Dissemination and adaptation strategies customized for trustworthy practice guidelines using the GRADE framework
Annette Kristiansen
# Table of Contents

ACKNOWLEDGEMENTS .......................................................................................................................... 4

LIST OF ABBREVIATIONS ...................................................................................................................... 6

LIST OF PUBLICATIONS .......................................................................................................................... 7

SUMMARY ............................................................................................................................................. 9

1. INTRODUCTION ............................................................................................................................... 11
   1.1 CLINICAL PRACTICE GUIDELINES ................................................................................................. 11
   1.1.1 Definitions and standards for trustworthy guidelines ............................................................... 11
   1.2 THE GRADING OF RECOMMENDATIONS ASSESSMENT, DEVELOPMENT AND EVALUATION (GRADE) FRAMEWORK .......................................................................................................................... 11
   1.2.1 Phrasing clinical questions and performing evidence assessments ......................................... 12
   1.2.2 Going from evidence to recommendation ................................................................................. 13
   1.3 KNOWLEDGE TRANSLATION ......................................................................................................... 14
   1.3.1 Definitions and purpose ............................................................................................................... 14
   1.3.2 Knowledge to action process framework ..................................................................................... 15
   1.3.3 Dissemination strategies ............................................................................................................. 16
   1.3.4 Guideline implementability ......................................................................................................... 17
   1.3.5 Communicating uncertainty ........................................................................................................ 17
   1.3.6 Adaptation of guidelines ............................................................................................................ 18
   1.4 ADAPTATION FRAMEWORKS: A SYSTEMATIC REVIEW ............................................................. 18

2. AIMS AND OBJECTIVES .................................................................................................................. 23

3. MATERIALS AND METHODS .......................................................................................................... 24
   3.1 GENERAL OVERVIEW .................................................................................................................... 24
   3.2 RELATED RESEARCH PROJECTS .................................................................................................. 24
      3.2.1 Making Grade the Irresistible Choice (MAGIC) ..................................................................... 24
      3.2.2 The DECIDE project ................................................................................................................ 25
   3.3 MULTILAYERED PRESENTATION FORMAT .................................................................................. 25
      3.3.1 Development of a new presentation format ........................................................................... 25
      3.3.2 Randomized controlled trial of new format vs. current standard ......................................... 27
   3.4 DEVELOPMENT OF AN ADAPTATION PROCESS ....................................................................... 28
   3.5 APPLYING NOVEL DISSEMINATION AND ADAPTATION STRATEGIES THROUGH THE NORWEGIAN ADAPTATION OF AT9 .............................................................................................................. 29
   3.6 ETHICAL CONSIDERATIONS .......................................................................................................... 31

4. SUMMARY OF RESULTS ................................................................................................................... 32
   4.1 PAPER I: DEVELOPMENT OF A NOVEL, MULTILAYERED PRESENTATION FORMAT FOR CLINICAL PRACTICE GUIDELINES .................................................................................................................. 32
   4.2 PAPER II: MULTILAYERED AND DIGITALLY STRUCTURED PRESENTATION FORMATS OF TRUSTWORTHY RECOMMENDATIONS: A RANDOMIZED TRIAL ................................................................. 32
   4.3 PAPER III: ADAPTATION OF TRUSTWORTHY GUIDELINES USING THE GRADE METHODOLOGY: A NOVEL FIVE-STEP PROCESS .................................................................................................................. 34
   4.4 PAPER IV: APPLYING NEW STRATEGIES FOR THE NATIONAL ADAPTATION, UPDATING AND DISSEMINATION OF TRUSTWORTHY GUIDELINES ...................................................................................... 36
ACKNOWLEDGEMENTS

I wish to thank the panel members of the Norwegian Antithrombotic Guideline for their hard work and generous contribution to the pilot testing of not only a new adaptation framework, but also a new way of authoring guidelines. My newfound appreciation of baseline estimates and fondness for weak recommendations was occasioned by work with the panelists, benefitting my current and future patients.

I am thankful to my main supervisor Dr. Per Olav Vandvik for introducing me to the colorful world of evidence-based medicine in his own pedagogical way. He has not shied away from challenging me with new responsibilities, and has not only learned me to be acutely critical of any assertions, however confidentially delivered, but to attempt to think of new and better ways of facing current challenges in our everyday clinical life. These notions also extend to the rest of the core MAGIC team; Linn Brandt, Gordon Guyatt, Thomas Agoritsas, Anja Fog Heen and Christopher Friis Berntsen. Their collected enthusiasm, cunning knowledge, creativity and fondness of cider has made for an interesting and engaging few years as I have completed my PhD education.

My work would not have been possible without the explicit support of Hospital Innlandet Trust, specifically the Department of Medicine at Gjøvik Hospital. I would especially like to thank Øystein Stubhaug, Even Reinertsen and Mette Farstad for their moral, financial and logistical support throughout the project.

Many more have provided invaluable support throughout. My co-supervisors Drs. Guyatt and Sandset have played an incremental part in educating me in all things evidence and thrombosis related. Sarah Rosenbaum has taught me about the advantages of combining design and research. Thank you to all members of the DECIDE work package 1 for being inspiring and welcoming research partners. And to my colleagues at Diakonhjemmet Hospital for keeping me grounded in the realities of day-to-day clinical life.

Thank you to the Norwegian Society of Hematology for enabling me to join the project at a critical time, with no guaranties that I would be able to follow through until the end. The South-Eastern Norway Regional Health Authority funded my further participation as a PhD student.
Lastly, a heartfelt thanks to my family and friends for their encouragement, welcomed distraction and all-round awesomeness.
LIST OF ABBREVIATIONS

ACCP = American College of Chest Physicians
AHRQ = The Agency for Healthcare Research and Quality
AT9 = Antithrombotic Therapy and Prevention of Thrombosis, 9th ed.: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines
DECIDE = Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence
DVT = Deep Venous Thrombosis
EBM = Evidence Based Medicine
GRADE = Grading of Recommendations Assessment, Development and Evaluation
IOM = Institute of Medicine
KT = Knowledge Translation
MAGIC = Making Grade the Irresistible Choice
NGC = National Guideline Clearing House
NSTH = Norwegian Society of Thrombosis and Hemostasis
SoF = Summary of Findings table
PICO = Patient, Intervention, Comparator, Outcomes
VTE = Venous thromboembolic disease
LIST OF PUBLICATIONS

I. Development of a Novel, Multilayered Presentation Format for Clinical Practice Guidelines.
Annette Kristiansen, MD; Linn Brandt, MD; Pablo Alonso-Coello, MD, PhD; Thomas Agoritsas, MD; Elie A. Akl, MD, PhD, MPH; Tara Conboy, RGN, MSc; Mahmoud Elbarbary, MD, PhD; Mazen Ferwana, MD, PhD; Wedad Medani, MSc; Mohammad Hassan Murad, MD, MPH; David Rigau, MD; Sarah Rosenbaum, PhD; Frederick A. Spencer, MD; Shaun Treweek, PhD; Gordon Guyatt, MD, FCCP; and Per Olav Vandvik, MD, PhD.
CHEST 2015; 147(3):754-763

II. Multilayered and digitally structured presentation formats of trustworthy recommendations: a randomised trial
Linn Brandt, Per Olav Vandvik, Pablo Alonso-Coello, Elie A. Akl, Judith Thornton, Paul O’Connor, Katie Adams, David Rigau, Gordon Guyatt, Annette Kristiansen.
(Submitted February 17th 2016)

