UpToDate's presentations a suitable reference standard for current presentation formats due to its widespread use, commonality with other guidelines, evidence-based approach and use of GRADE.[17,18] UpToDate provides a textual summary of its recommendations in bullet points with links to the supporting information in the

main text or in other articles. The recommendations are labelled with numbers (1 & 2) and letters (A-C), depicting the strength of the recommendation and quality of evidence respectively.

Figure 2: Standard and multilayered guideline presentation formats.

Randomisation and blinding

A research member colour marked the base of half the clickers and haphazardly rearranged them in their container. The presenter handed out the clickers to the participants front-up to conceal the marking. Participants with marked clickers were randomised to group B, being presented the multilayered format. We asked participants in either group to put on blindfolds when the other group were shown their allocated presentation format. Questions during randomisation were not read out loud.

Outcome measures

Our primary outcome was preference for either presentation format. We provided three response options: preference for the standard format, no preference and preference for the multilayered format.

Secondary outcomes included:

- Correct understanding/interpretation of 1) the evidence summaries and 2) the recommendation with four potential answers, one alternative being correct.
- Anticipated course of action to the clinical scenario with four potential answers, two alternatives being correct.
- Participants' perceived usefulness of 1) evidence summaries and 2) recommendations. Participants provided answers to the statement "This information/recommendation would help me manage my patient" on a 6 point Likert scale: 1= strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree and 6=strongly agree.

- Correct understanding of the strength of the following example recommendation and the confidence in effect estimates: "We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol). GRADE 2B." We provided participants with four potential answers, one alternative being correct.
- Perceived understanding of the strength of the recommendation. We asked the participants the following question twice, before and after being provided with a short written explanation: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision-making." They provided answers on a 6 point Likert scale with 2 anchors: 1= strongly disagree and 6 = strongly agree.
- Participants' perception of presenting absolute effect estimates. We asked participants "What is your first reaction to being presented with absolute effects?" The answers were collected using a 5-point scale: 1=confusing distraction, leave it out, 2=a little confusing, but not a big problem, 3=doesn't help, but doesn't hurt, 4=not crucial, but helpful and 5=crucial information, should always be included.

Statistical analysis

We dichotomized all outcomes to either correct/incorrect or agree/disagree, and analysed them by use of Pearson's Chi-square test. We included all randomised participants that answered more than one question in the final analysis (Intention To Treat). No subgroup analyses were specified in the protocol. The accompanying data set (Excel and original Turning Point data files from each centre) provides results subdivided per type of physician and centre. We used SPSS (version 23) for all analyses.

RESULTS

We performed the study from June 2013 until January 2015. We included 181 practicing physicians across four countries and nine centres (Lebanon 1, Norway 6, Spain 1 and UK 1), 177 were randomised and 174 (96%) answered > 1 question and are included in the final analysis (figure 3). Their demographics and information resource preferences are provided in table 1. The

two groups were fairly similar. One centre (27 participants) in Norway did not include demographic questions, so the demographic presentation includes the remaining eight centres.

Number of participants randomised	**Multilayered format** 92	**Standard format** 85
Country		
(# participants eligible for analysis)	(92)	(83)
Norway (%)	61 (66.3%)	57 (68.7%)
UK (%)	10 (10.9%)	10 (12%)
Lebanon (%)	11 (12.0%)	10 (13.3%)
Spain (%)	10 (10.9%)	5 (6%)
Professional status or Specialty		
(# participants eligible for analysis)	(76)	(72)
Medical student or intern (%)	13 (17.1%)	8 (11.1%)
Internist Resident (%)	21 (27.6%)	27 (37.5%)
Internist Attending/Consultant (%)	23 (30.3%)	21 (29.2)
General practitioner (%)	13 (17.1%)	12 (16.7%)
Unknown (% did not answer that question)	6 (7.9%)	4 (5.6%)
Training in health research methodology		
(# participants eligible for analysis)	(72)	(68)
No training in HRM (%)	34 (47.2%)	35 (51.2%)
≥ 1 HRM course (%)	26 (36.1%)	21 (30.9%)
Degree in HRM (%)	12 (16.7%)	12 (17.6)
Preferred knowledge source		
(# participants eligible for analysis)	(89)	(82)
Local guideline (%)	22 (24.7%)	14 (17.1%)
Systematic review (%)	2 (2.2%)	2 (2.4%)
EBM textbook (%)	17 (19.1%)	13 (15.9%)
National or international guideline (%)	34 (38.2%)	36 (43.9%)
Colleague (%)	14 (15.7%)	17 (20.7%)
Primary study (%)	0 (0%)	0 (0%)

Table 1: Demographics of the groups randomised to different formats. HRM=Health Research Methodology. EBM=Evidence Based Medicine.

Figure 3: Flow chart of design and enrolment of participants to the multilayered format (n=92) and standard format (n=85)

Preference for alternative presentation formats

When exposed to both formats after completing the randomised part of the study 113 of 156 (72%, 95% CI 65-79) participants preferred the new multilayered format, 25 (16%, 95% CI 10-22) preferred the standard format, and 18 (12%, 95% CI 7-17) reported no preference. Results were similar in those randomised to the standard vs. the multilayered format (p = 0.66, figure 4).

Figure 4: Preferences for standard format versus multilayered presentation format**.**

Understanding and anticipated clinical action

69 of 78 participants (88%, 95% CI 81-96) randomised to the standard format, and 57 of 64 participants (89%, 95% CI 82-97) randomised to the multilayered format correctly understood the evidence summaries, with no difference between groups (p-value = 0.91). A majority correctly understood the recommendations, with no statistically significant difference between the standard format (44/ 76, 58%) and the multilayered format (55/ 76, 72%) (p-value = 0.06).

Most participants correctly stated they would consider an oral anticoagulant as an appropriate course of treatment, with a majority preferring direct acting oral anticoagulants (DOACs) rather than warfarin. We observed no statistically significant difference between the presentation formats (standard format 66/72 (92%, 95% CI 85-98) versus the multilayered format 79/81 (98%, 95% CI 94-100), p-value = 0.10).

Perceived usefulness of evidence summaries and recommendations

Of the 69 participants in the standard format group, 53 (77%, 95% CI 67-87) agreed (somewhat agree, agree or strongly agree) that the background evidence summaries were helpful in the context of the clinical scenario, as did 60 of 76 participants in the multilayered format group (79%, 95% CI 70-88), with no difference between the randomised groups (p-value = 0.76). 72 of 79 participants (91%, 95% CI 85-97) randomised to the standard format and 74 of 86 participants (86%, 95% CI 79-93) to the multilayered format agreed that recommendations were helpful in the context of the clinical scenario (p-value = 0.31). When specifically asked, 84 of 102 (82%,

95% CI 75-90) participants considered absolute effect estimates provided in the multilayered format evidence summaries helpful or crucial.

Survey on conceptual understanding

Prior to randomisation, we provided the participants with a recommendation labelled with "2B" and asked them what the number 2 and the letter B meant. Few participants answered correctly that this represented a weak (2) recommendation based on moderate (B) quality evidence (20/154 (13%, 95% CI 8-18) stated weak, while 61/144 (42%, 95% CI 34-50) stated moderate). We furthermore twice asked to what extent they agreed to the following statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making." Prior to randomisation 63 of 158 participants (40%, 95% CI 32-48, figure 5) stated that they agreed (agreed or strongly agreed) to this statement. After randomisation, we provided the participants with a one slide explanation of the strength of the recommendation according to the GRADE system, defining the difference between strong and weak recommendations and posed the same question again. 71 of 89 participants (80%, 95% CI 70-88, figure 5) agreed with the statement. There was a borderline significant difference between participants according to randomised format (standard format 72% vs. multilayered format 88%, p = 0.051, figure 5).

Figure 5: Reported answers to the statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making."

DISCUSSION

In this study of practicing physicians, we demonstrated a clear preference for the new guideline multilayered presentation format rather than a traditional narrative format, with a majority agreeing that recommendations and underlying evidence summaries - regardless of presentation format - were useful in the context of the clinical scenario. A majority of the physicians (82%) also reported absolute effect estimates provided in the multilayered format evidence summaries

to be helpful or crucial for decision-making. Conceptual understanding of strength of recommendations and quality of evidence was limited when expressed through numbers and letters - as done in the standard format example from UpToDate. A short explanation of these concepts according to the GRADE system substantially increased reported understanding.

Strengths and weaknesses

We enlisted a substantial and diverse set of practising physicians across six different countries, targeting mandatory educational sessions, making our participant sample representative of the everyday clinician and our results generalizable. We mapped understanding of both key methodological concepts and preference for entire guideline formats, as opposed to single elements. We devised a clinical scenario with a highly prevalent disease (atrial fibrillation), in which there have been recent treatment innovations (DOACs), possibly not commonly known to all physicians when this study was conducted. Through this manoeuvre, we limited the possible bias of previous knowledge on the participants' performance. We targeted internists and family physicians as participants given their familiarity with atrial fibrillation.

There are limitations to the chosen study design. Firstly, devising multiple-choice questions that accurately test key outcomes such as understanding is challenging. Our approaches, including subjectively perceived understanding and anticipated clinical action are less satisfactory than detailed testing of understanding. Actual understanding may be less than our results suggest, and the correct clinical choice of action may have been highly influenced by previous knowledge on the subject. A substantial proportion of physicians reported to "somewhat agree" rather than agree or strongly agree with statements concerning usefulness of recommendations and evidence summaries in clinical practice guidelines. This finding highlights some limitations of our study related to using a hypothetical scenario - rather than assessing real life decision making - and the challenge of phrasing questions and devising appropriate response categories within this field of research. Lastly, as we have only tested one specific clinical scenario on a limited representation of health care professionals, the transferability of our findings to other settings and professional groups needs further validation.

Implications for practice and research

The finding that the majority of clinicians preferred to be informed about the absolute effects on patient-important outcomes in evidence summaries is encouraging, as numeracy among health care professionals is highly variable and many guidelines omit contextualised effect estimates necessary for shared decision making.[19–21] GRADE Summary of findings tables - as displayed in the multilayered guideline formats - have emerged as user-friendly and well accepted formats in the context of systematic reviews, with recent advances in formatting resulting in further increasing understanding.[22,23]

40% of the participants stated that they fully understood the difference between strong and weak recommendations, but only 13% could correctly recognize a weak recommendation applying the commonly used numbered labelling (1=strong, 2=weak). However, when given recommendation and evidence summary linked to a clinical scenario 95% of participants would treat according to current guidelines. There are thus two reasons for optimism: Most participants correctly interpreted the intent and meaning of the recommendations when communicated within a larger context. Furthermore, perceived understanding improved after being presented a one slide explanation on the meaning of key concepts.

We did not explore in detail why participants preferred the new DECIDE multilayered format. However, informed by feedback throughout the design process from stakeholders as well as informal discussions following the survey, clinicians seem to appreciate short and clear advice, provision of strength that is easy to interpret and details around the key factors that drove the recommendations, which is provided within the multilayered format.

Optimized guideline presentation formats and sufficient conceptual understanding can potentially facilitate the uptake of trustworthy guidelines and application of research evidence in practice.[24,25] The multilayered format serves several purposes: It is devised around the GRADE framework and thus directs guideline authors through the appropriate methodological steps. It

provides end-user clinicians with actionable, graded recommendations, as advocated by the Institute of Medicine.[26] Lastly, having the guideline digitally structured facilitates easy translation into several outputs; such as decision aids, clinical decision support systems within the electronic medical record, tablets, smartphones, online and as PDFs. It also facilitates continuous updating and adaptation to local settings,[14] thereby minimizing the workload for guideline developers and policy makers.

Further research into different ways of communicating key guideline concepts to health care professionals to improve understanding and adoption is still necessary. The multilayered guideline presentation formats are currently implemented in a handful of published guidelines from a variety of organizations, and more are under development. We are continuously performing usability testing, both with authors and end-users, informing further improvements.

CONCLUSION

Clinicians prefer a novel multilayered presentation format to the standard format. Optimized guideline presentation formats and sufficient conceptual understanding can potentially facilitate the uptake of trustworthy guidelines and application of research evidence in practice. Whether the preferred format improves decision-making and has an impact on patient-important outcomes merits further investigation.

CONTRIBUTORSHIP STATEMENT

LB, POV, EA, DR, KA and POC collected the data. LB, POV and AK had access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. LB was principal investigator on the study. LB, AK, POV, PAC and AE contributed to the conception, design, and ethical approval of the study. LB, POV and AK contributed to writing the first draft of the article; and LB, POV, PAC, EA, JT, DR, KA, POC, GG and AK contributed to editing and approval of the final manuscript.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form. LB, AK, GG and POV are members of a non-profit research and innovation project MAGIC: www.magicproject.org, which has an open technical platform where the new DECIDE multilayered formats were prototyped. All authors were either co-investigators or collaborators of the DECIDE project. LB, AK, GG, POV, PAC, EA, JT and DR are members of the Grade Working group. The strategy evaluated in the study is based on the GRADE approach. No other competing interests were declared.

FUNDING

This work was supported by Sykehuset Innlandet Hospital Trust under grant number 150194 (PLUGGED IN), The Research Council of Norway (VERDIKT) under grant 193022 (EVICARE), and European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 258583 (DECIDE project)

.

DATA SHARING MANAGEMENT

Technical appendix and full dataset available from the Dryad repository, DOI: 10.5061/dryad.2qv30

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):427–433.
2. Kung J, Miller RR, Mackowiak P a. Failure of clinical practice guidelines to meet institute of medicine standards: Two more decades of little, if any, progress. *Arch Intern Med*. 2012;172(21):1628–33. doi:10.1001/2013.jamainternmed.56.
3. Alonso-Coello P, Irfan A, Solà I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58. doi:10.1136/qshc.2010.042077.
4. Institute of Medicine, Graham R, Mancher M, Wolman DM. *Clinical Practice Guidelines We Can Trust*. Washington, DC: The National Academies Press.; 2011.
5. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med*. 2012;156(7):525–31. doi:10.1059/0003-4819-156-7-201204030-00009.
6. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. *Ann Intern Med*. 2011;154(11):774–5. doi:10.7326/0003-4819-154-11-201106070-00011.
7. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6. doi:10.1136/bmj.39489.470347.AD.
8. Goldet G, Howick J. Understanding GRADE : an introduction. *J Evid Based Med*. 2013;6:50–54. doi:10.1111/jebm.12018.
9. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: Time for a map? *J Contin Educ Health Prof*. 2006;26(1):13–24. doi:10.1002/chp.47.
10. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348(jun13 4):g3725–g3725. doi:10.1136/bmj.g3725.
11. Treweek S, Oxman AD, Alderson P, et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. *Implement Sci*. 2013;8:6. doi:10.1186/1748-5908-8-6.
12. Kristiansen A, Brandt L, Alonso-Coello P, et al. Development of a novel, multilayered presentation format for clinical practice guidelines. *Chest*. 2015;147(3):754–763. doi:10.1378/chest.14-1366.
13. Kristiansen A, Brandt L, Agoritsas T, et al. Applying new strategies for the national adaptation, updating and dissemination of trustworthy guidelines: Results from the Norwegian adaptation of the American College of Chest Physicians Evidence-based Guidelines on Antithrombotic Therapy and the Preven. *Chest*. 2014;146(3):735–61. doi:10.1378/chest.13-2993.
14. Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. *Chest*. 2013;144(2):381–9. doi:10.1378/chest.13-0746.
15. www.nsth.no. :Accessed February 2016.
16. 2013 AJ. http://www.uptodate.com/contents/atrial-fibrillation-anticoagulant-therapy-to-

prevent-embolization?source=machineLearning&search=atrial+fibrillation+anticoagulation&selectedTitle=1~150&sectionRank=1&anchor=H31#H31.
17. Isaac T, Zheng J, Jha A. Use of UpToDate and outcomes in US hospitals. *J Hosp Med*. 2012;7(2):85–90. doi:10.1002/jhm.944.
18. Kwag KH, González-Lorenzo M, Banzi R, Bonovas S, Moja L. Providing Doctors With High-Quality Information: An Updated Evaluation of Web-Based Point-of-Care Information Summaries. *J Med Internet Res*. 2016;18(1):e15. doi:10.2196/jmir.5234.
19. Manrai AK, Bhatia G, Strymish J, Kohane IS, Jain SH. Medicine's UncomfortableRelationship With Math: CalculatingPositive PredictiveValue. *JAMA Intern Med*. 2014;174(6):991–993. doi:10.1001/jamainternmed.2014.1001.Author.
20. Johnston B, Alonso-Coello P, Friedrich JO, Reem MA. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ*. 2015;(November):1–8. doi:10.1503/cmaj.150430.
21. Agoritsas T, Heen AF, Brandt L, et al. Decision aids that really promote shared decision making: the pace quickens. *BMJ*. 2015;350(Feb10):g7624. doi:10.1136/bmj.g7624.
22. 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(6):620–6. doi:10.1016/j.jclinepi.2009.12.014.
23. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, 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;(In press). doi:10.1016/j.jclinepi.2015.12.007.
24. Hsiao J, Chen R. Critical factors influencing physicians ' intention to use computerized clinical practice guidelines : an integrative model of activity theory and the technology acceptance model. *BMC Med Inform Decis Mak*. 2016:1–15. doi:10.1186/s12911-016-0241-3.
25. Cuello García C a, Pacheco Alvarado KP, Pérez Gaxiola G. Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems. *Arch Dis Child*. 2011;96(8):723–728. doi:10.1136/adc.2010.199307.
26. Neumann I, Santesso N, Akl EA, et al. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol*. 2015;(c). doi:10.1016/j.jclinepi.2015.11.017.