III. Adaptation of Trustworthy Guidelines Developed Using the GRADE Methodology.
A Novel Five-Step Process.
Annette Kristiansen, MD; Linn Brandt, MD; Thomas Agoritsas, MD; Elie A. Akl, MD, PhD, MPH; Eivind Berge, MD, PhD; Johan Bondi, MD, PhD; Anders E. Dahm, MD, PhD; Lars-Petter Granan, MD, PhD; Sigrun Halvorsen, MD, PhD; Pål-Andre Holme, MD, PhD; Anne Flem Jacobsen, MD, PhD; Eva-Marie Jacobsen, MD, PhD; Ignacio Neumann, MD; Per Morten Sandset, MD, PhD; Torunn Sætre, MD, PhD; Arnljot Tveit, MD, PhD; Trond Vardal, MD, PhD; Gordon Guyatt, MD, FCCP; and Per Olav Vandvik, MD, PhD.
CHEST 2014; 146(3):727-734

IV. Applying New Strategies for the National Adaptation, Updating, and Dissemination of Trustworthy Guidelines.
Results From the Norwegian Adaptation of the Antithrombotic therapy and the Prevention of thrombosis, 9th Ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines.
Annette Kristiansen, MD; Linn Brandt, MD; Thomas Agoritsas, MD; Elie A. Akl, MD, PhD, MPH; Eivind Berge, MD, PhD; Anne Flem Jacobsen, MD, PhD; Lars-Petter Granaan, MD, PhD; Sigrun Halvorsen, MD, PhD; Gordon Guyatt, MD, FCCP; and Per Olav Vandvik, MD, PhD.

CHEST 2014; 146(3):735-761
SUMMARY

Background
Over the last few decades, the world of health care has seen a growing productivity, amassing an increasing amount of research evidence. Clinical practice guidelines incorporating not only the best available research evidence, but also clinical experience and patient preferences can provide a potential solution to the resulting message fatigue. Unfortunately they often suffer from subpar rigor of development, applicability and usability. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework provides a comprehensive system for development of trustworthy guidelines. GRADE has yet to provide specific solutions on how to communicate and disseminate recommendations to narrow the evidence to practice gap.

This thesis aimed to 1) create and test a novel guideline presentation format (paper I & II), 2) develop and evaluate a new adaptation framework (paper III), and 3) apply the new strategies for dissemination and adaptation through development of a Norwegian guideline for antithrombotic therapy (paper IV).

Material and methods
This project is part of the research and innovation program MAGIC (MAking GRADE the Irresistible Choice), which applies a holistic approach to improving the development, dissemination and updating of guidelines. In a collaborative effort between the MAGIC organization and the DECIDE (Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence) project funded by the European Union, we applied a mixed-methods approach to develop and test new guideline presentation formats. We performed multiple cycles of brainstorming, sketching and usability testing (paper I). We undertook a randomized, controlled trial by use of clickers to vet our new format against a standard format, examining the outcomes of preference, perceived understanding, usefulness and intended clinical action (paper II).

Informed by current standards for trustworthy guidelines and the widely applied ADAPTE framework, an international group of clinicians and experts in health research methodology devised an adaptation framework for evaluating and modifying individual recommendations (paper III).

We applied and evaluated the strategies for dissemination and adaptation by developing a guideline for antithrombotic therapy (paper IV). We elected the 9th iteration of the American College of
Chest Physicians (ACCP) Evidence-based Guidelines for Antithrombotic Therapy (AT9) as the source guideline due to its currency and stringent application of GRADE.

**Results**

We created a new digitally structured, multilayered presentation format for communicating recommendations to end-user clinicians. Clinicians are presented with the recommendations first, together with the strength of the recommendation. By accessing the recommendation, the clinician can delve into deeper layers of relevant information, such as the key information (certainty and magnitude of the net benefit of the considered intervention, patient preferences and values, resources and other feasibility considerations), rationale, practical information, evidence summary tables, references and adaptation disclaimers. We randomized 181 practicing physicians in five countries to either the multilayered format or a current standard presentation format. 72% of the participants preferred the multilayered format and a majority expressed that they found recommendations useful. Both during usability testing and in the randomized trial we observed common struggles in comprehending key concepts of uncertainty that could be partly alleviated by clearer phrasing and short explanations.

We developed, evaluated and revised a five-step adaptation process and accompanying taxonomy for modifying recommendations tailored to the GRADE methodology. The steps include 1) planning, 2) individual assessment, 3) modifications, 4) publication and 5) evaluation.

We applied the five-step adaptation process with 46 recruited clinicians adapting the original AT9 guideline to the Norwegian setting, rewriting the recommendations in the multilayered format. We excluded 30 (9%), modified 131 (39%) of the original recommendations and developed eight new recommendations. The most common reason for modification was feasibility issues (e.g. availability of drugs). In November 2013, we published 249 adapted and updated recommendations on the web-based authoring and publication platform (MAGICapp). The adapted guideline can be accessed on any electronic device at www.nsth.no.

**Conclusion**

In conclusion, we developed, iteratively revised, tested and demonstrated superior preference for a novel, multilayered guideline presentation format. We also developed a five-step adaptation process and accompanying taxonomy and demonstrated the feasibility of both strategies by applying them in the making of a Norwegian antithrombotic guideline.
1 INTRODUCTION

1.1 Clinical practice guidelines

Ideally, patients are correctly diagnosed by well-informed health care professionals (also referred to as clinicians) and offered safe and effective treatment based on best current research evidence. Clinicians usually face an abundance of questions concerning diagnosis, treatment and prognosis during any given work day, the majority of which remain unanswered. Much clinical practice is thus based solely on tradition passed on from senior to junior colleagues, which might contribute to preventable harm and resource waste. Given the vast amount of available research evidence, with more than 1 200,000 citations indexed in Medline in 2015, including over 35,000 randomized trials of varying quality, clinicians need help to identify best current research evidence and guidance to make well informed decisions at the point of care. Knowledge summaries can be a necessary, preferable and more accessible option for acquiring answers, however they suffer from frequently being difficult to locate, access and use.

1.1.1 Definitions and standards for trustworthy guidelines

According to international and national standards for trustworthy guidelines, clinical practice guidelines are defined as « 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 ». Guidelines set a professional standard for diagnosis and treatment to promote quality, good practice and equity in care. They are not legally binding, and clinicians are still required to show professional judgment and take into account the individual patient's needs when applying the recommendations. Table 2 provides an overview of the trustworthy standards defined by the Institute of Medicine.

1.2 The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework

The GRADE framework provides a systematic and transparent approach to evidence assessment and guideline development, facilitating adherence to standards for evidence based practice. The GRADE working group (www.gradeworkinggroup.org) is an international group of
epidemiologists, health care professionals and guideline developers. The working group formed under the GRADE name back in 2000 and since then, the GRADE system has been adopted by more than 90 organizations worldwide, including the Norwegian Directorate of Health. The GRADE working group continues to grow and evolve. The framework is embedded in a web based Guideline Development Tool (GDT at www.guidelinedevelopment.org) for summarizing evidence and presenting recommendations. There is continuously ongoing development and research on evidence based methodological issues; such as, but not limited to, diagnostic tests, network meta-analysis and evidence to decision frameworks.

The five steps to making GRADE guidelines are: Phrase clinical questions in the PICO format, systematically identify and assess relevant evidence, evaluate the net benefit and phrase graded recommendations (figure 1).

Figure 1: Overview of evidence to recommendation process. Illustration adapted from the original by Holger Schunemann and Yngve Falck Ytter.

1.2.1 Phrasing clinical questions and performing evidence assessments
A guideline panel starts off by defining their clinical questions according to the Population, Intervention, Comparator and the Outcomes (PICO format). The chosen outcomes are preferably
limited to patient important outcomes; i.e. that surrogate outcomes (such as evaluating change in the value of HbA1c as opposed to diabetic neuropathy) are avoided as far as possible.

Next, they systematically and transparently assess the current best evidence, typically represented by a systematic review. Critical appraisal follows predefined assessment criteria: risk of bias, imprecision, indirectness, heterogeneity and publication bias.\textsuperscript{27-30} Observational studies are always considered to be of low quality due to the inherent risk of bias, but can be upgraded if there is a large magnitude of effect, all plausible confounding factors would serve to decrease an apparent effect or there is a clear dose-response gradient.\textsuperscript{31} GRADE uses four levels of confidence in the effect estimates, as outlined in table 1.\textsuperscript{32}

<table>
<thead>
<tr>
<th>HIGH</th>
<th>We are very confident that the true effect lies close to that of the estimate of the effect.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODERATE</td>
<td>We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.</td>
</tr>
<tr>
<td>LOW</td>
<td>Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.</td>
</tr>
<tr>
<td>VERY LOW</td>
<td>We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.</td>
</tr>
</tbody>
</table>

Table 1: Quality of the evidence as defined by GRADE

1.2.2 Going from evidence to recommendation

Lastly, guideline panels move from evidence to development of recommendations. Guideline developers determine the direction and strength of each recommendation through the integration of the following factors:\textsuperscript{25,26}

- The magnitude and balance between the benefits (desirable consequences) and harms (undesirable consequences) of the considered treatment options.
- Quality of the evidence (confidence in effect estimates).
- Assumed or confirmed typical patient values and preferences.
- Resources use, acceptability and feasibility.

There are two levels of recommendations; strong and weak/conditional. GRADE defines the strength of the recommendations as follows:\textsuperscript{33}
**Strong recommendation:** Strong recommendations reflect a high confidence that the potential benefits outweigh the harms. The implications of a strong recommendation are that all, or nearly all, patients would want to comply with the recommended treatment, clinicians would agree that virtually all patients should receive the recommended treatment, and the recommendation should be routinely applied and is suitable as a quality indicator.

**Weak recommendation:** Weak recommendations reflect that there is either a close balance between benefits and harms, uncertainty regarding the magnitude of the net benefit or that the cost or burden is questionably justified. The implications of a weak recommendation are that most patients would still want the recommended treatment, but some will probably opt against it. Clinicians should recognize that different choices will be appropriate for different patients and that they must help patients make a decision in accordance with his or her own values and preferences, and that there may be a need to involve other stakeholders.

By appropriate use of GRADE a majority of recommendations will end up being weak, reflecting either a fine balance between the advantages and disadvantages of the considered options and/or expected large variation in individual patients' values and preferences. When dealing with weak recommendations it is thus essential to make the evidence base and rationale easily available to help clinicians tailor care when interacting with individual patients.

1.3 Knowledge translation

1.3.1 Definitions and purpose
In 2000 the Canadian Institutes of Health Research (CIHR) defined knowledge translation as “the exchange, synthesis and ethically-sound application of knowledge - within a complex system of interactions among researchers and users - to accelerate the capture of the benefits of research for Canadians through improved health, more effective services and products, and a strengthened health care system”.

Since then several more organizations have made their own definitions of knowledge translation.

Evidence-based clinical practice is defined as the integration of best current research evidence with clinical, experience-based expertise and the patient’s values and preferences. While increasingly recognized as a global standard for health care its implementation is lagging behind. This could help explain the current gap between what doctors are doing in practice and what they ought to do.
Knowledge translation aims at closing this “knowledge to practice” gap, i.e. the difference between existing evidence and clinical practice. It is a nonlinear, interactive, dynamic process that includes knowledge production, synthesis, dissemination, communication and utilization at the point of care, health policy changes and continuous monitoring and evaluation to inform future studies and amendments to the process. Translation of research evidence into correct diagnosis, treatment and follow-up of patients requires systematic efforts that include key concepts of evidence-based practice.

1.3.2 Knowledge to action process framework
Several different knowledge translation models exist, one of which is the knowledge to action cycle. As shown in figure 2, the process starts with knowledge creation, moving from inquiry through synthesis and lastly the production of knowledge tools. In a dynamic fashion, knowledge creation is combined with a planned action cycle to aid knowledge application. The starting point is identifying an issue at hand, followed by customizing the knowledge to a specific context, assessing potential barriers and facilitators to adoption and designing specific implementation strategies to address these issues. Lastly, use and application is monitored to evaluate the effectiveness and impact of the strategies and adjust if necessary to sustain knowledge use. Knowledge adaptation to tailor recommendations to local needs may occur both during knowledge synthesis and as part of a planned action cycle when implementing evidence.

Figure 2: Knowledge to Action Framework. Courtesy of Harrison et al. Guideline adaptation and implementation planning: a prospective observational study. Implementation Science 2013 8:49.
1.3.3 Dissemination strategies

Dissemination of evidence to clinicians represents a key step in knowledge translation. Dissemination can be defined as an actively planned process for targeted distribution of evidence-based knowledge and represents an intermediary step in the continuum from evidence synthesis to implemented practice and evaluation.

In 2013, McCormack et al. published a report commissioned by AHRQ on strategies that focus on “making evidence interpretable, persuasive and actionable”. They found no existing overarching research framework to address this issue. When collating data for the report, they were faced with strategies of communication, diffusion (passive spread) and dissemination often being combined in the included studies. However, they conceded that it is challenging to clearly define and separate these concepts and their individual effects.

They investigated a wide range of different dissemination strategies. These included strategies to 1) improve reach of evidence by having a wide distribution, such as delivery by post, electronic, digital, social or mass media, interpersonal verbal group or individual outreach. 2) Strategies to motivate recipients to use and apply evidence, by use of champions (people who promote specific knowledge within their own organization), stakeholders and social networks. 3) Strategies to enhance ability to use and apply evidence, such as provision of support materials, skills training and capacity building.

Based on 38 studies, they found insufficient evidence to support any one strategy, but multicomponent approaches that addressed a combination of reach, ability or motivation seemed more effective than single strategies on clinician’ behavior and knowledge. The multicomponent strategies varied from study to study, making it difficult to conclude which combination is superior. They did not find that any particular single strategy directed at increasing ability or motivation was better than reach strategies. To add to this complexity, McCormack et al.’s findings are only partly supported by other reports. The impact of multifaceted approaches compared to singular strategies have shown inconsistent results, and none yield a large effect. Targeted strategies seemingly yield somewhat improved effects, the crux being that currently there is no validated way to match an implementation strategy to an identified barrier.

Meanwhile, Bernhardt et al. highlighted the conundrum of dissemination strategies not being oriented to the needs and preferences of end users. They suggested it be high time for researchers to
take advantage of technological advances to improve translation of research to practice, such as online videos, search engine optimization, electronic networks and social data mining.

1.3.4 Guideline implementability

Shiffman et al. defines implementability as “a set of characteristics that predict the relative ease of implementation of guideline recommendations.” In 2015, Kastner et al. performed a realist review to ascertain which intrinsic characteristics of guidelines improve their uptake (implementability). They identified 16 categories across six domains:

- **Stakeholder involvement:** Credibility of guideline development group, disclosure of conflicts of interest.
- **Evidence synthesis:** Sufficient reporting (scope, resource requirements, outcomes data, patient preferences), transparent and evidence-based approach to evidence synthesis (validity and reproducibility), currency.
- **Considered judgment:** Clinical relevance, values (including provider and patient freedom), net benefit, costs, organizational needs.
- **Implementation feasibility:** Local applicability, existing resource constraints, novelty (compatibility with existing knowledge).
- **Message:** Simple (information overload, complexity), clear (actionable) and persuasive message (framing).
- **Format:** Multiple versions, components (purpose, target audience, methods etc.), presentation (layout, structure, visualization).