Target journal BMJ Open

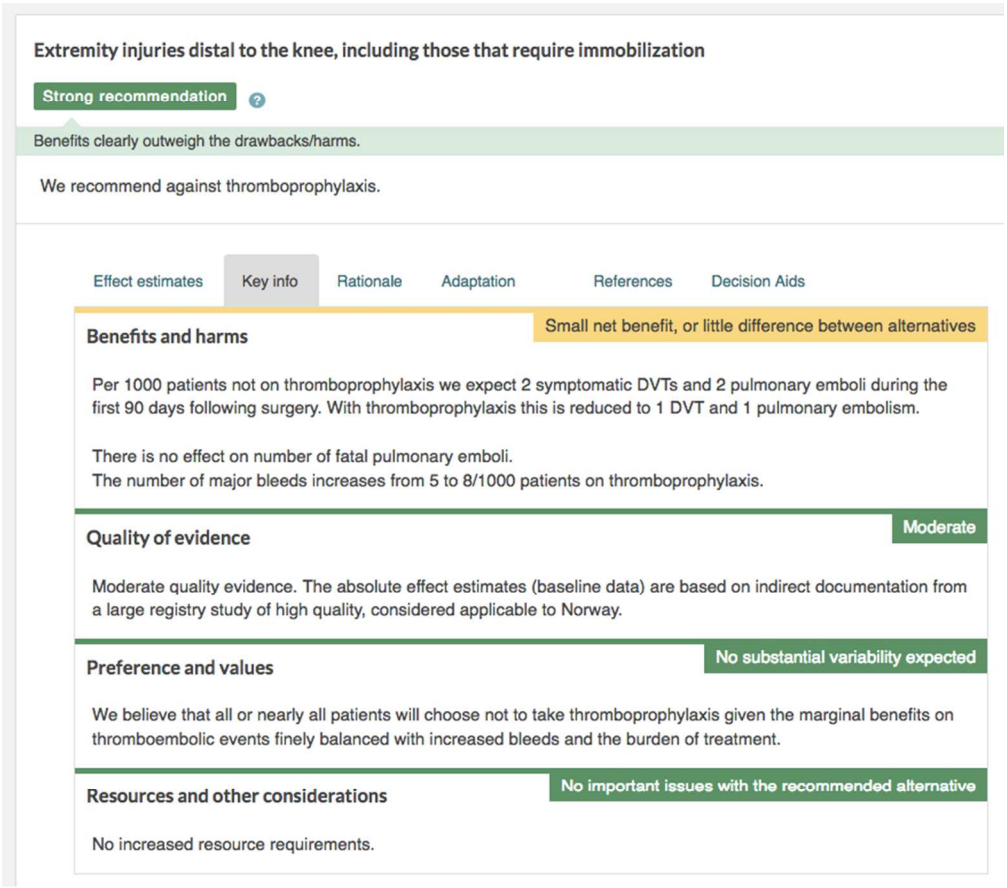


Figure 1a and 1b: Current version of the multilayered guideline presentation formats
Figure 1a
308x271mm (72 x 72 DPI)

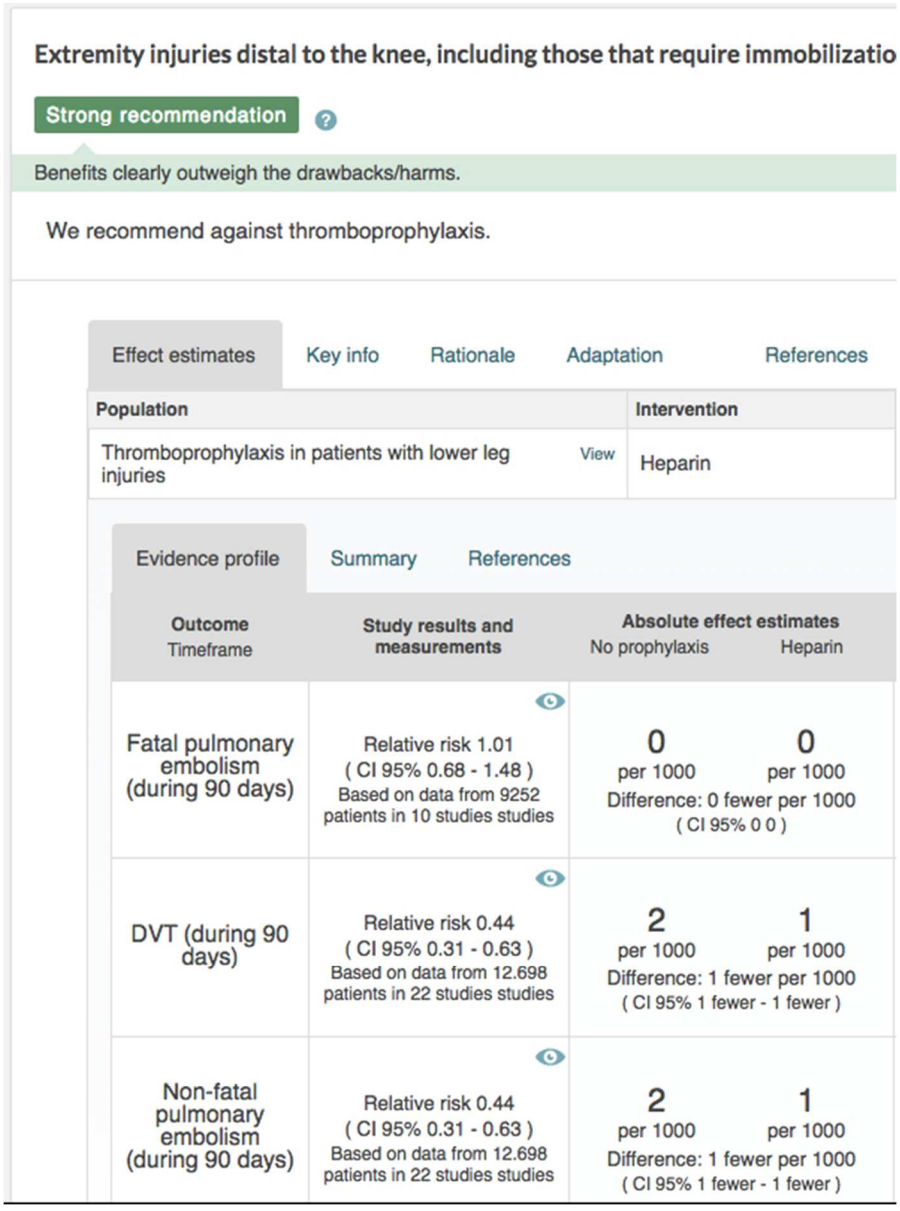


Figure 1a and 1b: Current version of the multilayered guideline presentation formats and 1b
210x283mm (72 x 72 DPI)

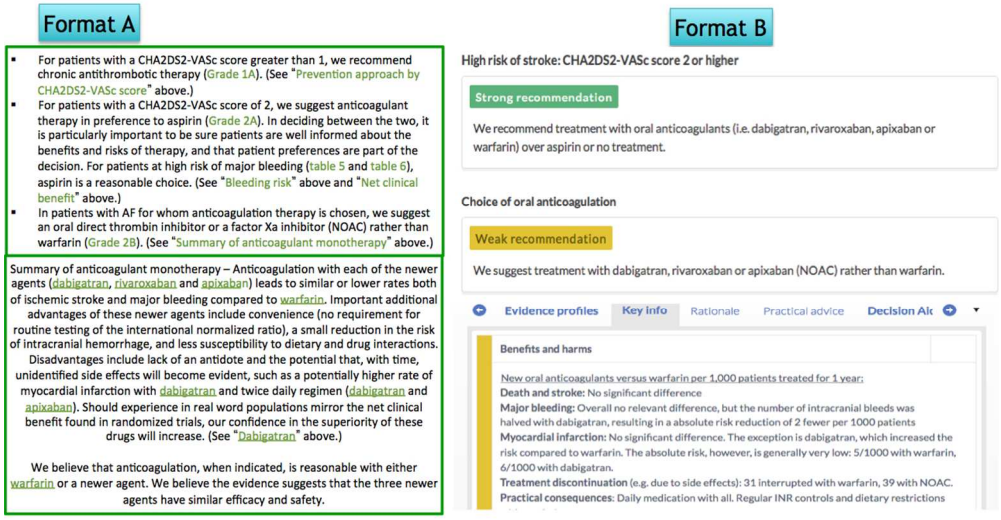


Figure 2: Standard (A) and multilayered guideline (B) presentation formats
Figure 2
455x238mm (72 x 72 DPI)

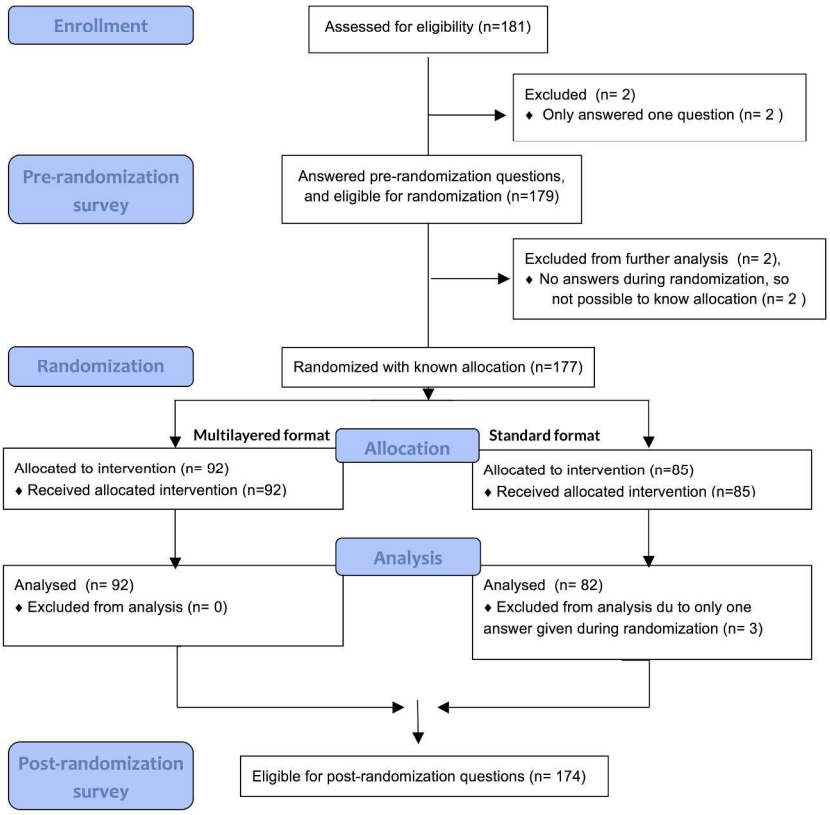


Figure 3: Flow chart of design and enrolment of participants to the multilayered format (n=92) and standard format (n=85)
Figure 3
279x361mm (300 x 300 DPI)

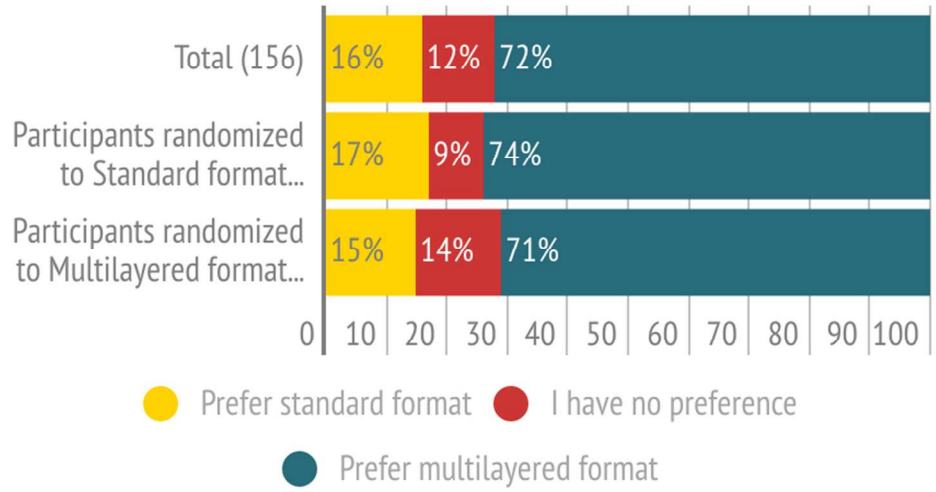


Figure 4: Preferences for standard format versus multilayered presentation format.
Figure 4
322x176mm (72 x 72 DPI)

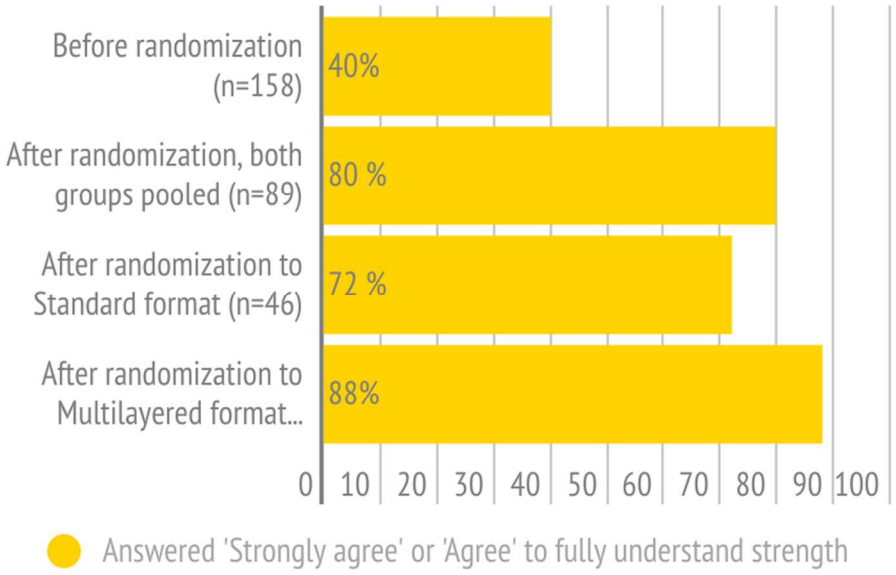


Figure 5: Reported answers to the statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making.

Figure 5
335x206mm (72 x 72 DPI)

DECIDE

Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence

GRADE

New standards for trustworthy guidelines

Clinical scenario:
Anticoagulation treatment
for prevention of stroke
in atrial fibrillation

CLICK-IT studies in educational sessions
DECIDE WP1 2013

1

CLICK-IT

How do clinicians like and understand trustworthy guidelines?
Mixed methods study using Clickers in educational sessions

Objectives:

- ✓ Determine understanding and preferences for guideline presentation formats
- ✓ Teach about new concepts for trustworthy guidelines

- Registered results and data will be used for research.
- We regard answering the questions is to give informed consent for us to use this in research. (You can walk out of the room now)
- The questions are in …… (if another language) , but some of the examples are in English

First, some demographic questions

2

Q 1: My age is

1. 25-35
2. 36-46
3. 46-55
4. 56-65
5. 66-100

Q 2: My position is

1. Intern, medical student
2. Resident physician
3. Consultant physician

Q3: In terms of training in health research methodology (HRM), you have:

1. Never completed a formal course in HRM or epidemiology
2. Completed one or more formal courses in HRM or epidemiology
3. A masters degree or PhD degree in HRM or epidemiology

Meet Gabriel 68y

- **Medical history:** Type 2 diabetes. No medications
- **Chief complaint:** For the past 6 months intermittent episodes of heart palpitations and rapid heart rate, duration between 30 minutes to 3 days
- **Diagnosis:** Atrial fibrillation
- No risk factors indicating increased risk of bleeding
- Risik for stroke? Anticoagulation a prophylaxis?

6

CHA2DS2-VASc Score for Atrial Fibrillation Stroke Risk

Calculates stroke risk for patients with atrial fibrillation, possibly better than the CHADS2 score.

Age?	○ < 65 years old +0 ◉ 65-74 years old +1 ○ ≥ 75 years old +2
Congestive Heart Failure History?	☐ Yes +1
Hypertension History?	☐ Yes +1
Stroke/TIA/Thromboembolism History?	☐ Yes +2
Vascular Disease History? (previous MI, peripheral arterial disease or aortic plaque)	☐ Yes +1
Diabetes Mellitus?	☑ Yes +1
Female?	☐ Yes +1
Score	2

Baseline risk for stroke (no anticoagulation treatment) with **CHA2DS2-VASc score 2**

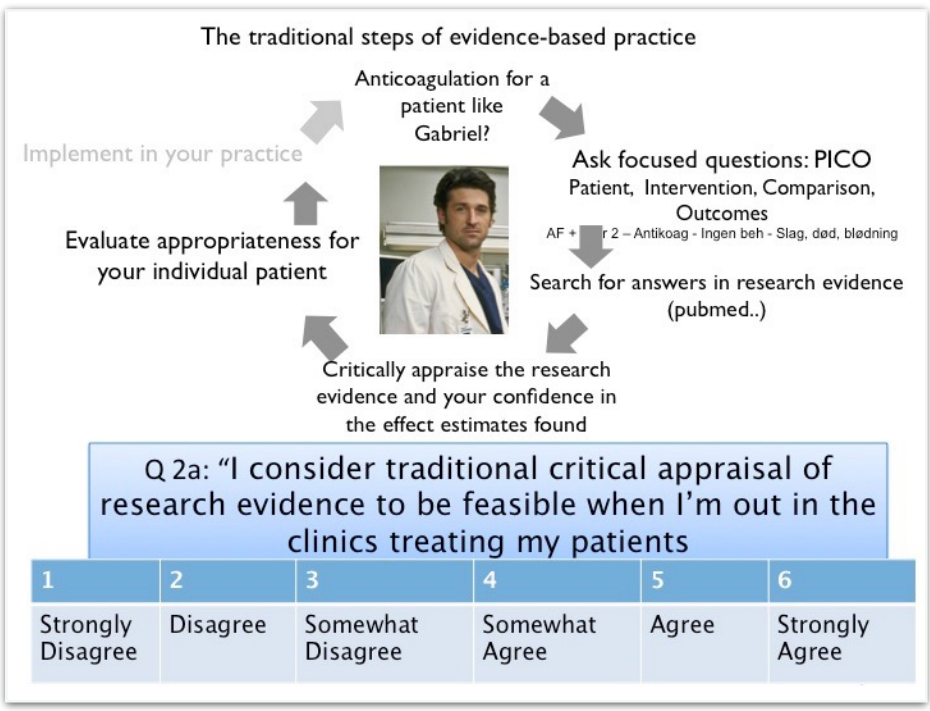
This means an assumed risk of **22 strokes per 1000** patients over 1 year

Treatment for Gabriel?

- **Diagnosis:** Atrial fibrillation
- Moderate risk of stroke (CHAD2S2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

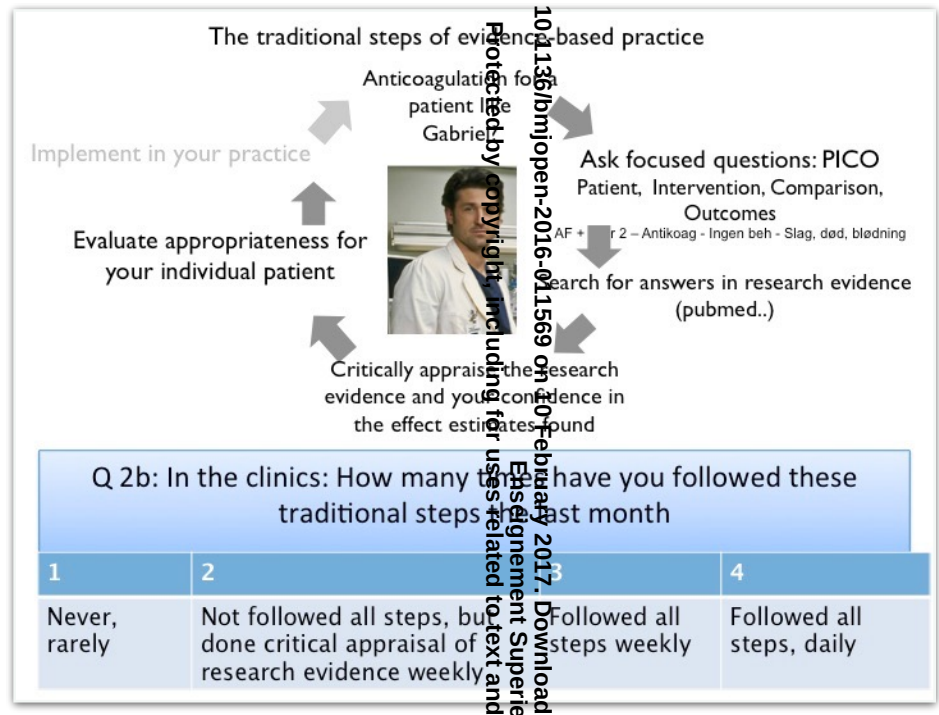
Q 1: If you were unsure of which, if any, therapy to offer the patient, where would you first look for an answer?

1. Local guideline
2. Systematic review
3. EBM textbook (e.g. UpToDate)
4. Practice guideline (national or international)
5. Ask a colleague
6. Individual study



Q 2a: "I consider traditional critical appraisal of research evidence to be feasible when I'm out in the clinics treating my patients

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree



Q 2b: In the clinics: How many times have you followed these traditional steps last month

1	2	3	4
Never, rarely	Not followed all steps, but done critical appraisal of research evidence weekly	Followed all steps weekly	Followed all steps, daily

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 3 : What does the number (2) reflect?

1	2	3	4
It´s a strong recommendation	It´s moderate quality evidence	It´s a moderate recommendation	It´s a weak recommendation

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 4 : What does the letter (B) reflect?

1	2	3	4
Moderate quality evidence	Low quality evidence	Study design (based on a single Randomized study)	Study design (based on a smaller systematic review)

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 5: GRADE provides either strong or weak recommendations. To what extent do you agree with the following statement:

"I fully understand the difference between strong and weak recommendations and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Now, let's get back to Gabriel

1. Two groups, 5 questions each
2. Different formats of guidelines for atrial fibrillation and anticoagulation
3. One group gets blindfolds (they are "blinded")
4. I will not read the questions or text out loud.
5. Read the text and give me a sign (waive/ raise you hand) when you are ready to answer
6. Then switch and the other group gets blindfolds

14

Imagine you search online for an answer to what to do with Gabriel,

and you found the guideline on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

15

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 7 : How do you interpret these recommendations?

1	2	3	4
Strong recommendation for NOAC. Weak recommendation for warfarin or aspirin.	Weak recommendation for NOAC and warfarin.	Weak recommendation for any option.	Strong recommendation for treatment. Weak recommendation for NOAC over warfarin.

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 6: "These recommendations will help me manage my patient"

1	2	3	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Agree	Strongly Agree

You also find the summary you will see on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

18

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and safety.

Q 8 : "This information helps me apply the recommendation on my patient"

1	2	3	4	5	6
Strongly disagree	Disagree	Somewhat disagree	Somewhat Agree	Agree	Strongly Agree

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and safety.

Q 9 : What does this information tell you about NOAC vs warfarin?

1	2	3	4
Vastly superior treatment effect	Less burden of treatment and slightly better treatment effect	Large reduction in side effects	No difference in effect or side effects

How would you have treated Gabriel?

- Diagnosis: Atrial fibrillation
- Moderate risk of stroke (CHA2DS2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

Let's look at the recommendations again

21

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 1, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 10 : Which, if any, antithrombotic treatment would you consider appropriate for Gabriel?

1	2	3	4
NOAC (Dabigatran, rivaroxaban or apixaban)	Aspirin	Warfarin	No therapy

Now give your blindfold to an unblinded colleague

23

Imagine you search online for an answer to what to do with Gabriel,

and you found the guideline on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

24

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 11: "The recommendations will help me manage my patient"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 12 : How do you interpret these recommendations?

1	2	3	4
Strong recommendation for NOAC. Weak recommendation for warfarin or aspirin.	Weak recommendation for NOAC and warfarin.	Weak recommendation for any option.	Strong recommendation for any treatment. Weak recommendation for NOAC over warfarin.

When you click one of the recommendations you find the summary you will see on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

27

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but there was seen a halving of the number intracranial bleeds with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high

Q 13 : "This information helps me apply the recommendation on my patient"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but there was seen a halving of the number intracranial bleeds with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high

Q 14 : What does this information tell you about NOAC vs warfarin?

1	2	3	4
Vastly superior treatment effect	Less burden of treatment and slightly better treatment effect	Large reduction in side effects	No difference in effect or side effects

How would you have treated Gabriel?

- Diagnosis: Atrial fibrillation
- Moderate risk of stroke (CHA2DS2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

Let's look at the recommendations again

30

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 15 : Which, if any, antithrombotic treatment would you consider appropriate for Gabriel?

1	2	3	4
NOAC (Dabigatran, rivaroxaban or apixaban)	Aspirin	Warfarin	No therapy

Now, let us all get unblinded and have a look at both formats for a few minutes

32

Format A

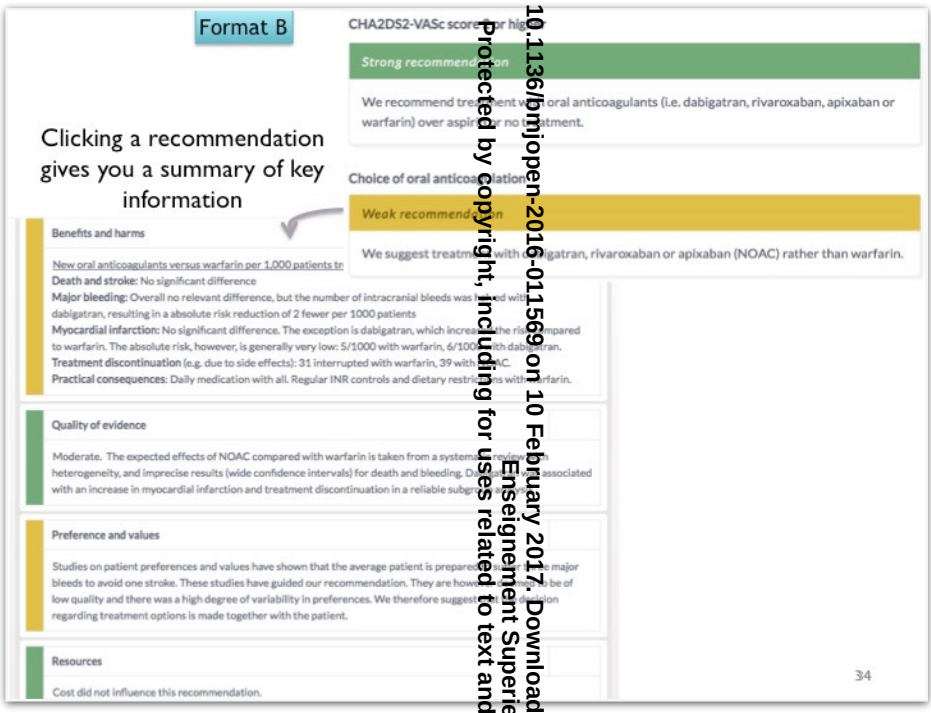
- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and

Format B

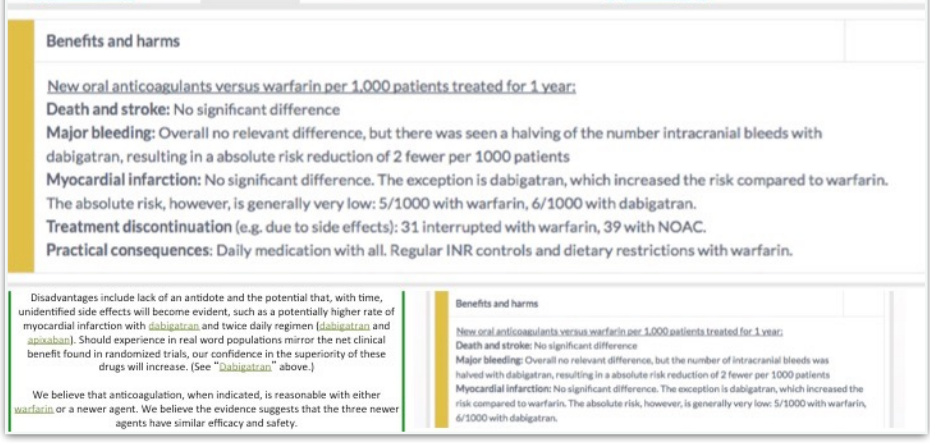
Clicking a recommendation gives you a summary of key information



34

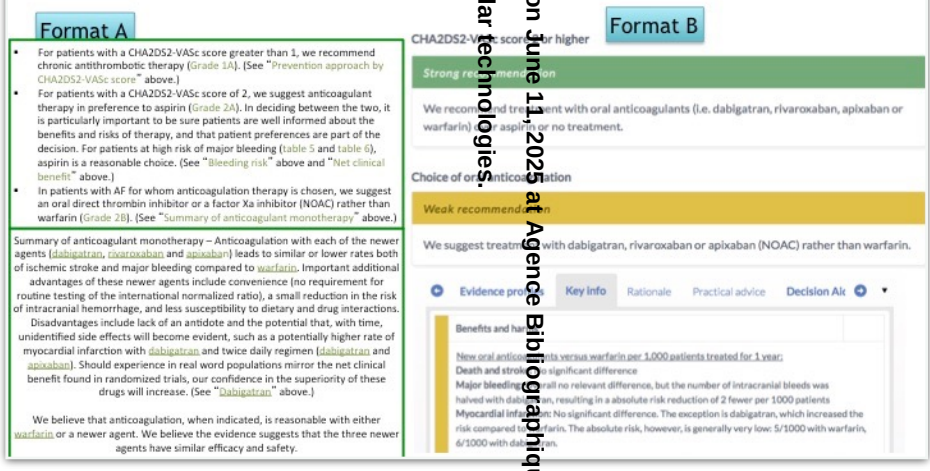
Q 16: Format B presents absolute effects for benefits and harms, whereas format A does not.
What is your first reaction to being presented with the absolute effects?

1	2	3	4	5
Confusing distraction, leave it out	A little confusing but not a big problem	Doesn´t help but doesn´t hurt	Not crucial but helpful	Crucial information, should always be included



Q 17: Overall, do you prefer format A or format B?