Successful implementation requires addressing both extrinsic and intrinsic factors. However, amending intrinsic factors to increase implementability will most probably come at less of a cost, and be more feasible for most guideline developers with limited resources. It is considered prudent to plan and incorporate these strategies in conjunction with guideline development, especially since incorporating guideline implementation tools has been shown to improve provider compliance.

1.3.5 Communicating uncertainty

McCormack et al. also investigated strategies to explain uncertain evidence. They included nine studies in total, only one of which was directed at health care providers, and none addressed the concepts of strength of evidence, risk of bias, consistency, or applicability. They found insufficient evidence to support any one strategy of communicating precision or strength of the evidence.
They did find some evidence to support certain strategies for communicating directness and net benefit. Factual statements with explicit advice might influence patients to choose interventions with proven net effect on patient important outcomes (as opposed to effect on surrogate outcomes such as cholesterol, glucose levels etc.). Additionally, there seemed to be some support for the framing effect of receiving information on harms, which significantly influenced choices in the form of test refusals. Providing additional information on benefits did not seem to yield the same difference of effect.

1.3.6 Adaptation of guidelines
Guideline adaptation is the process of modifying recommendations developed for one particular setting to another context, ensuring local applicability. Prior to 2006 no validated framework for trans-contextual adaptation existed. Aiming to develop a framework to ensure a systematic and evidence-based process, the ADAPTE Collaboration was founded in 2006.\textsuperscript{55} Their effort resulted in the comprehensive ADAPTE Manual and Resource Toolkit, which consists of three phases (Set-up, Adaptation and Finalization) with 24 steps and 18 tools.\textsuperscript{56}

In 2010 the Guidelines International Network (GIN) assumed main responsibility for the toolkit. That same year GIN performed a survey among its members to explore current views, preferences, and anticipated benefits and barriers of the ADAPTE framework.\textsuperscript{57} Common feedback included the perceived benefit of an increased transparency of the development process, while several stated the complexity of the process being a possible barrier to its successful use. Several organizations worldwide have since made use of the ADAPTE Manual and Resource Toolkit. In addition several new frameworks have emerged, mirroring the growing frequency of guideline development and adaptation.

1.4 Adaptation frameworks: a systematic review

We have - in the absence of an available overview - recently performed a systematic review mapping existing adaptation frameworks, and adapted guidelines’ adherence to standards for trustworthy guidelines. Here we briefly report this systematic review, which has not yet been submitted for publication.
Systematic review of adaptation frameworks and their adherence to standards for trustworthy guidelines
Annette Kristiansen, Christopher Friis Berntsen, Per Olav Vandvik

Material and methods

Systematic literature search and study selection:
With the assistance of a medical librarian we developed a search strategy and searched MEDLINE, EMBASE, PsycINFO, AMED, Pedro, CINAHL, the Cochrane Central Register, NICE and the GIN database for relevant publications from their inception to September 2014. There were no language restrictions. We included original reports describing an adaptation framework or depicting experiences with adaptation of clinical practice guidelines. We excluded conference proceedings and abstracts, reports on de novo guideline development, editorials/commentaries, text books, dissertations, studies omitting a description of the adaptation methodology, as well as adaptation of anything other than recommendations (e.g. risk scores and multi-factorial community based interventions).

Screening and data-extraction:
Two reviewers (AK, POV, CFB) independently screened the abstracts and titles of all retrieved records, followed by full text assessment. Disagreements were resolved by discussion. One reviewer (AK) has thus far evaluated the included full text articles according to the pre-defined outcomes of interest: Applied adaptation methodology, adherence to a modified set of criteria for trustworthy guidelines (table 2) and resource use. Possible difference in level of trustworthiness over time was assessed. We used the web-based systematic review tool Covidence for screening and data extraction. Descriptive statistics were applied.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Establishing transparency by giving explicit description of the development process</td>
</tr>
<tr>
<td>II</td>
<td>Defined management of conflicts of interest</td>
</tr>
<tr>
<td>III</td>
<td>Multidisciplinary and balanced guideline development group</td>
</tr>
<tr>
<td>IV</td>
<td>Systematic search of the underlying evidence (guidelines)</td>
</tr>
<tr>
<td>V</td>
<td>Systematic quality assessment of the underlying evidence (seed guidelines)</td>
</tr>
<tr>
<td>VI</td>
<td>Systematic update of the source guidelines/recommendations</td>
</tr>
<tr>
<td>VII</td>
<td>Providing an explicit rationale and evidence summary</td>
</tr>
<tr>
<td>VIII</td>
<td>Providing information on both the benefits and harms of the considered interventions</td>
</tr>
<tr>
<td>IX</td>
<td>Rating the quality of the underlying evidence</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>X</td>
<td>State the strength of the recommendation</td>
</tr>
<tr>
<td>XI</td>
<td>Clearly articulated and actionable recommendations</td>
</tr>
<tr>
<td>XII</td>
<td>Description of how and why the recommendation was modified/adapted</td>
</tr>
<tr>
<td>XIII</td>
<td>External review</td>
</tr>
<tr>
<td>XIV</td>
<td>Future updating</td>
</tr>
</tbody>
</table>

Table 2: Criteria for trustworthy guidelines. Standards added to original IOM criteria in **bold**

**Results**

We identified 12,740 citations of which 460 were selected for full-text assessment. Of these 102 met the eligibility criteria (figure 3: flowchart). Agreement between the reviewers was moderate (kappa 0.59). We were not able to retrieve 21 guidelines from the GIN database. We further excluded 32 Ukrainian guidelines, as we were not able to obtain a person fluent in that language. Notably, 56 adapted guidelines included no description of their development process, and could not be included in the final analysis.

![Flowchart screening of adaptation frameworks](image)

**Adaptation methodology**

14 guideline developers reported using the ADAPTE framework,\(^{59-71}\) while a further 13 reported using a modified version or a new framework informed by ADAPTE.\(^{72-85}\) Accordingly the applied methodologies have more similarities than dissimilarities. More extensively described frameworks are summarized in table 1 in the appendix.
Both Chakraborty\textsuperscript{74} and Harstall\textsuperscript{79} performed case studies adapting a guideline for primary care, while simultaneously performing a formal evaluation of the ADAPTE framework. Challenges reported by their groups, commonly affecting other guideline groups, included finding the process more complex than anticipated and being in need of additional resources such as methodological expertise and other supporting staff to help with organizing and overseeing the completion of the guideline development. Chakraborty also identified the challenge of having several different source guidelines applying different rating scales and need of guidance on how to unify these.

Harrison et al.\textsuperscript{41} observed five guideline development groups as they applied the ADAPTE framework. Commonalities across groups included a sense that the ADAPTE approach lacked practical aspects. The guideline groups required extensive facilitation throughout and sought a separate implementation phase. The experiences made resulted in the re-formatted framework CAN-IMPLEMENT.

Two organizations, the National Comprehensive Cancer Network (NCCN)\textsuperscript{86–89} and the American Association of Clinical Oncology (ASCO)\textsuperscript{83,84}, have developed their own guideline adaptation frameworks applicable to in house guidelines. Representatives from the original guideline development group participate and assist in the adaptation process and the final guideline product is peer reviewed by the NCCN/ASCO.

\textit{Adherence to standards for trustworthy guidelines}

We had access to 43 records with sufficient data to assess adherence to our modified set of IOM criteria. Out of a total of 13 criteria, we found that the guidelines met a median of 8 (interquartile range 7 to 10). As GIN took on the responsibility of ADAPTE in 2010, and the IOM published their updated standards in 2011, we assessed if there was a tendency towards increased transparency and trustworthiness from 2012 and onward. 15 records predated 2012\textsuperscript{62,66,79,85,90–99}, while 28 were published after 2012.\textsuperscript{16,52,55,56,58,61,62,64,67,72,73,88–104} We found no difference in number of criteria met (p = 0.91). The guidelines most often omitted to include statements on future updates (63%), conflicts of interest (44%) and how the recommendations were modified (42%). Most guidelines had an overwhelming focus on the benefits of their recommended strategies, but were given points if they included any harm. The imbalance between the focus on benefits rather than harms is for this reason not reflected in the final tally of degree of trustworthiness.

\textit{Resource use}
We had insufficient data to make any formal assessment of resource requirements (i.e. financial, time, or personnel) in the identified studies.

**Limitations**

A significant amount of adapted guidelines have potentially been inadvertently ignored by the limited search and retrieval approach. Screening and inclusion of full text articles was performed in duplicate to minimize selection bias. A single reviewer performed data extraction, potentially skewing the results in favor of our adaptation approach. Several of the described guidelines in the identified records were not accessible in full text, limiting the final tally of guidelines assessable for adherence to the modified IOM criteria. This especially applies to older guidelines, possibly affecting the observation of no change over time.

**Conclusion**

The systematic review revealed more similarities than dissimilarities between the identified adaptation frameworks, with most of them including steps such as search and retrieval of source guidelines, overall quality assessments, modifications of individual recommendations, peer review and future updates. No guidelines met all of the standards for trustworthy guidelines, with many omitting any description of mode of adaptation.
2 AIMS AND OBJECTIVES

I. To explore strategies for the targeted communication of evidence-based recommendations to health care professionals, by developing and examining the usefulness of a new digital presentation format for clinical practice guidelines (paper 1).

II. To examine clinicians’ preferences, perceived usefulness and understanding of our new presentation format compared to an existing guideline format (paper 2).

III. To explore the process of contextualizing guidelines by developing a framework for guideline adaptation (paper 3).

IV. To examine the feasibility and usefulness of the digital presentation format and adaptation framework, by applying both strategies in a national case study in Norway (paper 4).
3 MATERIALS AND METHODS

3.1 General overview

This project combined a national guideline adaptation initiative with research on guideline adaptation methodology and presentation formats. Three separate research projects have run in parallel:

1) The development and mixed-method testing of new multilayered presentation formats for clinical practice guidelines developed with the GRADE system.
2) The development of an adaptation framework and stepwise process for GRADE guidelines.
3) The application of both strategies in a case study of the Norwegian adaptation of the 9th iteration of the American College of Chest Physicians (ACCP) Evidence-based Guidelines for Antithrombotic Therapy (AT9).

The material and methods applied are presented separately as the research projects differ in design and execution. Limitations of the studies are reserved for the general discussion.

3.2 Related research projects

This project is part of a research and innovation program called MAGIC. In addition, we have performed the studies on guideline presentation formats in collaboration with the "Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence" (DECIDE) project, funded by the European Union (www.decide-collaboration.eu).

3.2.1 Making Grade the Irresistible Choice (MAGIC)

MAGIC launched in 2010 and aims to improve the creation, dissemination, dynamic updating and monitoring/evaluation of trustworthy guidelines, evidence summaries and decision aids (www.magicproject.org and figure 4). Key methodological concepts are based on the GRADE framework and put into action by utilizing technological advantages. A key innovation in MAGIC – partly prompted by the difficulties encountered during the Norwegian guideline adaptation (detailed under Discussion) - is a web-based authoring and publication platform (www.MAGICapp.org) which allows for digitally structured guideline presentation formats to be authored, published, adapted and dynamically updated. Creating guideline content in a digitally structured format enables dissemination through a range of outputs including electronic medical record systems, web portals, and applications for handheld devices. Modifications, such as
recommendation updates, are relayed automatically to all guideline outputs, making dynamic updating of individual recommendations possible. Semi-automated creation of decision aids linked to guideline recommendations can possibly facilitate face-to-face shared decision-making in the clinical encounter.  

![Diagram of MAGIC app](image)

Figure 4: Overview of the MAGIC project

3.2.2 The DECIDE project

In 2011, the GRADE working group launched the DECIDE project, a multinational, multidisciplinary research project funded by the European Union. DECIDE aims to develop and provide empirical evidence for different ways of targeted communication of recommendations within health care. Ultimately, the goal is to optimize communication and uptake of evidence-based recommendations to healthcare professionals, policymakers, managers, patients and the general public. The DECIDE consortium did not pre-specify and explicitly place their aims, objectives or methodology within a defined knowledge translation framework. However the project did address key communication features supported by current evidence to improve implementability.

3.3 Multilayered presentation format

3.3.1 Development of a new presentation format (paper 1)
At the start of the project, we sought out and collated information on existing guideline and EBM textbook formats in popular use and previously tested communication strategies. The development process included cycles of brainstorming, sketching, usability testing and stakeholder feedback. We applied a common design approach of rapid iterations, which allowed us to eliminate design flaws as they were identified and retest components apparently working well in later iterations with multiple participants until saturation. The research group included frontline clinicians, guideline developers, clinical epidemiologists and interface designers with a broad international representation.

**Material:** We followed a “mobile first” design approach and customized the first prototype to a smartphone screen, as it is the smallest platform for guideline viewing and represents the most challenging presentation format. We used Blueprint, a tablet application, to create a semi-interactive guideline mock-up. We selected recommendations from published guidelines developed using the GRADE framework and restructured the content into a multilayered presentation. For each recommendation, we developed a clinical scenario introducing a treatment decision.

**Participants:** We performed usability testing with practicing physicians as representatives of clinicians frequently using guidelines. We recruited a convenience sample of physicians in six countries who reported to spend a minimum of 50% of their time in patient care. Members of the DECIDE advisory group, which included experts on guideline development, design, clinical epidemiology, implementation science, communication, and psychology, were selected as stakeholders.

**Data collection:** Usability testing was performed with one researcher, one participant and sometimes an observer. We devised a semi-structured interview, based on interview guides successfully used in previous studies, to explore relevant facets of a modified version of Morville’s user experience model (figure 5): findability, usefulness, usability, understandability, credibility, and desirability. The participants engaged in a think-aloud protocol in which they were encouraged to state any and all thoughts as they reviewed the guideline. The usability sessions were audiotaped.

Members of the advisory group provided their feedback on the first version of the presentation format by completing a seven-item e-mail questionnaire.
Data analysis: Reviewers categorized feedback according to five predefined criteria: showstoppers (i.e. preventing further use), major problems (hindering further use, but where the participant eventually figured the problem out), minor problems/cosmetic issues, positive feedback, and suggestions for improvement. Two reviewers independently assessed the stakeholder feedback and results from the first round of usability testing. The interviewer analyzed results from the second round of usability testing. We focused particularly on showstoppers and major problems in the development process.

3.3.2 Randomized controlled trial of new format versus current standard (paper 2)

Study design, setting and participants: We applied a mixed survey and randomized controlled trial. We included practicing physicians in internal and family medicine. In order to recruit representative samples of participants, investigators targeted compulsory educational sessions at teaching hospitals. In addition, we invited general practitioners to attend an educational session performed within the context of this study. Consent was given if the participant chose to answer at least one question.

We performed a standardized lecture with three components (see appendix for full presentation): 1) a clinical scenario, 2) presentation of a current guideline relevant to the scenario, displayed in a standard narrative format from UpToDate\textsuperscript{129} and our multilayered format, and 3) key communicative concepts of trustworthy guidelines and GRADE methodology. We randomized participants to view either format first by handing out clickers front-up, half of which had markings at the base. Participants were blindfolded during presentation of the format they were
Participants provided anonymous answers to the same multiple-choice questions by use of the clickers. We showed both formats to the whole group at the end.