1	2	3
Format A	I have no preference	Format B



Understand new definitions and standards for trustworthy guidelines

"Clinical Practice Guidelines are statements that include **recommendations** intended to optimize patient care. They are informed by a **systematic review** of evidence and an assessment of the **benefits and harms of alternative** care options"(2011)

CLINICAL PRACTICE GUIDELINES WE CAN TRUST

Broad consensus

GRADE

Common methods for guidelines

37

GRADE defines **strength** of recommendation as:

"The strength of a recommendation reflects the extent to which we can, across the range of patients for whom the recommendations are intended, be confident that desirable effects of a management strategy outweigh undesirable effects."

Strong recommendations:
Reflects clear benefit of the recommended treatment alternative.
Implications: Recommendation applies to all or nearly all patients. "Just do it"

Weak recommendations:
Reflects fine balance between benefits and harms for the treatment alternatives.
Implications: Recommendation applies to the majority of patients.
"Maybe", "Depends on patient values and preferences"

Q18a: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

GRADE defines **strength** by always considering 4 factors

Strong recommendations:
Reflects clear benefit of the recommended treatment alternative.
Applies to all or nearly all patients. "Just do it"

Weak recommendations:
Reflects fine balance between benefits and harms .
The majority of patients.
"Maybe", "Depends on patient values and preferences"

Process it using the 4 factors

- B&H (Benefits & Harms): What are the expected benefits and harms?
- QoE (Quality of Evidence): What are the overall confidence in these estimates?
- V&P (Values & Preference): What are patients values and preferences? Do they vary? Do you know? Important societal values to consider?
- R (Resources): Are there resource issues to consider? For whom?

Q18b: "This explanation is necessary to understand the difference between strong and weak recommendations and the implications for clinical decision

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

GRADE defines **quality of evidence** as:

High quality: We are very confident that the true effect lies close to our effect estimates

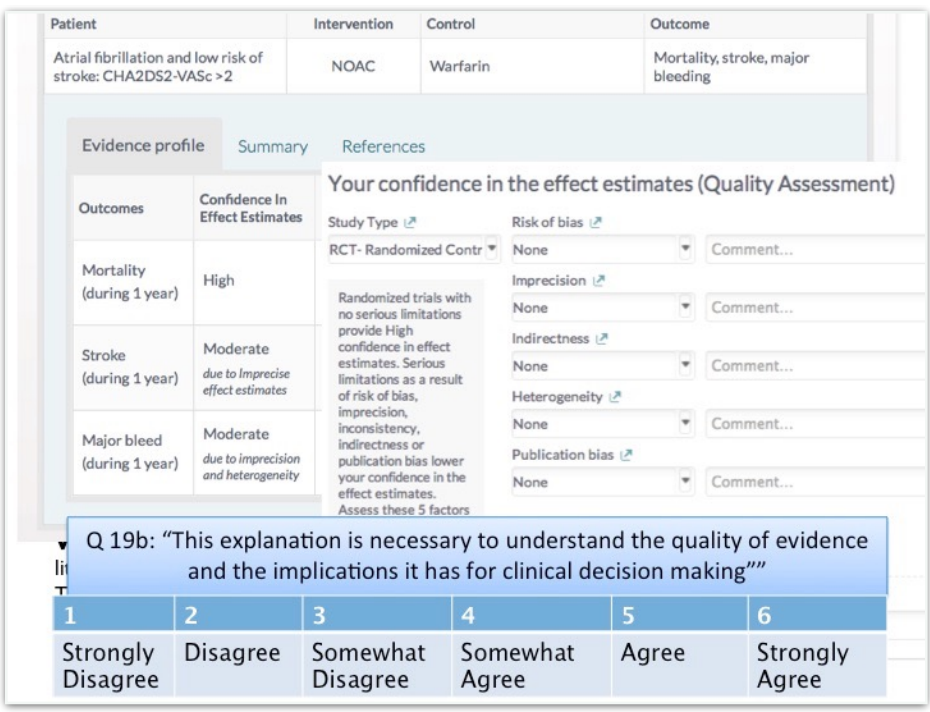
Moderate quality: We are moderately confident in our effect estimates. The true effect is likely to be close to our effect estimates, but with the possibility to be substantially different.

Low quality: Our confidence in the effect estimates are limited. The true effect may be substantially different from our effect estimates.

Very low quality: We have very little confidence in our effect estimates. The true effect is likely to be substantially different from our effect estimates.

Q 19a: "I fully understand the difference between the different categories of quality and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree



Q 19b: "This explanation is necessary to understand the quality of evidence and the implications it has for clinical decision making""

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

There is a lot of information included in a guideline

Do we need to see it all, all the time?

How do we like multilayered guidelines?

42

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

43

Choice of oral anticoagulation

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Effect estimates | **Key info** | Rationale | Practical advice | **Adaptation** | Reference

The new oral anticoagulants have equal effect to warfarin with regards to stroke reduction, they lower the incidence of intracranial bleeds and are more convenient to use. We therefore suggest the new oral anticoagulants over warfarin as first treatment of choice.
For patients that are already on warfarin therapy with stable INR values the cost/benefit ratio is similar to treatment with NOACs. We therefore suggest that patients well-established on warfarin therapy continue with this if they wish.

44

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but the number of intracranial bleeds was halved with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high degree of variability in preferences. We therefore suggest that the decicion regarding treatment options is made together with the patient.

Resources

Cost did not influence this recommendation.

45

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Effect estimates | **Key info** | Rationale | Practical advice | **Adaptation** | Reference

Patient	Intervention	Control	Outcome
Atrial fibrillation and low risk of stroke: CHA2DS2-VASc >2	NOAC	Warfarin	Mortality, stroke, major bleeding

Evidence profile | Summary | References

Outcomes	Confidence In Effect Estimates	Relative Effect	Warfarin	NOAC	Absolute Difference	Participants (Studies), Follow-Up
Mortality (during 1 year)	High	RR 0.88 (CI 0.82 - 0.96)	7 per 1000	6 per 1000	1 fewer per 1000 (CI 1 fewer - 0 fewer)	44.442 (3), 2 years
Stroke (during 1 year)	Moderate due to imprecise effect estimates	RR 0.89 (CI 0.78 - 1.02)	3 per 1000	3 per 1000	0 fewer per 1000 (CI 1 fewer - 0 fewer)	44.442 (3), 2 years
Major bleed (during 1 year)	Moderate due to imprecision and heterogeneity	RR 0.8 (CI 0.63 - 1.01)	25 per 1000	20 per 1000	5 fewer per 1000 (CI 9 fewer - 0 fewer)	44.442 (3), 2 years

46

"This is a smart layout for guidelines"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Thank you for your participation

For more information go to:

http://www.decide-collaboration.eu

http://www.gradeworkinggroup.org

http://www.magicproject.org

If there is time left we can look at your results

48

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	p1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance, see CONSORT for abstracts)	p2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	p4 (last paragraph)-p5
	2b	Specific objectives or hypotheses	p5 (second paragraph)
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	p6 (second paragraph). Flowchart p12
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	No, described p5 (second paragraph)
pParticipants	4a	Eligibility criteria for participants	p5 (second paragraph)
	4b	Settings and locations where the data were collected	p5 (second paragraph)
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	p5 -6 + Figure 2 + attached lecture
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	p9-10
	6b	Any changes to trial outcomes after the trial commenced, with reasons	No changes
Sample size	7a	How sample size was determined	Power calculation, in

			protocol
	7b	When applicable, explanation of any interim analyses and stopping guidelines	Not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	p8 (first paragraph)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	p8 (first paragraph)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	p8 (first paragraph)
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	p8 (first paragraph)
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	p8 (first paragraph)
	11b	If relevant, description of the similarity of interventions	p7 (last paragraph) -8
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	P10 (last paragraph)
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	None, explained p11 (first paragraph)
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	p11 (first paragraph)
	13b	For each group, losses and exclusions after randomisation, together with reasons	p11 (first paragraph)
Recruitment	14a	Dates defining the periods of recruitment and follow-up	P11 (second paragraph)
	14b	Why the trial ended or was stopped	Stopped only as planned.
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	p11
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was	Table1 p11

		by original assigned groups	and Figure3 p12
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	p13-14
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	p13-14
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	None performed
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	n/a
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	p16 (second paragraph)
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	p16 (second paragraph)-p17 (first and second paragraph)
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	p17-18
Other information			
Registration	23	Registration number and name of trial registry	None
Protocol	24	Where the full trial protocol can be accessed, if available	Appendix
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	FUNDING statement

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

BMJ Open

Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011569.R2
Article Type:	Research
Date Submitted by the Author:	29-Nov-2016
Complete List of Authors:	Brandt, Linn; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Vandvik, Per; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society Alonso-Coello, Pablo; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Akl, Elie; American University of Beirut, Department of Internal Medicine Thornton, Judith; National Institute for Health and Care Excellence Rigau, David ; Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau Adams, Katie; National Institute for Health and Care Excellence O'Connor, Paul; National Institute for Health and Care Excellence Guyatt, Gordon; Mcmaster University, Department of Clinical Epidemiology and Biostatistics Kristiansen, Annette; Sykehuset Innlandet Hospital Trust, Department of Internal Medicine; Faculty of Medicine, University of Oslo , Institute of Health and Society
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Health informatics, Epidemiology, Communication
Keywords:	Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, EPIDEMIOLOGY

Target journal BMJ Open

Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial

Linn Brandt[1,2], Per Olav Vandvik[1,2], Pablo Alonso-Coello[3,6], Elie A. Akl[4,6], Judith Thornton[5], David Rigau[3], Katie Adams[5], Paul O'Connor[5], Gordon Guyatt[6], and Annette Kristiansen[1,2].

Author Affiliations: [1)] Department of Internal Medicine, Sykehuset Innlandet Hospital Trust, Gjøvik, Norway (Drs. Brandt, Kristiansen and Vandvik), [2)]Institute of Health and Society, Faculty of Medicine, University of Oslo (Drs. Brandt, Kristiansen and Vandvik), [3)] Iberoamerican Cochrane Centre, Biomedical Research Institute (IIB Sant Pau-CIBERESP), Barcelona, Spain (Drs. Alonso-Coello and Rigau), [4)] Department of Internal Medicine, American University of Beirut, Lebanon (Dr. Akl), [5)] National Institute for Health and Care Excellence, Manchester UK (Dr. Judith Thornton, Katie Adams, Paul O'Connor), [6)] Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (Drs. Alonso-Coello, Akl and Guyatt).

Corresponding Author: Linn Brandt, Department of Internal Medicine, Sykehuset Innlandet Gjøvik, Kyrre Grepps gate 11, 2819 Gjøvik, Norway. E-mail: brandtlinn@gmail.com

Designers: Rob Fracisco, Frankie Achille and Sarah Rosenbaum helped designing the novel multilayered format.

Abstract

Objectives: To investigate practicing physicians' preferences, perceived usefulness and understanding of a new multilayered guideline presentation format - compared to a standard format - as well as conceptual understanding of trustworthy guideline concepts.

Design: Participants attended a standardized lecture in which they were presented with a clinical scenario and randomised to view a guideline recommendation in a multilayered format or standard format after which they answered multiple-choice questions using clickers. Both groups were also presented and asked about guideline concepts.

Setting: Mandatory educational lectures in seven non-academic and academic hospitals, and two settings involving primary care in Lebanon, Norway, Spain and United Kingdom.

Participants: 181 practicing physicians in internal medicine (156) and general practice (25).

Interventions: A new digitally structured, multilayered guideline presentation format and a standard narrative presentation format currently in widespread use.

Primary and secondary outcome measures: Our primary outcome was preference for presentation format. Understanding, perceived usefulness and perception of absolute effects were secondary outcomes.

Results: 72% (95% CI 65-79) of participants preferred the multilayered format and 16% (95% CI 10-22) preferred the standard format. A majority agreed that recommendations (multilayered 86% vs. standard 91%, p-value = 0.31) and evidence summaries (79% vs. 77%, p-value = 0.76) were useful in the context of the clinical scenario. 72% of participants randomised to the multilayered format vs. 58% for standard formats reported correct understanding of the recommendations (p-value = 0.06). Most participants elected an appropriate clinical action after

viewing the recommendations (98% vs. 92%, p-value = 0.10). 82% of the participants considered absolute effect estimates in evidence summaries helpful or crucial.

Conclusions: Clinicians clearly preferred a novel multilayered presentation format to the standard format. Whether the preferred format improves decision-making and has an impact on patient important outcomes merits further investigation.

Trial registration: None.

Strengths and limitations of this study

- We conducted a multicentre trial targeting regular educational sessions.
- Both formats were taken from published guidelines, and we used a comparator format that most participants were familiar with.
- To avoid peer pressure, we ensured participants anonymity through use of clickers (audience response technology).
- A weakness of the study is having researchers involved in development of the new presentation format perform most of the educational sessions.
- We did not measure impact of alternative formats on clinical decisions or patient important outcomes.

BACKGROUND

Clinical practice guidelines that provide recommendations addressing diagnosis and treatment can help clinicians optimize their evidence-based practice (EBP) at the point of care.[1] An abundance of guidelines are available, but many have shortcomings with their trustworthiness and dissemination strategies.[2,3] New standards for trustworthy guidelines developed by the Institute of Medicine and the Guideline International Network highlight the need for more rigorous development processes.[4–6]

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (www.gradeworkinggroup.org) represents a systematic, explicit and transparent process for evaluating and reporting quality of research evidence and for moving from evidence to recommendations.[7,8] GRADE facilitates the creation of trustworthy guidelines and has been adopted by more than 100 organisations worldwide.

Implementation of guidelines requires effective dissemination of recommendations.[9] Guidelines should generally answer clinicians' informational needs within two minutes, which implies that recommendations need to be easy to find, understand, apply and share.[10] Traditionally, guidelines have often been distributed as comprehensive PDFs, impeding efficient use at the point of care. With these challenges in mind, the GRADE working group initiated the Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE) project.[11] DECIDE aimed to improve dissemination of evidence-based recommendations for a range of stakeholders, including health care professionals, policy makers and patients, as well as to ensure and facilitate adherence to trustworthy guideline standards.[4–6]

As detailed in articles by Treweek et al.[11] and Kristiansen et al.,[12] a multidisciplinary group of clinicians, guideline developers, methodologists and graphical designers developed a multilayered guideline presentation format (figure 1a and 1b) targeted at health care professionals. We hypothesized that clinicians' apparent discomfort with more complex

methodological concepts could be alleviated with proper education and an optimal user interface. Based on the groups´ extensive experience from clinical practice and guideline development and informed by a narrative review of guideline formats, we designed a prototype presentation format through brainstorming sessions. This prototype format was iteratively improved based on results from stakeholder feedback and usability testing with clinicians. [12]

During usability testing we confirmed previous research demonstrating limitations in clinicians' conceptual understanding of key standards for trustworthy guidelines. We have since deployed the multilayered format in real-life guidelines.[13] Uncertain of the relative merits of our novel versus existing formats, we conducted a combined survey and randomised controlled trial to determine clinicians' preferences for the new multilayered presentation format versus a traditional format for guidelines, the perceived usefulness of guideline recommendations and understanding of key concepts of trustworthy guidelines.

Figure 1a and 1b: Current version of the multilayered guideline presentation formats.

MATERIALS AND METHODS

This study is the result of a collaboration of DECIDE and MAGIC (Making GRADE the Irresistible choice), a non-profit innovation and research program (www.magicproject.org), which aims at facilitating the efficient creation, dissemination and dynamic updating of trustworthy guidelines and evidence summaries. As previously reported, MAGIC has created a web-based guideline authoring and publication platform (www.magicapp.org), incorporating the digitally structured multilayered presentation format used in this study.[14] Through DECIDE these novel formats have also been incorporated in GRADEpro (http://gradepro.org/). Organizations can use these platforms or freely adopt research outputs from the DECIDE project, including but not limited to the multilayered presentation format, into their own workflow, tools or platforms.

Study design, setting and participants

We applied a combined survey and randomised controlled trial. We included practicing

physicians in internal or family medicine. In order to recruit representative samples of internal medicine physicians, investigators targeted compulsory educational sessions at teaching hospitals. Two facilities recruited general practitioners in family medicine. One by targeting a compulsory educational session at a larger family practice centre (Spain), the other by inviting individual physicians, including general practitioners, to a specific CLICK-IT study session (NICE, UK). All participants attended a standardized educational session on key guideline standards performed within the context of this study. We performed four pilot sessions and made revisions to the survey questions based on experiences in these sessions. The revisions were minor and concerned mainly phrasing of the questions asked. Our reported primary and secondary outcomes are in accordance with the original study protocol.

Participants were considered to provide consent by accepting to answer the survey questions. Research ethics boards in each participating country approved the study: Institutional Review Board (IRB) of the American University of Beirut (Libanon), Ethics committee and R&D at NICE (UK), Ethics committee of the Hospital Sant Pau (Spain) and Regional Ethical Committee South East (Norway).

Standardized lecture and study procedure

At each site, a member of the research team delivered a standardized lecture according to a predefined protocol. The lecture was titled "New standards for trustworthy guidelines" (appendix 1) and was given in power point on a standard projector screen. Participants provided anonymous answers to questions with predefined response categories using clickers. The questions (both for the survey and the randomised part of the trial) - as well as screenshots of presentation formats - were embedded in the lecture slides using an audience response software, TurningPoint©, and read out loud by the presenter. For the Norwegian and Spanish sites we translated the presentation to their native language; while in Lebanon and the UK the questions were presented in English, which is commonly used in medical education.

The lecture and study procedure were developed by the investigators and included the following

components:

1. Collection of the demographics of the participants, their current preference for information resources and understanding of the GRADE system.

2. Presentation of a clinical scenario concerning choice of oral anticoagulation in a patient with atrial fibrillation and high risk of stroke (CHA2DS2-VASc score 2).

3. Presentation of guideline recommendations and evidence summaries relevant to the clinical scenario, sequentially presented in the two alternative guideline presentation formats through randomisation and blinding as outlined below. Both formats were shown side by side to both groups at the end.

4. Short presentation of key conceptual definitions of trustworthy guidelines, specifically explanation of the strength of recommendations and quality of the evidence using the GRADE methodology.

Guideline presentation formats

Figure 2 shows the standard narrative guideline presentation format (format A) and the new multilayered guideline presentation format (format B). We extracted both formats from existing guidelines, but masked the publishers' identity to avoid potentially biasing participants' responses. Both formats were displayed as screenshots in the presentation slides.

Multilayered format: The experimental multilayered format displays recommendations up front with supporting information as collapsible boxes provided by clicking on the recommendation itself. The strength of the recommendation is communicated by use of text and colour coding, and a header describes the population for which the recommendation applies. The example was taken from a Norwegian guideline for antithrombotic therapy applying the multilayered presentation formats published in MAGICapp and translated to English for the purpose of this study.[13,15]

Standard format: The control standard format displays recommendations and an abridged evidence summary from UpToDate.[16] We considered UpToDate's presentations a suitable

reference standard for current presentation formats due to its widespread use, commonality with other guidelines, evidence-based approach and use of GRADE.[17,18] UpToDate provides a textual summary of its recommendations in bullet points with links to the supporting information in the main text or in other articles. The recommendations are labelled with numbers (1 & 2) and letters (A-C), depicting the strength of the recommendation and quality of evidence respectively.

Figure 2: Standard and multilayered guideline presentation formats.

Randomisation and blinding

A research member colour marked the base of half the clickers and haphazardly rearranged them in their container. The presenter handed out the clickers to the participants, randomly and with the front-side up, thus concealing the allocation marking for both presenter and participant. The participants were not informed that the marking was part of the group allocation until the survey part of the lecture was completed, and they had no opportunity to swap clickers during the lecture. Participants with marked clickers were randomised to group B, being presented the multilayered format. We asked participants in either group to put on blindfolds while participants in the other group were shown their allocated presentation format and answered questions. Questions during randomisation were not read out loud, only the group not wearing blindfolds at the time answered the questions. The presenters watched while the participants put on blindfolds, making sure they adhered to their allocated group.

Outcome measures

Our primary outcome was preference for either presentation format. We provided three response options: preference for the standard format, no preference and preference for the multilayered format.

Secondary outcomes included:

- Correct understanding/interpretation of 1) the evidence summaries and 2) the recommendation with four potential answers, one alternative being correct.
- Anticipated course of action to the clinical scenario with four potential answers, two alternatives being correct.
- Participants' perceived usefulness of 1) evidence summaries and 2) recommendations. Participants provided answers to the statement "This information/recommendation would help me manage my patient" on a 6 point Likert scale: 1= strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree and 6=strongly agree.
- Correct understanding of the strength of the following example recommendation and the confidence in effect estimates: "We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol). GRADE 2B." We provided participants with four potential answers, one alternative being correct.
- Perceived understanding of the strength of the recommendation. We asked the participants the following question twice, before and after being provided with a short written explanation: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision-making." They provided answers on a 6 point Likert scale with 2 anchors: 1= strongly disagree and 6 = strongly agree.
- Participants' perception of presenting absolute effect estimates. We asked participants "What is your first reaction to being presented with absolute effects?" The answers were collected using a 5-point scale: 1=confusing distraction, leave it out, 2=a little confusing, but not a big problem, 3=doesn't help, but doesn't hurt, 4=not crucial, but helpful and 5=crucial information, should always be included.

Statistical analysis

We dichotomized all outcomes to either correct/incorrect or agree/disagree, and analysed them by use of Pearson's Chi-square test. We included all randomised participants that answered more than one question in the final analysis (Intention To Treat). No subgroup analyses were specified in the protocol. The accompanying data set (Excel and original Turning Point data files from

each centre) provides results subdivided per type of physician and centre. We used SPSS (version 23) for all analyses.

RESULTS

We performed the study from June 2013 until January 2015. We included 181 practicing physicians across four countries and nine centres (Lebanon 1, Norway 6, Spain 1 and UK 1), 177 were randomised and 174 (96%) answered > 1 question and are included in the final analysis (figure 3). Their demographics and information resource preferences are provided in table 1. The two groups were fairly similar. One centre (27 participants) in Norway did not include demographic questions, so the demographic presentation includes the remaining eight centres.

Number of participants randomised	**Multilayered format** 92	**Standard format** 85
Country		
(# participants eligible for analysis)	(92)	(83)
Norway (%)	61 (66.3%)	57 (68.7%)
UK (%)	10 (10.9%)	10 (12%)
Lebanon (%)	11 (12.0%)	10 (13.3%)
Spain (%)	10 (10.9%)	5 (6%)
Professional status or Specialty		
(# participants eligible for analysis)	(76)	(72)
Medical student or intern (%)	13 (17.1%)	8 (11.1%)
Internist Resident (%)	21 (27.6%)	27 (37.5%)
Internist Attending/Consultant (%)	23 (30.3%)	21 (29.2)
General practitioner (%)	13 (17.1%)	12 (16.7%)
Unknown (% did not answer that question)	6 (7.9%)	4 (5.6%)
Training in health research methodology		
(# participants eligible for analysis)	(72)	(68)
No training in HRM (%)	34 (47.2%)	35 (51.2%)
≥ 1 HRM course (%)	26 (36.1%)	21 (30.9%)
Degree in HRM (%)	12 (16.7%)	12 (17.6)
Preferred knowledge source		
(# participants eligible for analysis)	(89)	(82)
Local guideline (%)	22 (24.7%)	14 (17.1%)
Systematic review (%)	2 (2.2%)	2 (2.4%)
EBM textbook (%)	17 (19.1%)	13 (15.9%)
National or international guideline (%)	34 (38.2%)	36 (43.9%)

Colleague (%)	14 (15.7%)	17 (20.7%)
Primary study (%)	0 (0%)	0 (0%)

Table 1: Demographics of the groups randomised to different formats. HRM=Health Research Methodology. EBM=Evidence Based Medicine.

Figure 3: Flow chart of design and enrolment of participants to the multilayered format (n=92) and standard format (n=85)

Preference for alternative presentation formats

When exposed to both formats after completing the randomised part of the study 113 of 156 (72%, 95% CI 65-79) participants preferred the new multilayered format, 25 (16%, 95% CI 10-22) preferred the standard format, and 18 (12%, 95% CI 7-17) reported no preference. Results were similar in those randomised to the standard vs. the multilayered format (p = 0.66, figure 4).

Figure 4: Preferences for standard format versus multilayered presentation format**.**

Understanding and anticipated clinical action

69 of 78 participants (88%, 95% CI 81-96) randomised to the standard format, and 57 of 64 participants (89%, 95% CI 82-97) randomised to the multilayered format correctly understood the evidence summaries, with no difference between groups (p-value = 0.91). A majority correctly understood the recommendations, with no statistically significant difference between the standard format (44/ 76, 58%) and the multilayered format (55/ 76, 72%) (p-value = 0.06).

Most participants correctly stated they would consider an oral anticoagulant as an appropriate course of treatment, with a majority preferring direct acting oral anticoagulants (DOACs) rather than warfarin. We observed no statistically significant difference between the presentation formats (standard format 66/72 (92%, 95% CI 85-98) versus the multilayered format 79/81 (98%, 95% CI 94-100), p-value = 0.10).

Perceived usefulness of evidence summaries and recommendations

Of the 69 participants in the standard format group, 53 (77%, 95% CI 67-87) agreed (somewhat agree, agree or strongly agree) that the background evidence summaries were helpful in the context of the clinical scenario, as did 60 of 76 participants in the multilayered format group (79%, 95% CI 70-88), with no difference between the randomised groups (p-value = 0.76). 72 of 79 participants (91%, 95% CI 85-97) randomised to the standard format and 74 of 86 participants (86%, 95% CI 79-93) to the multilayered format agreed that recommendations were helpful in the context of the clinical scenario (p-value = 0.31). When specifically asked, 84 of 102 (82%, 95% CI 75-90) participants considered absolute effect estimates provided in the multilayered format evidence summaries helpful or crucial.

Survey on conceptual understanding

Prior to randomisation, we provided the participants with a recommendation labelled with "2B" and asked them what the number 2 and the letter B meant. Few participants answered correctly that this represented a weak (2) recommendation based on moderate (B) quality evidence (20/154 (13%, 95% CI 8-18) stated weak, while 61/144 (42%, 95% CI 34-50) stated moderate). We furthermore twice asked to what extent they agreed to the following statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making." Prior to randomisation 63 of 158 participants (40%, 95% CI 32-48, figure 5) stated that they agreed (agreed or strongly agreed) to this statement. After randomisation, we provided the participants with a one slide explanation of the strength of the recommendation according to the GRADE system, defining the difference between strong and weak recommendations and posed the same question again. 71 of 89 participants (80%, 95% CI 70-88, figure 5) agreed with the statement. There was a borderline significant difference between participants according to randomised format (standard format 72% vs. multilayered format 88%, p = 0.051, figure 5).

Figure 5: Reported answers to the statement: "I fully understand the difference between strong

and weak recommendations and the implications for clinical decision making."

DISCUSSION

In this study of practicing physicians, we demonstrated a clear preference for the new guideline multilayered presentation format rather than a traditional narrative format, with a majority agreeing that recommendations and underlying evidence summaries - regardless of presentation format - were useful in the context of the clinical scenario. A majority of the physicians (82%) also reported absolute effect estimates provided in the multilayered format evidence summaries to be helpful or crucial for decision-making. Conceptual understanding of strength of recommendations and quality of evidence was limited when expressed through numbers and letters - as done in the standard format example from UpToDate. A short explanation of these concepts according to the GRADE system substantially increased reported understanding.

Strengths and weaknesses

We enlisted a substantial and diverse set of practising physicians across six different countries, targeting mandatory educational sessions, making our participant sample representative of the everyday clinician and our results generalizable. We mapped understanding of both key methodological concepts and preference for entire guideline formats, as opposed to single elements. We devised a clinical scenario with a highly prevalent disease (atrial fibrillation), in which there have been recent treatment innovations (DOACs), possibly not commonly known to all physicians when this study was conducted. Through this manoeuvre, we limited the possible bias of previous knowledge on the participants' performance. We targeted internists and family physicians as participants given their familiarity with atrial fibrillation.

There are limitations to the chosen study design. Firstly, devising multiple-choice questions that accurately test key outcomes such as understanding is challenging. Our approaches, including subjectively perceived understanding and anticipated clinical action are less satisfactory than detailed testing of understanding. Actual understanding may be less than our results suggest, and the correct clinical choice of action may have been highly influenced by previous knowledge on

the subject. A substantial proportion of physicians reported to "somewhat agree" rather than agree or strongly agree with statements concerning usefulness of recommendations and evidence summaries in clinical practice guidelines. This finding highlights some limitations of our study related to using a hypothetical scenario - rather than assessing real life decision making - and the challenge of phrasing questions and devising appropriate response categories within this field of research. Lastly, as we have only tested one specific clinical scenario on a limited representation of health care professionals, the transferability of our findings to other settings and professional groups needs further validation.

Implications for practice and research

The finding that the majority of clinicians preferred to be informed about the absolute effects on patient-important outcomes in evidence summaries is encouraging, as numeracy among health care professionals is highly variable and many guidelines omit contextualised effect estimates necessary for shared decision making.[19–21] GRADE Summary of findings tables - as displayed in the multilayered guideline formats - have emerged as user-friendly and well accepted formats in the context of systematic reviews, with recent advances in formatting resulting in further increasing understanding.[22,23]

40% of the participants stated that they fully understood the difference between strong and weak recommendations, but only 13% could correctly recognize a weak recommendation applying the commonly used numbered labelling (1=strong, 2=weak). However, when given recommendation and evidence summary linked to a clinical scenario 95% of participants would treat according to current guidelines. There are thus two reasons for optimism: Most participants correctly interpreted the intent and meaning of the recommendations when communicated within a larger context. Furthermore, perceived understanding improved after being presented a one slide explanation on the meaning of key concepts.

We did not explore in detail why participants preferred the new DECIDE multilayered format. However, informed by feedback throughout the design process from stakeholders as well as

informal discussions following the survey, clinicians seem to appreciate short and clear advice, provision of strength that is easy to interpret and details around the key factors that drove the recommendations, which is provided within the multilayered format.

Optimized guideline presentation formats and sufficient conceptual understanding can potentially facilitate the uptake of trustworthy guidelines and application of research evidence in practice.[24,25] The multilayered format serves several purposes: It is devised around the GRADE framework and thus directs guideline authors through the appropriate methodological steps. It provides end-user clinicians with actionable, graded recommendations, as advocated by the Institute of Medicine.[26] Lastly, having the guideline digitally structured facilitates easy translation into several outputs; such as decision aids, clinical decision support systems within the electronic medical record, tablets, smartphones, online and as PDFs. It also facilitates continuous updating and adaptation to local settings,[14] thereby minimizing the workload for guideline developers and policy makers.

Further research into different ways of communicating key guideline concepts to health care professionals to improve understanding and adoption is still necessary. The multilayered guideline presentation formats are currently implemented in a handful of published guidelines from a variety of organizations, and more are under development. We are continuously performing usability testing, both with authors and end-users, informing further improvements.

CONCLUSION

Clinicians prefer a novel multilayered presentation format to the standard format. Optimized guideline presentation formats and sufficient conceptual understanding can potentially facilitate the uptake of trustworthy guidelines and application of research evidence in practice. Whether the preferred format improves decision-making and has an impact on patient-important outcomes merits further investigation.