**Outcomes:**

1) **Primary outcome:** Preference for either presentation format.
2) **Secondary outcomes:** Understanding and perceived usefulness of the evidence summaries and recommendations.
3) Anticipated course of action to the clinical scenario.
4) Perceived understanding of the strength of the recommendation.
5) Perception of presenting absolute effect estimates.

**Statistical analysis:** We analysed categorical outcomes by use of Pearson’s Chi-square test, while for continuous variables provided on a 6-point Likert scale we used the non-parametric Wilcoxon rank sum test. We included all randomised participants that answered more than one question in the final analysis. We used SPSS (version 23) for all analyses.

### 3.4 Development of an adaptation process (paper 3)

Using the ADAPTE framework as a template, through discussions and informal consensus, we sketched out a revised adaptation process that was largely influenced by personal experience with developing AT9\textsuperscript{130} and trustworthy guideline standards.\textsuperscript{15} Through brainstorming and an informal consensus process within our extended research group, utilizing their extensive expertise in evidence based practice and guideline methodology, we formulated the taxonomy for assessing individual recommendations devised according to the GRADE system outlined in the introduction. We revised the taxonomy based on judgments made by the Norwegian panelists during step 2 (see details under *Results*). An outline of the proposed adaptation process and accompanying taxonomy was distributed to a group of stakeholders, all guideline experts, for feedback.

We approached the chapter editors following submission for publication of the adapted guideline, and asked them to provide a formal evaluation of the process. A person not otherwise affiliated with the adaptation work performed a short semi-structured interview by teleconference, audiotaping and transcribing the interviews. The editors were asked to score their overall experience on a scale from 1 to 10, provide suggestions for improvements of the framework, and state positive and negative experiences. The interviewer and one member of the
research team reviewed the transcribed responses. We distributed an anonymized summary to all chapter editors for verification.

3.5 Applying novel dissemination and adaptation strategies through the Norwegian adaptation of AT9 (paper 4)

*Mandate:* Thrombosis-related disease is the leading cause of death in Norway, affecting patients treated across a wide range of medical specialties. In 2010 the Norwegian Society of Hematology decided to develop a guideline on antithrombotic therapy and prophylaxis, and mandated dr. Per Morten Sandset to carry out the work. Coinciding with this, the American College of Chest Physicians (ACCP) had commenced production on the 9th edition of their evidence-based guideline on Antithrombotic Therapy and Prevention of Thrombosis (AT9). Dr. Per Olav Vandvik - being an editor on several of the chapters in AT9 - contacted Dr. Sandset in 2010, proposing that AT9 be adapted and translated for Norwegian use. As appointed co-chairs of the guideline they formalized the collaboration with the establishment of an editorial committee given main responsibility for overseeing the project. Drs. Vandvik and Kristiansen, representing MAGIC, performed the duties of operative project management and co-ordination with research activities.

*Material:* The ACCP published AT9 in February 2012 in the scientific journal CHEST. AT9 represented a laboratory for research within clinical epidemiology and guideline methodology according to the GRADE framework. 128 panelists, including methodologists, leading clinical experts and researchers participated in the development of more than 600 recommendations over a 2 year period, their efforts summing up to a near complete overhaul, more than a mere update of AT8.

The guideline panels tread the following steps when making AT9:

1) **Formulating PICO questions.** One major change made in the 9th iteration included excluding non-patient important outcomes, such as asymptomatic deep vein thrombosis (DVT). This reduced the absolute incidence of DVT in most populations and settings, resulting in much stricter recommendations for thromboprophylactic therapy.

2) **A systematic literature search** was performed with the support of librarians. The chapter editors selected the most relevant articles.

3) **Summarizing and evaluating the retrieved evidence according to the GRADE methodology and producing evidence tables** with absolute estimates of effect in the target population.
4) When moving from evidence to recommendation, the panels systematically considered the key factors of certainty and magnitude of net benefit and values and preferences. With regards to antithrombotic treatment one must typically weigh a reduction of thrombotic events against the downside of an increased number of bleeds and practical disadvantages of treatment. A systematic review of studies on patient preferences was conducted. With regards to antithrombotic treatment one must typically weigh a reduction of thrombotic events against the downside of an increased number of bleeds and practical disadvantages of treatment. A systematic review of studies on patient preferences was conducted. Each chapter included a health economist with clinical expertise ensuring that resource considerations were addressed, but resource evaluations were only included in the recommendations when it was likely that they would change the direction or strength of the recommendation.

5) Each AT9 panelist reported intellectual and financial conflicts of interest. AT9 distinguishes between primary and secondary conflicts. Participants with primary conflicts of interest were supposed to be excluded from participation in the final phrasing of recommendations in AT9. A chapter editor without primary conflicts of interest and with methodological expertise led each chapter together with a deputy editor with content expertise.

6) AT9 underwent an internal peer review process with the Health and Science Policy Committee and Board of Regents at the ACCP. Following revisions, the manuscripts were externally reviewed before publication in the journal CHEST.

Norwegian guideline panelists: The co-chairs approached relevant medical specialty societies and asked them to nominate candidate panelists. We recruited 42 content experts by snowball sampling, 11 particularly interested frontline clinicians were designated as chapter editors and took primary responsibility for completing the adapted chapters. Chapter editors underwent practical training in the GRADE methodology and were asked to create evidence profiles for a relevant question in their chapter prior to initiating the guideline work. We gave the remaining panelists an introduction to GRADE at a kick-off conference in February 2012. Each panel had a designated methods expert from the MAGIC team.

Adaptation process: We set up an Editorial Committee to oversee and coordinate the adaptation work. The committee consisted of three content and three method experts.

Directly following the publication of AT9, we provided each panel with a printed copy of their chapter. Each panelist individually assessed and recorded their initial evaluation of the recommendations in separate forms (see appendix), using the taxonomy as a guide. Panelists stated any conflicts of interest on a per recommendation level. The chapter editor collated the
forms, highlighting any disagreements between panelists. The evaluation summaries formed the basis for group deliberations, and helped structure the discussions to reach informal consensus on which recommendations to exclude or modify, and to identify the need for new recommendations.

All panel members adduced new evidence identified through their clinical work. In addition, colleagues at McMaster University conducted systematic searches for new evidence using McMaster PLUS (Premium Literature Service) database entries from November 2010 to May 2013. The chapter editors screened the updated feeds, focusing on evidence that could potentially require modifications or new recommendations. During development of new recommendations, the designated methods expert performed the initial evidence assessment and rating according to GRADE, providing the panels with a tabular evidence summary (summary of findings table) as the basis for further discussions in which the methods experts actively participated.

When rewriting the guidelines in the new presentation format, the panelists had to explicitly re-evaluate the key factors of each recommendation. One methods and one content expert from the editorial committee read each chapter in duplicate. The adapted guideline underwent external review by user representatives, relevant specialty associations in Norway, the Norwegian Medicines Agency and the American College of Chest Physicians (ACCP). The chapter editors were primarily responsible for reading the feedback and assessing the need for adjustments. If major adjustments were potentially necessary, the chapter editors consulted with the panels and methods experts. All controversial recommendations were extensively discussed within the editorial committee until reaching consensus.

*Measurements and data collection:* We categorized all modifications according to the taxonomy criteria in duplicate (AK and POV), and recorded factors not included in the proposed taxonomy. We mapped the frequency of the different reasons for modification and the mode of modification.

Each panelist reported time spent on the adaptation work on a biweekly basis by responding to an e-mail prompt.