CONTRIBUTORSHIP STATEMENT

LB, POV, EA, DR, KA and POC collected the data. LB, POV and AK had access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. LB was principal investigator on the study. LB, AK, POV, PAC and AE contributed to the conception, design, and ethical approval of the study. LB, POV and AK contributed to writing the first draft of the article; and LB, POV, PAC, EA, JT, DR, KA, POC, GG and AK contributed to editing and approval of the final manuscript.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form. LB, AK, GG and POV are members of a non-profit research and innovation project MAGIC: www.magicproject.org, which has an open technical platform where the new DECIDE multilayered formats were prototyped. All authors were either co-investigators or collaborators of the DECIDE project. LB, AK, GG, POV, PAC, EA, JT and DR are members of the Grade Working group. The strategy evaluated in the study is based on the GRADE approach. No other competing interests were declared.

FUNDING

This work was supported by Sykehuset Innlandet Hospital Trust under grant number 150194 (PLUGGED IN), The Research Council of Norway (VERDIKT) under grant 193022 (EVICARE), and European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 258583 (DECIDE project)

.

DATA SHARING MANAGEMENT

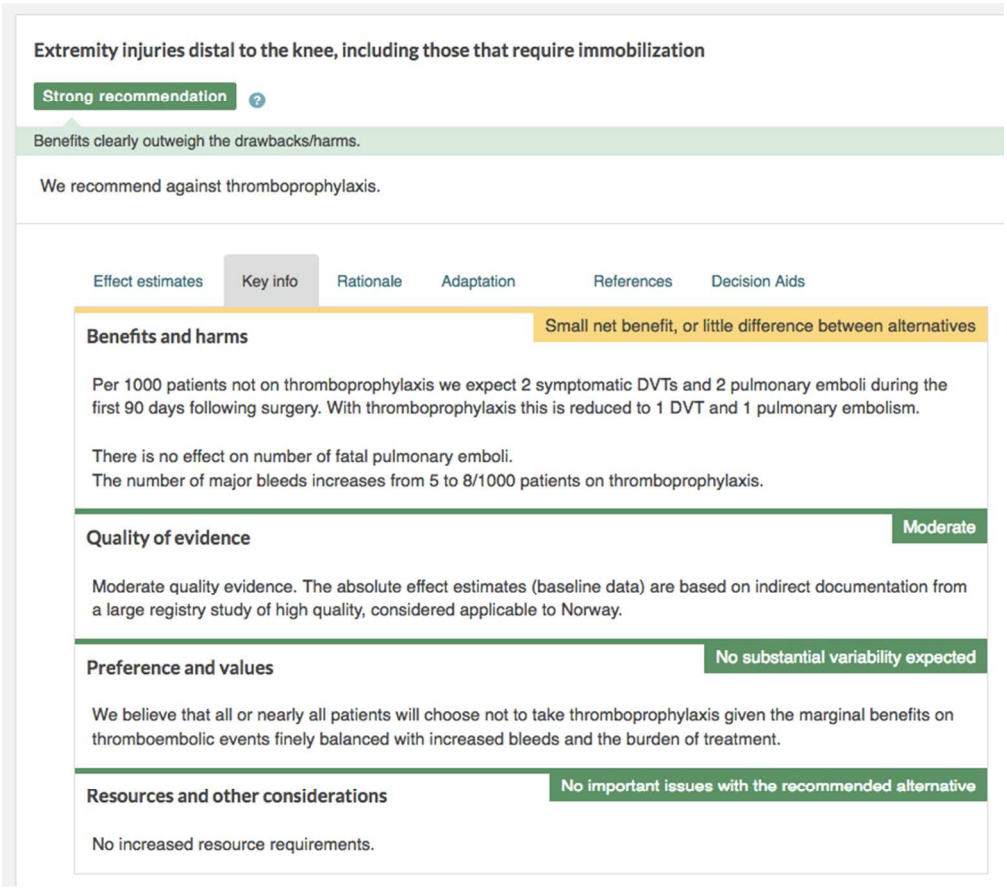
Technical appendix and full dataset available from the Dryad repository, DOI: 10.5061/dryad.2qv30

REFERENCES

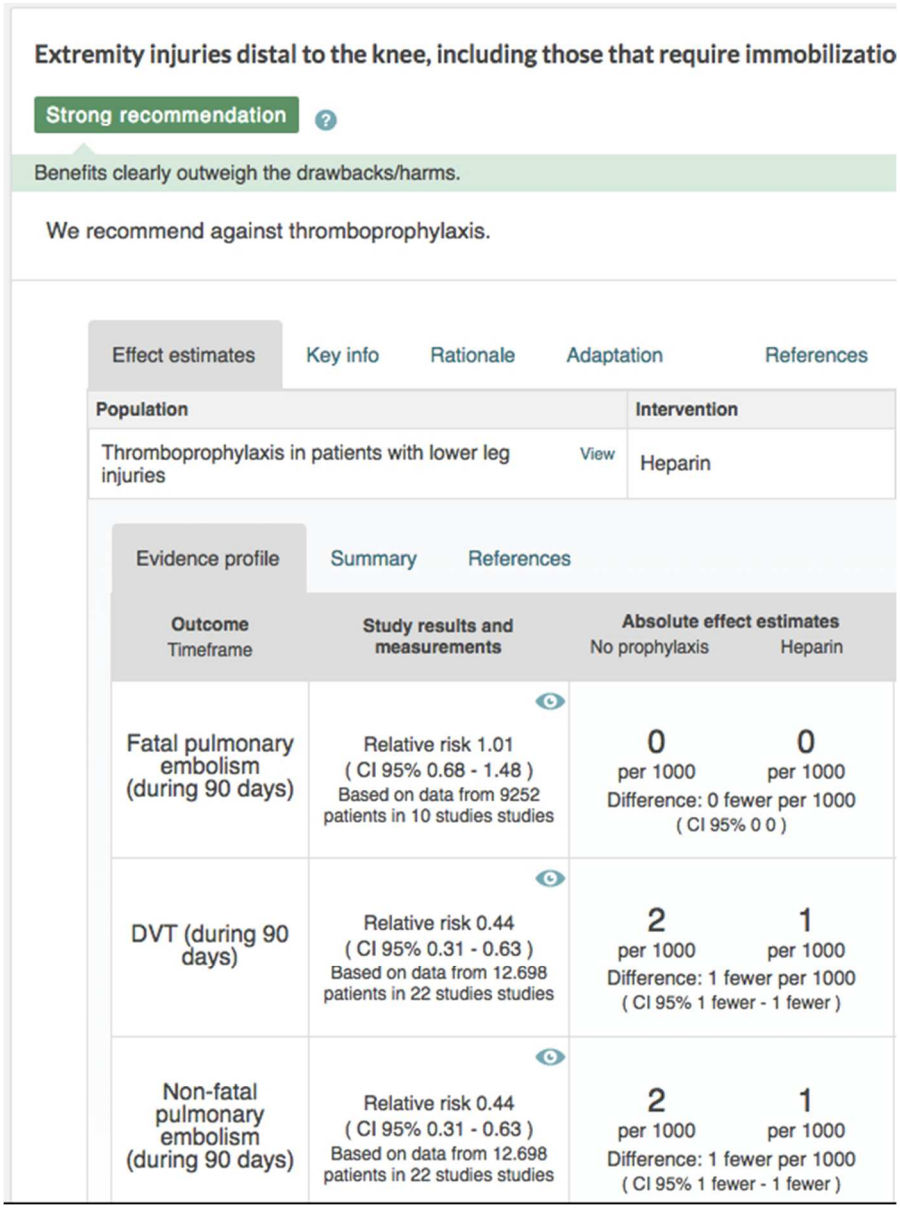
1. Bahtsevani C, Udén G, Willman A. Outcomes of evidence-based clinical practice guidelines: A systematic review. *Int J Technol Assess Health Care*. 2004;20(4):427–433.
2. Kung J, Miller RR, Mackowiak P a. Failure of clinical practice guidelines to meet institute of medicine standards: Two more decades of little, if any, progress. *Arch Intern Med*. 2012;172(21):1628–33. doi:10.1001/2013.jamainternmed.56.
3. Alonso-Coello P, Irfan A, Solà I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58. doi:10.1136/qshc.2010.042077.
4. Institute of Medicine, Graham R, Mancher M, Wolman DM. *Clinical Practice Guidelines We Can Trust*. Washington, DC: The National Academies Press.; 2011.
5. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med*. 2012;156(7):525–31. doi:10.1059/0003-4819-156-7-201204030-00009.
6. Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. *Ann Intern Med*. 2011;154(11):774–5. doi:10.7326/0003-4819-154-11-201106070-00011.
7. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6. doi:10.1136/bmj.39489.470347.AD.
8. Goldet G, Howick J. Understanding GRADE : an introduction. *J Evid Based Med*. 2013;6:50–54. doi:10.1111/jebm.12018.
9. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: Time for a map? *J Contin Educ Health Prof*. 2006;26(1):13–24. doi:10.1002/chp.47.
10. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348(jun13 4):g3725–g3725. doi:10.1136/bmj.g3725.
11. Treweek S, Oxman AD, Alderson P, et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. *Implement Sci*. 2013;8:6. doi:10.1186/1748-5908-8-6.
12. Kristiansen A, Brandt L, Alonso-Coello P, et al. Development of a novel, multilayered presentation format for clinical practice guidelines. *Chest*. 2015;147(3):754–763. doi:10.1378/chest.14-1366.
13. Kristiansen A, Brandt L, Agoritsas T, et al. Applying new strategies for the national adaptation, updating and dissemination of trustworthy guidelines: Results from the Norwegian adaptation of the American College of Chest Physicians Evidence-based Guidelines on Antithrombotic Therapy and the Preven. *Chest*. 2014;146(3):735–61. doi:10.1378/chest.13-2993.
14. Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. *Chest*. 2013;144(2):381–9. doi:10.1378/chest.13-0746.
15. www.nsth.no. :Accessed February 2016.
16. 2013 AJ. http://www.uptodate.com/contents/atrial-fibrillation-anticoagulant-therapy-to-

prevent-embolization?source=machineLearning&search=atrial+fibrillation+anticoagulation&selectedTitle=1~150&sectionRank=1&anchor=H31#H31.
17. Isaac T, Zheng J, Jha A. Use of UpToDate and outcomes in US hospitals. *J Hosp Med*. 2012;7(2):85–90. doi:10.1002/jhm.944.
18. Kwag KH, González-Lorenzo M, Banzi R, Bonovas S, Moja L. Providing Doctors With High-Quality Information: An Updated Evaluation of Web-Based Point-of-Care Information Summaries. *J Med Internet Res*. 2016;18(1):e15. doi:10.2196/jmir.5234.
19. Manrai AK, Bhatia G, Strymish J, Kohane IS, Jain SH. Medicine's UncomfortableRelationship With Math: CalculatingPositive PredictiveValue. *JAMA Intern Med*. 2014;174(6):991–993. doi:10.1001/jamainternmed.2014.1001.Author.
20. Johnston B, Alonso-Coello P, Friedrich JO, Reem MA. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ*. 2015;(November):1–8. doi:10.1503/cmaj.150430.
21. Agoritsas T, Heen AF, Brandt L, et al. Decision aids that really promote shared decision making: the pace quickens. *BMJ*. 2015;350(Feb10):g7624. doi:10.1136/bmj.g7624.
22. 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(6):620–6. doi:10.1016/j.jclinepi.2009.12.014.
23. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, 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;(In press). doi:10.1016/j.jclinepi.2015.12.007.
24. Hsiao J, Chen R. Critical factors influencing physicians ' intention to use computerized clinical practice guidelines : an integrative model of activity theory and the technology acceptance model. *BMC Med Inform Decis Mak*. 2016:1–15. doi:10.1186/s12911-016-0241-3.
25. Cuello García C a, Pacheco Alvarado KP, Pérez Gaxiola G. Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems. *Arch Dis Child*. 2011;96(8):723–728. doi:10.1136/adc.2010.199307.
26. Neumann I, Santesso N, Akl EA, et al. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol*. 2015;(c). doi:10.1016/j.jclinepi.2015.11.017.

Target journal BMJ Open



1a and 1b: Current version of the multilayered guideline presentation formats
1a
73x65mm (300 x 300 DPI)



1a and 1b: Current version of the multilayered guideline presentation formats and 1b
50x67mm (300 x 300 DPI)

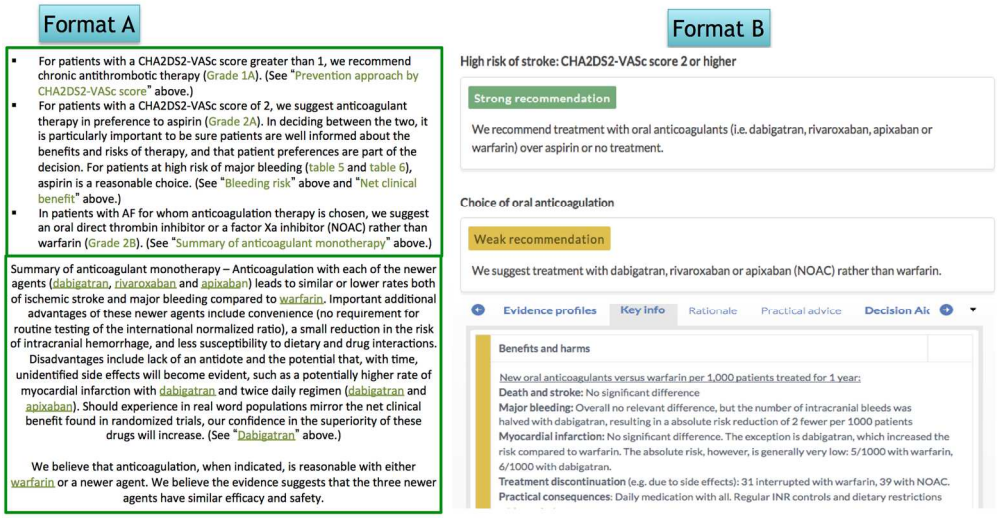


Figure 2: Standard (A) and multilayered guideline (B) presentation formats
Figure 2
173x90mm (300 x 300 DPI)

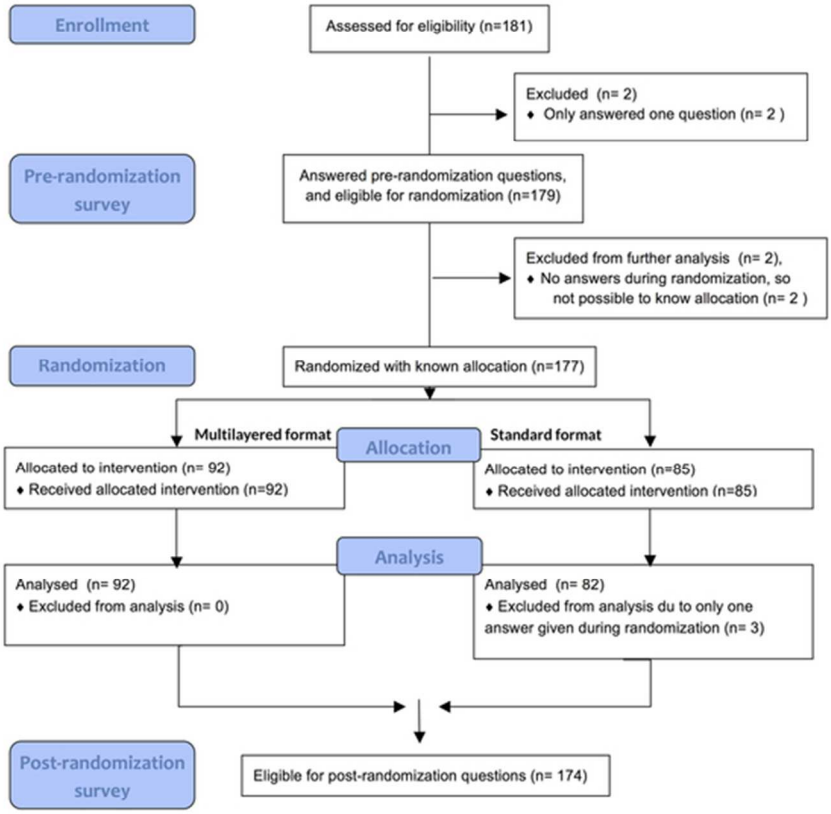


Figure 3: Flow chart of design and enrolment of participants to the multilayered format (n=92) and standard format (n=85)
Figure 3
51x66mm (300 x 300 DPI)

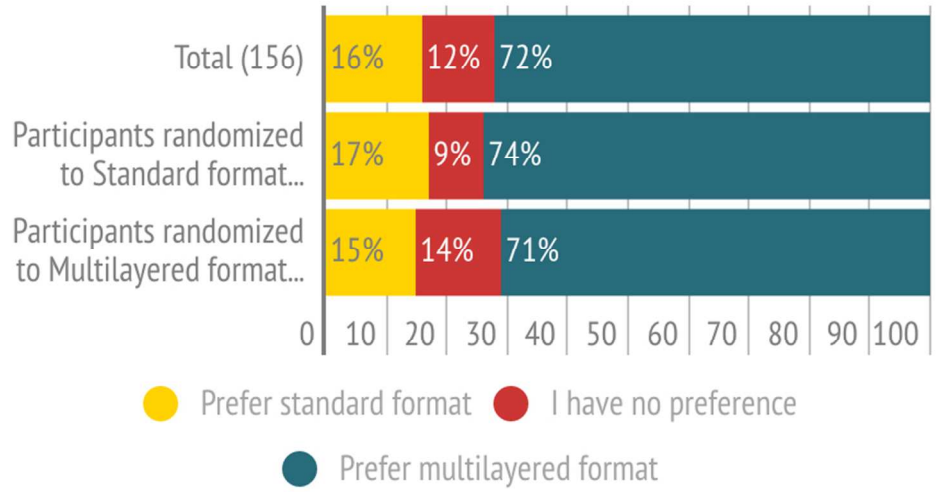


Figure 4: Preferences for standard format versus multilayered presentation format.
Figure 4
77x42mm (300 x 300 DPI)

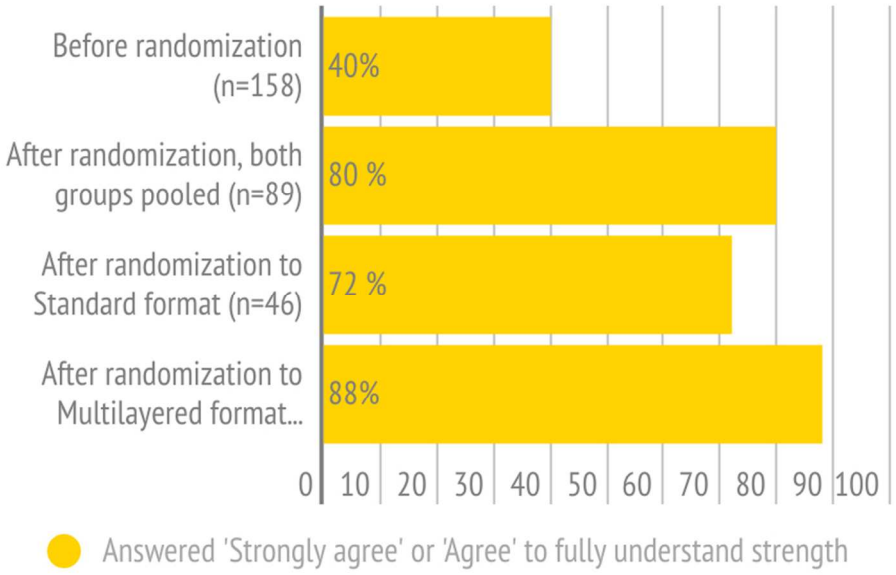


Figure 5: Reported answers to the statement: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making.

Figure 5
80x49mm (300 x 300 DPI)

DECIDE

Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence

GRADE

New standards for trustworthy guidelines

Clinical scenario:
Anticoagulation treatment
for prevention of stroke
in atrial fibrillation

CLICK-IT studies in educational sessions
DECIDE WP1 2013

1

CLICK-IT

How do clinicians like and understand trustworthy guidelines?
Mixed methods study using Clickers in educational sessions

Objectives:

- ✓ Determine understanding and preferences for guideline presentation formats
- ✓ Teach about new concepts for trustworthy guidelines

- Registered results and data will be used for research.
- We regard answering the questions is to give informed consent for us to use this in research. (You can walk out of the room now)
- The questions are in …… (if another language) , but some of the examples are in English

First, some demographic questions

2

Q 1: My age is

1. 25-35
2. 36-46
3. 46-55
4. 56-65
5. 66-100

Q 2: My position is

1. Intern, medical student
2. Resident physician
3. Consultant physician

Q3: In terms of training in health research methodology (HRM), you have:

1. Never completed a formal course in HRM or epidemiology
2. Completed one or more formal courses in HRM or epidemiology
3. A masters degree or PhD degree in HRM or epidemiology

Meet Gabriel 68y

- **Medical history:** Type 2 diabetes. No medications
- **Chief complaint:** For the past 6 months intermittent episodes of heart palpitations and rapid heart rate, duration between 30 minutes to 3 days
- **Diagnosis:** Atrial fibrillation
- No risk factors indicating increased risk of bleeding
- Risik for stroke? Anticoagulation a prophylaxis?

6

CHA2DS2-VASc Score for Atrial Fibrillation Stroke Risk

Calculates stroke risk for patients with atrial fibrillation, possibly better than the CHADS2 score.

Age?	○ < 65 years old +0 ◉ 65-74 years old +1 ○ ≥ 75 years old +2
Congestive Heart Failure History?	☐ Yes +1
Hypertension History?	☐ Yes +1
Stroke/TIA/Thromboembolism History?	☐ Yes +2
Vascular Disease History? (previous MI, peripheral arterial disease or aortic plaque)	☐ Yes +1
Diabetes Mellitus?	☑ Yes +1
Female?	☐ Yes +1
Score	2

Baseline risk for stroke (no anticoagulation treatment) with **CHA2DS2-VASc score 2**

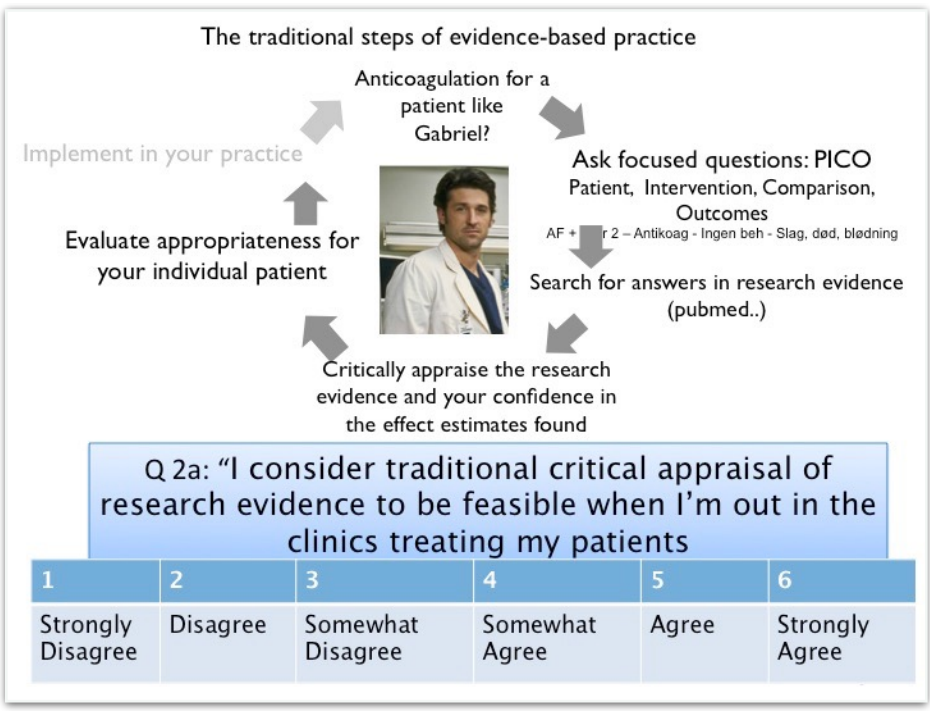
This means an assumed risk of **22 strokes per 1000** patients over 1 year

Treatment for Gabriel?

- **Diagnosis:** Atrial fibrillation
- Moderate risk of stroke (CHAD2S2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

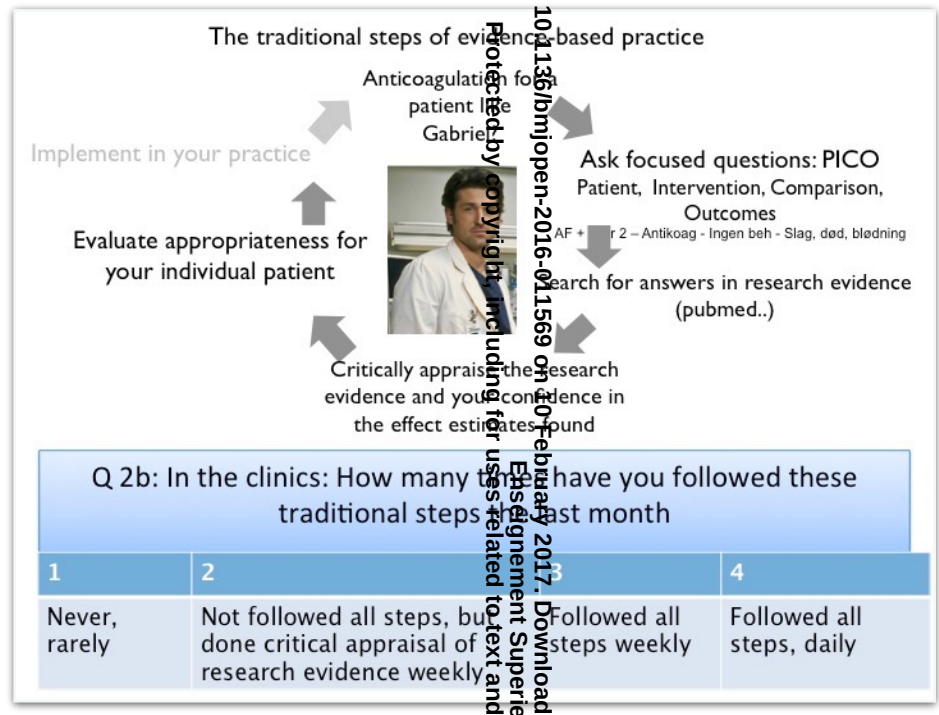
Q 1: If you were unsure of which, if any, therapy to offer the patient, where would you first look for an answer?

1. Local guideline
2. Systematic review
3. EBM textbook (e.g. UpToDate)
4. Practice guideline (national or international)
5. Ask a colleague
6. Individual study



Q 2a: "I consider traditional critical appraisal of research evidence to be feasible when I'm out in the clinics treating my patients

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree



Q 2b: In the clinics: How many times have you followed these traditional steps last month

1	2	3	4
Never, rarely	Not followed all steps, but done critical appraisal of research evidence weekly	Followed all steps weekly	Followed all steps, daily

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 3 : What does the number (2) reflect?

1	2	3	4
It´s a strong recommendation	It´s moderate quality evidence	It´s a moderate recommendation	It´s a weak recommendation

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 4 : What does the letter (B) reflect?

1	2	3	4
Moderate quality evidence	Low quality evidence	Study design (based on a single Randomized study)	Study design (based on a smaller systematic review)

Several guidelines and EBM textbooks (e.g. UpToDate) use the GRADE system and label their recommendations with a number + letter.

We suggest that older patients receive supplementation with vitamin D3 (cholecalciferol) GRADE 2B

Q 5: GRADE provides either strong or weak recommendations. To what extent do you agree with the following statement:

"I fully understand the difference between strong and weak recommendations and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Now, let's get back to Gabriel

1. Two groups, 5 questions each
2. Different formats of guidelines for atrial fibrillation and anticoagulation
3. One group gets blindfolds (they are "blinded")
4. I will not read the questions or text out loud.
5. Read the text and give me a sign (waive/ raise you hand) when you are ready to answer
6. Then switch and the other group gets blindfolds

14

Imagine you search online for an answer to what to do with Gabriel,

and you found the guideline on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

15

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 7 : How do you interpret these recommendations?

1	2	3	4
Strong recommendation for NOAC. Weak recommendation for warfarin or aspirin.	Weak recommendation for NOAC and warfarin.	Weak recommendation for any option.	Strong recommendation for treatment. Weak recommendation for NOAC over warfarin.

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 6: "These recommendations will help me manage my patient"

1	2	3	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Agree	Strongly Agree

You also find the summary you will see on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

18

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and safety.

Q 8 : "This information helps me apply the recommendation on my patient"

1	2	3	4	5	6
Strongly disagree	Disagree	Somewhat disagree	Somewhat Agree	Agree	Strongly Agree

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and safety.

Q 9 : What does this information tell you about NOAC vs warfarin?

1	2	3	4
Vastly superior treatment effect	Less burden of treatment and slightly better treatment effect	Large reduction in side effects	No difference in effect or side effects

How would you have treated Gabriel?

- Diagnosis: Atrial fibrillation
- Moderate risk of stroke (CHA2DS2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

Let's look at the recommendations again

21

- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 1, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Q 10 : Which, if any, antithrombotic treatment would you consider appropriate for Gabriel?

1	2	3	4
NOAC (Dabigatran, rivaroxaban or apixaban)	Aspirin	Warfarin	No therapy

Now give your blindfold to an unblinded colleague

23

Imagine you search online for an answer to what to do with Gabriel,

and you found the guideline on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

24

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 11: "The recommendations will help me manage my patient"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 12 : How do you interpret these recommendations?

1	2	3	4
Strong recommendation for NOAC. Weak recommendation for warfarin or aspirin.	Weak recommendation for NOAC and warfarin.	Weak recommendation for any option.	Strong recommendation for any treatment. Weak recommendation for NOAC over warfarin.

When you click one of the recommendations you find the summary you will see on next slide!

Read through the text first and you'll get some questions later.

The questions will always come together with the text so there is no need to memorize!

27

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but there was seen a halving of the number intracranial bleeds with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high

Q 13 : "This information helps me apply the recommendation on my patient"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but there was seen a halving of the number intracranial bleeds with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high

Q 14 : What does this information tell you about NOAC vs warfarin?

1	2	3	4
Vastly superior treatment effect	Less burden of treatment and slightly better treatment effect	Large reduction in side effects	No difference in effect or side effects

How would you have treated Gabriel?

- Diagnosis: Atrial fibrillation
- Moderate risk of stroke (CHA2DS2-VASc score: 2)
- Low risk of bleeding
- Currently no antithrombotic treatment

Let's look at the recommendations again

30

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Q 15 : Which, if any, antithrombotic treatment would you consider appropriate for Gabriel?