### 3.6 Ethical considerations

We have not recorded any patient data or sensitive information about the participating clinicians. We presented the project to the Regional Ethics Board and the Data Protection Official for Research who considered the studies to fall outside of their area of responsibility.
4 SUMMARY OF RESULTS

4.1 PAPER I: DEVELOPMENT OF A NOVEL, MULTILAYERED PRESENTATION FORMAT FOR CLINICAL PRACTICE GUIDELINES

Following several rounds of brainstorming and sketching, we agreed on a prototype multilayered presentation format custom-made for a digital output. Applying a multilayered approach highlights the recommendation, with more in-depth information one click away. We dubbed the new format the “top layer,” defined as the minimum set of information necessary for understanding and appropriately applying the recommendation in clinical encounters. The top layer consists of:

1. The recommendation statement including strength of the recommendation (strong vs. weak).
2. A rationale: a statement presenting the guideline panel’s reasoning regarding integration of key factors into the recommendation.
3. The key information: the balance between benefits and harms of patient important outcomes, the confidence in the estimates of the effect of the interventions under consideration, extent of variability in patient values and preferences, and resource/other considerations.

We collected feedback from 27 stakeholders and performed usability testing with 47 practicing clinicians from six countries. Based on feedback and usability results on the prototype, which uncovered showstoppers such as difficulties with conceptual understanding of guideline methodology and navigating the complexity of the layout, we made major revisions on the second version. Findings in the ensuing usability testing included an increase in positive feedback on usability, findability, usefulness and desirability. Challenges with the format’s credibility and understandability remained, with participants especially struggling with the strength of the recommendation and the concept of «preferences and values».

4.2 PAPER II: MULTILAYERED AND DIGITALLY STRUCTURED PRESENTATION FORMATS OF TRUSTWORTHY RECOMMENDATIONS: A RANDOMIZED TRIAL

We included 181 physicians from six departments of internal medicine and three primary care centers in Lebanon, Norway, Spain and The United Kingdom.

Most participants (72% vs. 16%) expressed preference for the multilayered format compared to the standard format. We detected no difference in secondary outcomes between the randomized groups,
with p-values presented in parenthesis. 89% correctly understood the evidence summaries (p-value = 0.91), while 65% correctly understood the recommendations (p-value = 0.06). A slight majority of participants stated that they found recommendations to be useful (56%, p-value = 0.33). When asked about intended clinical action in relation to the clinical scenario, 95% would treat according to the provided guidelines (p-value = 0.10).

We twice asked to what extent participants agreed to the following statement: “I fully understand the difference between strong and weak recommendations and the implications for clinical decision making.” Prior to randomization 63 of 158 participants (40%, 95% CI 32-48) stated that they agreed or strongly agreed to this statement. After randomization, we provided the participants with a one slide explanation of the strength of the recommendation, and posed the same question again. 71 of 89 participants (80%, 95% CI 70-88) agreed or strongly agreed with the statement. When specifically asked, 84 of 102 (82%) participants considered absolute effect estimates provided in the multilayered format helpful or crucial.
4.3 PAPER III: ADAPTATION OF TRUSTWORTHY GUIDELINES DEVELOPED USING THE GRADE METHODOLOGY: A NOVEL FIVE-STEP PROCESS

Figure 7 outlines the final non-linear guideline adaptation process. The process includes five distinctive steps: 1) planning, 2) initial individual assessment of recommendations, 3) final modifications, 4) publication of the guideline and 5) evaluation and planning for the future.

We have equipped steps 2 and 3 with a taxonomy based on the GRADE system. The recommendation, strength of the recommendation and the overall quality of the evidence, are the starting entities of step 2 (individual assessment). If the panelists identify a need to modify recommendations they are asked to systematically consider factors ranging from formulation of PICO questions to assessment of quality of evidence, baseline risk estimates, values and preferences and cost considerations. The final recommendations should be provided with a rationale including statements on how (e.g. from strong to weak recommendation) and why (e.g. new baseline risk estimates provided) modifications were made. These rationales - published in the guideline - provide transparency in the adaptation process and will also secure any copyright issues demanded by the source guideline.
Practical application of these five steps is exemplified by reports on the Norwegian adaptation of AT9. Final publication was delayed for nine months while the Editorial Committee deliberated with the ACCP regarding a licensing agreement for legal adaptation of AT9. We excluded 30 (9%), modified 131 (39%) of the original recommendations and developed eight new recommendations. Across chapters, there was a median of 6 major modifications or new recommendation. The most common reason for modification was feasibility issues (e.g. availability of drugs). Based on experiences during this case study, we added disagreement with feasibility and final modification: minor rephrasing to the taxonomy. We also pointed out that any guideline group wishing to adapt should investigate possible copyright issues, and added a separate tab dubbed Adaptation (for explicitly stating mode of modification, disclaimers etc.) within the MAGICapp.

We invited 28 stakeholders from Europe and North America to provide feedback on the proposed adaptation process, and received two responses, precluding any amendments to be made. Of the 11 invited chapter editors, four provided format evaluation. They gave an average score of 7 out of 10 for the overall adaptation process. Positive experiences included learning the GRADE methodology and improving and tailoring the recommendations on the basis of current best evidence and national circumstances. The chapter editors expressed frustration by the long delay in the timeline resulting from insufficient planning and recommended that this step be clarified.

The chapter editors reported a median of 116 hours spent on the adaptation work (range, 50-218 hours). We observed a non-significant trend of higher workload correlating with number of major modifications or new recommendations developed (p = 0.058).
4.4 PAPER IV: APPLYING NEW STRATEGIES FOR THE NATIONAL ADAPTATION, UPDATING, AND DISSEMINATION OF TRUSTWORTHY GUIDELINES

We assessed 333 recommendations from 11 of the 15 management chapters in AT9, condensed, restructured and rewrote them into 249 actionable recommendations in the multilayered format. We incorporated 29 best practice statements about direct oral anticoagulants developed by the Norwegian Directorate of Health, and practical information sections for 121 recommendations. Across all chapters a total of 29 new evidence profiles were developed based on new PICO's and/or new evidence.

Decisions regarding the modification of the original recommendations took into account national circumstances, including the availability of recommended interventions, treatment traditions and existing guidelines in current use in Norway. The panels spent a substantial amount of time focusing on two major issues: (1) finding credible baseline estimates, preferably based on Norwegian data and (2) reevaluating the balance between the benefits and harms of interventions in relation to assumed typical patient preferences and values.
With the help of colleagues at McMaster in Canada, we updated AT9 by performing a systematic search in the McMaster PLUS database, first in November 2012 and again in May 2013. This database contains quality-assessed studies with relevance for clinical practice from 120 leading journals. Chapter editors received the results in two stages, screened and identified relevant studies that could potentially lead to modifications or new recommendations.

We sent all relevant specialist societies and patient organizations the adapted guideline for peer review, following the rules of the Norwegian Medical Association. We published the adapted guideline online by use of the MAGICapp in November 2013. Each recommendation includes an adaptation disclaimer referencing AT9, stating how and why the recommendation has been modified.
5 DISCUSSION

The studies within this thesis have addressed certain key elements within the knowledge translation continuum, specifically exploring new strategies for improving dissemination and adaptation of trustworthy guidelines. This has resulted in a new way of structuring and presenting guidelines that during testing performed well compared to a standard guideline format with regards to preference. We demonstrated that as many, and a clear majority of clinicians could correctly interpret the recommendation and evidence summary when presented the new and standard formats. In addition, we developed a new adaptation framework and demonstrated the real-life feasibility of both the framework and the presentation format through the Norwegian adaptation of AT9. Papers I, II and IV cover the three development phases of the DECIDE project (figure 8).^120

An international and multidisciplinary representation within the research and development group, applying a mixed methods approach to explore and examine various aspects of the novel presentation format,^144 and executing tests at multiple cites, are strengths of these studies. Furthermore, thrombosis is a field fortunate enough to be informed by large randomized trials (e.g. in cardiology), but also dealing with a dearth of knowledge in themes such as safety of treatment in pregnancy. In this regard AT9 is a representative example of the widely differing available evidence base that guideline developers face and thus adds more potential for generalizability to our case study.