1	2	3	4
NOAC (Dabigatran, rivaroxaban or apixaban)	Aspirin	Warfarin	No therapy

Now, let us all get unblinded and have a look at both formats for a few minutes

32

Format A

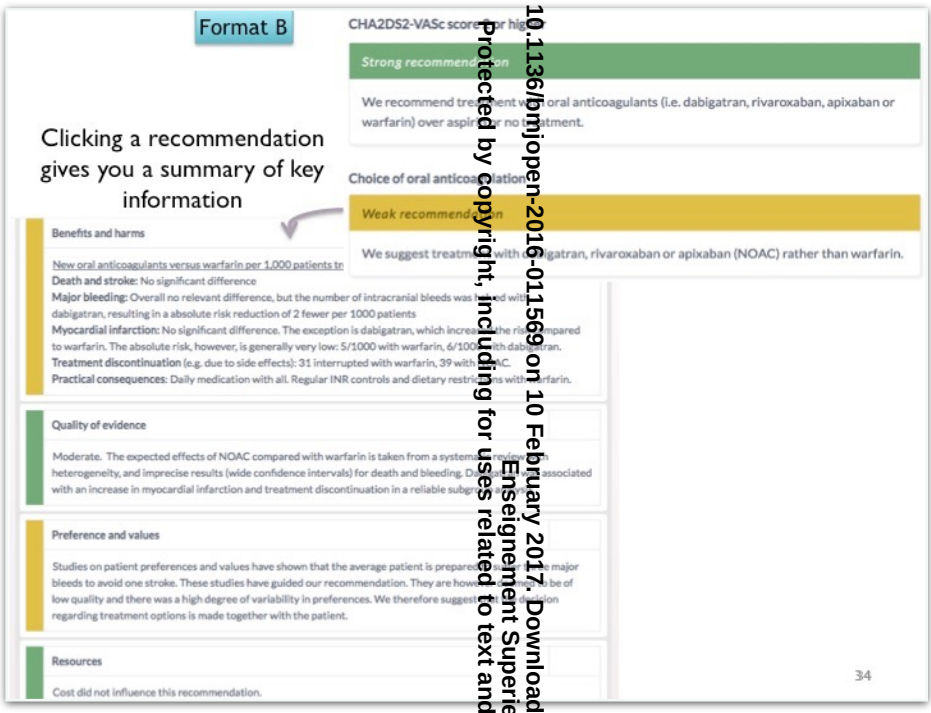
- For patients with a CHA2DS2-VASc score greater than 1, we recommend chronic antithrombotic therapy (Grade 1A). (See "Prevention approach by CHA2DS2-VASc score" above.)
- For patients with a CHA2DS2-VASc score of 2, we suggest anticoagulant therapy in preference to aspirin (Grade 2A). In deciding between the two, it is particularly important to be sure patients are well informed about the benefits and risks of therapy, and that patient preferences are part of the decision. For patients at high risk of major bleeding (table 5 and table 6), aspirin is a reasonable choice. (See "Bleeding risk" above and "Net clinical benefit" above.)
- In patients with AF for whom anticoagulation therapy is chosen, we suggest an oral direct thrombin inhibitor or a factor Xa inhibitor (NOAC) rather than warfarin (Grade 2B). (See "Summary of anticoagulant monotherapy" above.)

Summary of anticoagulant monotherapy – Anticoagulation with each of the newer agents (dabigatran, rivaroxaban and apixaban) leads to similar or lower rates both of ischemic stroke and major bleeding compared to warfarin. Important additional advantages of these newer agents include convenience (no requirement for routine testing of the international normalized ratio), a small reduction in the risk of intracranial hemorrhage, and less susceptibility to dietary and drug interactions. Disadvantages include lack of an antidote and the potential that, with time, unidentified side effects will become evident, such as a potentially higher rate of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real word populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See "Dabigatran" above.)

We believe that anticoagulation, when indicated, is reasonable with either warfarin or a newer agent. We believe the evidence suggests that the three newer agents have similar efficacy and

Format B

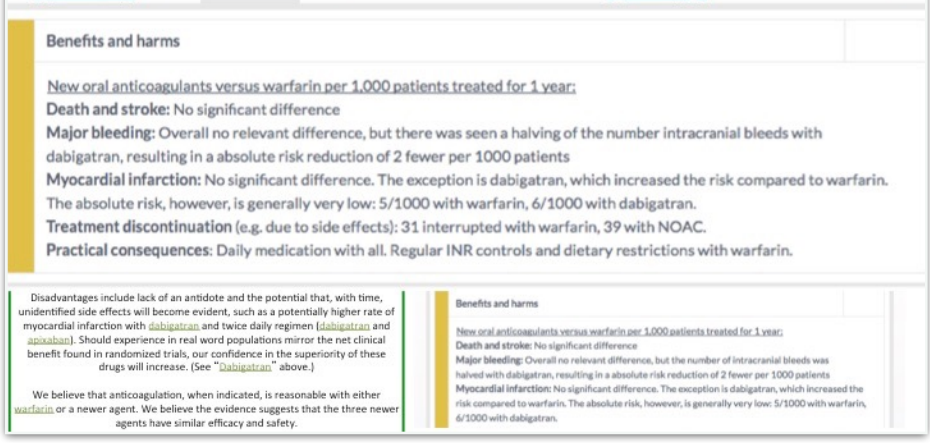
Clicking a recommendation gives you a summary of key information



34

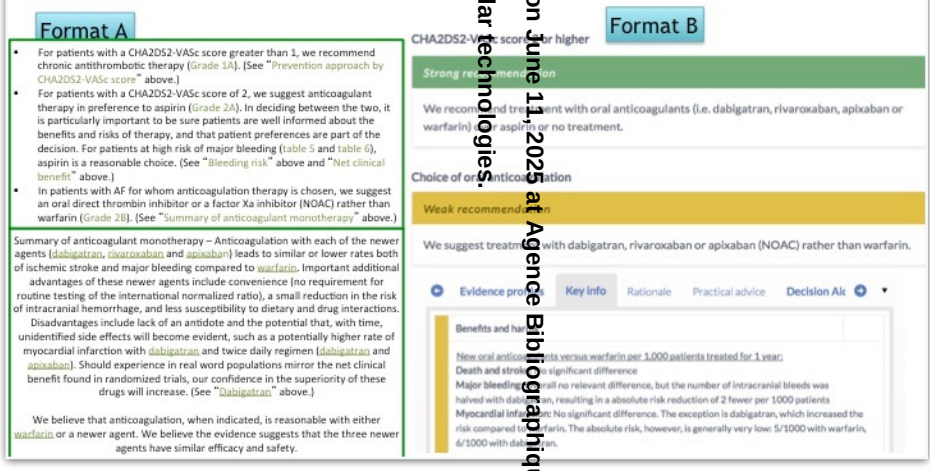
Q 16: Format B presents absolute effects for benefits and harms, whereas format A does not.
What is your first reaction to being presented with the absolute effects?

1	2	3	4	5
Confusing distraction, leave it out	A little confusing but not a big problem	Doesn´t help but doesn´t hurt	Not crucial but helpful	Crucial information, should always be included



Q 17: Overall, do you prefer format A or format B?

1	2	3
Format A	I have no preference	Format B



Understand new definitions and standards for trustworthy guidelines

"Clinical Practice Guidelines are statements that include **recommendations** intended to optimize patient care. They are informed by a **systematic review** of evidence and an assessment of the **benefits and harms of alternative** care options"(2011)

CLINICAL PRACTICE GUIDELINES WE CAN TRUST

Broad consensus

GRADE

Common methods for guidelines

37

GRADE defines **strength** of recommendation as:

"The strength of a recommendation reflects the extent to which we can, across the range of patients for whom the recommendations are intended, be confident that desirable effects of a management strategy outweigh undesirable effects."

Strong recommendations:
Reflects clear benefit of the recommended treatment alternative.
Implications: Recommendation applies to all or nearly all patients. "Just do it"

Weak recommendations:
Reflects fine balance between benefits and harms for the treatment alternatives.
Implications: Recommendation applies to the majority of patients.
"Maybe", "Depends on patient values and preferences"

Q18a: "I fully understand the difference between strong and weak recommendations and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

GRADE defines **strength** by always considering 4 factors

Strong recommendations:
Reflects clear benefit of the recommended treatment alternative.
Applies to all or nearly all patients. "Just do it"

Weak recommendations:
Reflects fine balance between benefits and harms .
The majority of patients.
"Maybe", "Depends on patient values and preferences"

Process it using the 4 factors

- B&H (Benefits & Harms): What are the expected benefits and harms?
- QoE (Quality of Evidence): What are the overall confidence in these estimates?
- V&P (Values & Preference): What are patients values and preferences? Do they vary? Do you know? Important societal values to consider?
- R (Resources): Are there resource issues to consider? For whom?

Q18b: "This explanation is necessary to understand the difference between strong and weak recommendations and the implications for clinical decision

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

GRADE defines **quality of evidence** as:

High quality: We are very confident that the true effect lies close to our effect estimates

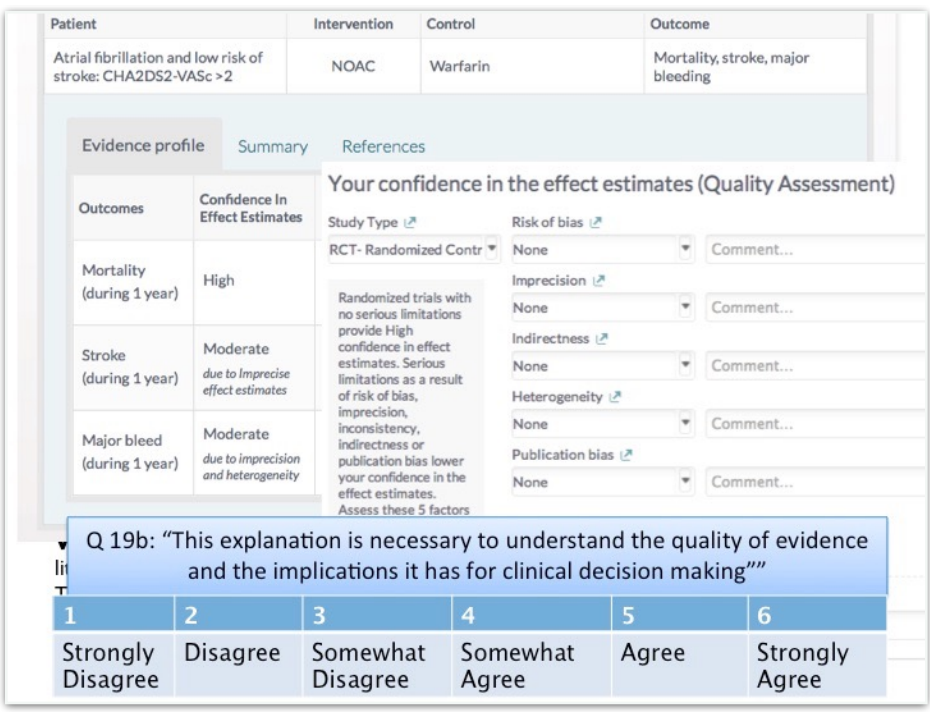
Moderate quality: We are moderately confident in our effect estimates. The true effect is likely to be close to our effect estimates, but with the possibility to be substantially different.

Low quality: Our confidence in the effect estimates are limited. The true effect may be substantially different from our effect estimates.

Very low quality: We have very little confidence in our effect estimates. The true effect is likely to be substantially different from our effect estimates.

Q 19a: "I fully understand the difference between the different categories of quality and the implications for clinical decision making"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree



Q 19b: "This explanation is necessary to understand the quality of evidence and the implications it has for clinical decision making""

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

There is a lot of information included in a guideline

Do we need to see it all, all the time?

How do we like multilayered guidelines?

42

CHA2DS2-VASc score 2 or higher

Strong recommendation

We recommend treatment with oral anticoagulants (i.e. dabigatran, rivaroxaban, apixaban or warfarin) over aspirin or no treatment.

Choice of oral anticoagulation

Weak recommendation

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

43

Choice of oral anticoagulation

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Effect estimates | **Key info** | Rationale | Practical advice | **Adaptation** | Reference

The new oral anticoagulants have equal effect to warfarin with regards to stroke reduction, they lower the incidence of intracranial bleeds and are more convenient to use. We therefore suggest the new oral anticoagulants over warfarin as first treatment of choice.
For patients that are already on warfarin therapy with stable INR values the cost/benefit ratio is similar to treatment with NOACs. We therefore suggest that patients well-established on warfarin therapy continue with this if they wish.

44

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Benefits and harms

New oral anticoagulants versus warfarin per 1,000 patients treated for 1 year:
Death and stroke: No significant difference
Major bleeding: Overall no relevant difference, but the number of intracranial bleeds was halved with dabigatran, resulting in a absolute risk reduction of 2 fewer per 1000 patients
Myocardial infarction: No significant difference. The exception is dabigatran, which increased the risk compared to warfarin. The absolute risk, however, is generally very low: 5/1000 with warfarin, 6/1000 with dabigatran.
Treatment discontinuation (e.g. due to side effects): 31 interrupted with warfarin, 39 with NOAC.
Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

Quality of evidence

Moderate. The expected effects of NOAC compared with warfarin is taken from a systematic review with heterogeneity, and imprecise results (wide confidence intervals) for death and bleeding. Dabigatran was associated with an increase in myocardial infarction and treatment discontinuation in a reliable subgroup analysis.

Preference and values

Studies on patient preferences and values have shown that the average patient is prepared to suffer three major bleeds to avoid one stroke. These studies have guided our recommendation. They are however deemed to be of low quality and there was a high degree of variability in preferences. We therefore suggest that the decicion regarding treatment options is made together with the patient.

Resources

Cost did not influence this recommendation.

45

Weak recommendation

It is less clear whether the benefits outweigh the drawbacks. We believe there will be variation in patients preferences

We suggest treatment with dabigatran, rivaroxaban or apixaban (NOAC) rather than warfarin.

Effect estimates | **Key info** | Rationale | Practical advice | **Adaptation** | Reference

Patient	Intervention	Control	Outcome
Atrial fibrillation and low risk of stroke: CHA2DS2-VASc >2	NOAC	Warfarin	Mortality, stroke, major bleeding

Evidence profile | Summary | References

Outcomes	Confidence In Effect Estimates	Relative Effect	Warfarin	NOAC	Absolute Difference	Participants (Studies), Follow-Up
Mortality (during 1 year)	High	RR 0.88 (CI 0.82 - 0.96)	7 per 1000	6 per 1000	1 fewer per 1000 (CI 1 fewer - 0 fewer)	44.442 (3), 2 years
Stroke (during 1 year)	Moderate due to imprecise effect estimates	RR 0.89 (CI 0.78 - 1.02)	3 per 1000	3 per 1000	0 fewer per 1000 (CI 1 fewer - 0 fewer)	44.442 (3), 2 years
Major bleed (during 1 year)	Moderate due to imprecision and heterogeneity	RR 0.8 (CI 0.63 - 1.01)	25 per 1000	20 per 1000	5 fewer per 1000 (CI 9 fewer - 0 fewer)	44.442 (3), 2 years

46

"This is a smart layout for guidelines"

1	2	3	4	5	6
Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree

Thank you for your participation

For more information go to:

http://www.decide-collaboration.eu

http://www.gradeworkinggroup.org

http://www.magicproject.org

If there is time left we can look at your results

48

CONSORT 2010 checklist of information to include when reporting a randomised trial*

The page numbers refer to the main manuscript "Clean copy" unless otherwise stated.

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	p1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	p2 –p3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	p4 - p5
	2b	Specific objectives or hypotheses	p5 (second paragraph)
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	p6 (first paragraph). Consort Flowchart Figure 3
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	No changes to methods, described p6 (first paragraph)
pParticipants	4a	Eligibility criteria for participants	p6 (first paragraph)
	4b	Settings and locations where the data were collected	p6 (first paragraph)
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	p6 (last paragraph) –p7 + Figure 2 + attached lecture
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	p8 (last paragraph)- p9
	6b	Any changes to trial outcomes after the trial commenced, with reasons	No changes to outcomes, described p6 (first paragraph)
Sample size	7a	How sample size was determined	Power calculation, in protocol
	7b	When applicable, explanation of any interim analyses and stopping guidelines	Not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	p8 (second paragraph)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	p8 (second paragraph), no restrictions.

Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	p8 (second paragraph)
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Generated allocation sequence and assigned participants: p8 (second paragraph), Enrolled participants: p6 (first paragraph)
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	p8 (second paragraph)
	11b	If relevant, description of the similarity of interventions	p7 (second paragraph) -p8 (first paragraph)
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	P9 (last paragraph)
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	No additional analysis
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	p10 and Figure 3 flow chart
	13b	For each group, losses and exclusions after randomisation, together with reasons	p10 and Figure 3 flow chart
Recruitment	14a	Dates defining the periods of recruitment and follow-up	p10
	14b	Why the trial ended or was stopped	Stopped as planned after all centers had performed the study
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	p10, Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 1 p10 and Figure 3 p11
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	p11-12
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	p11-12
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	No additional analysis performed
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	n/a
Discussion			

Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	p14 (second paragraph)
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	p13 (last paragraph) –p14
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	p13 (first paragraph)
Other information			
Registration	23	Registration number and name of trial registry	None
Protocol	24	Where the full trial protocol can be accessed, if available	Appendix
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	p16 (Funding statement)

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.