Despite not formally being placed within a particular knowledge transfer framework, the DECIDE research approach is supported by an existing framework for guideline implementability as described by Gagliardi et al.\textsuperscript{121} Some important methodological issues related to the studies are discussed below. These highlight challenges in designing and conducting knowledge translation
research and call for consensus methods to assess and define key concepts such as understanding of guideline content and real life use of guidelines.  

5.1 Methodological considerations

5.1.1 Qualitative usability testing
Researchers involved in the development also performed parts of the usability testing of the multilayered format, possibly causing participants to censure some of their comments. We transcribed the interviews and analyzed them in duplicate after the first round of usability testing. Due to a tight time frame, we refrained from this following the second round. These pragmatic decisions might have led to a biased interpretation of the results.

Not all facets of usability could be adequately evaluated due to the constraints of a semi-interactive mock-up prototype and a fictitious clinical scenario. We observed animosity and confusion regarding more complex methodological issues, somewhat alleviated by clearer labeling, but also by more everyday phrasing of the guideline text. The distinction between the effects of the presentation format (fixed variables including layout, structure and labeling) and guideline content (invariably changeable) was not always clearly distinguishable or discernable.

All the participants were physicians and volunteered to provide feedback. We can thus not conclude that their perceptions are externally valid, i.e. transferrable to a wider population of clinicians.

5.1.2 Randomized trial and survey of the multilayered format vs. a standard format
Researchers involved in the development of the multilayered format performed parts of the presentations, which might have biased the participants’ answers on preference and perceived usability in favor of the new format. We did not inform them beforehand which format we had developed, which can have alleviated some of this potential bias.

Due to randomization, the participants were fairly well balanced with regards to baseline characteristics between the groups. We collected answers by use of multiple-choice questions, and participants provided answers anonymously. Lastly, the participants were blinded, all of which reduces risk of systematic error. However we did not blind the data collectors or analysts, potentially introducing a biased assessment. We asked the participants both about their intended clinical action and how they interpreted the recommendations’ strength and treatment intent. Most of the participants answered the questions correctly. We cannot directly infer from this a causality
between the format and correct understanding, as we have not corrected for possible confounding from previous knowledge on the clinical topic or possible immediate recognition (as opposed to understanding) of provided labeling of the strength of the recommendations (we used the terms “strong”/”weak” in both the recommendation presentations and in the questions).

5.1.3 Five-step adaptation process

Feedback and evaluation of the adaptation framework: We received brief responses from only two stakeholders, giving insufficient data to perform alterations to the proposed framework. We aimed at interviewing the 11 chapter editors at the end of the Norwegian case study, but only four accepted the invitation within the provided timeframe. We refrained from interviewing the panel members, given their limited involvement in the adaptation process. We did distribute a summary of the provided evaluations to all the chapter editors, providing them a second chance to comment and provide suggestions for improvements. However, overall we have limited feedback and evaluation of the process to base any alterations on.

Adapting AT9 to the Norwegian setting proved the feasibility of the framework. The process is underpinned by existing systems for guideline development and adaptation, providing some support for its validity. However, further testing and evaluation by other guideline development groups in different settings and on different topics is needed before conclusions on the generalizability and validity of the framework can be made.

The panelists reported time spent on the adaptation work retrospectively and incompletely despite our attempts at rigorously evaluating these aspects. This can have introduced recall bias. The methodologists did not report time spent. We have no available data on time spent on making AT9, precluding statements on the cost and resource efficiency of the adaptation.

5.2 Discussion of remaining challenges and main results

Guidelines that meet standards for trustworthiness can be useful tools for clinicians, and ultimately their patients, having to sift through an ever increasing body of evidence to make trade-off decisions on an everyday basis.9,21,145,146 Development and appropriate updating of guidelines in the context of new evidence is highly resource demanding; requiring sufficient manpower and logistical, methodological and financial support. Local or national adaptation can possibly alleviate some of these, while in parallel facilitate dissemination by reducing language barriers and taking specific contextual issues into account.41,44,49,142 Effective dissemination can also be achieved by integrating
guidelines in a variety of information resources (e.g. web-guideline publications, smartphone applications, EBM textbooks such as UpToDate or as decision support in the electronic medical record). Such integration is facilitated by having a digitally structured guideline format with components that are interpretable both by computers and end users.

Pathman et al. cleverly phrased four stages from evidence to action: the clinician needs to be aware, then agree, then adopt, and then adhere. Below we outline some of the current challenges facing safe, efficient and effective knowledge translation, restrained within the scope of this thesis, which does not include implementation.

5.2.1 Current quality of clinical practice guidelines
In 2011, Kung et al. assessed a random sample of 130 guidelines indexed at the National Guideline Clearing House (NGC), mapping out their compliance with the 2011 IOM standards. They found that less than half the guidelines met ≥ 50% of the standards, and concluded that guidelines generally came up short in terms of an objective and transparent development process and reporting. They found no improvements over time.

Our systematic review on adaptation frameworks failed to identify any improvement in level of trustworthiness of guidelines since 2012. We identified several trends and issues giving some cause for concern. We had to exclude 56 articles as they lacked any description of their guideline development process. Almost half of the included guidelines omitted to report conflicts of interest, while a few less had no mention of future updates or how they had modified the recommendations.

Other reviews have supported our findings of suboptimal quality. Though they have found signs of improvement in quality over time, guidelines consistently failed to address applicability, defined by AGREE as pertaining to “the likely barriers and facilitators to implementation, strategies to improve uptake, and resource implications of applying the guideline”.

The Agency for Healthcare Research and Quality (AHRQ) have adopted the 2011 IOM standards, implementing a revised set of criteria for any new guideline published at the NGC website as of June 2014. Guidelines meeting these criteria are labeled to differentiate them from those accepted under the previous set of criteria. A year in, the revised criteria seem to be effectively stringent, with a drop in approved guidelines to about 20 percent of the number included under the 1997 criteria. AHRQ now provides assistance to guideline developers to help them meet the updated 2013 standards.
5.2.2 Information seeking behavior

In 2003, Dawes and Sampson\textsuperscript{155} looked at the information seeking behavior of clinicians, conducting a systematic review on the topic including 2788 clinicians across 19 studies from 1978 until 2001. They found that most clinicians consulted text sources and/or asked a colleague, a practice that might not always be prudent,\textsuperscript{156} and that only a third of clinical questions were sought out. Common barriers to information retrieval included lack of time, an overwhelming amount of material to evaluate and forgetfulness. Factors increasing the successful attainment of information included accessibility, habit, reliability, speed of use and applicability. They concluded that: «What is clear is that it [information and knowledge provision] needs to be useful, relevant and fast.»

In 2014 Del Fiol et al.\textsuperscript{1} performed a systematic review on the same topic. They identified 72 studies from 1978 to 2011. On average, clinicians had one clinical question for every second patient they encountered, half of which they pursued, finding answers in 78\% of the cases. Studies applying direct observation of clinicians found higher frequencies of questions asked combined with lower frequencies of questions pursued, indicating some level of recall bias. Clinicians stated lack of time and not believing that a useful answer could be found as common reasons for leaving queries unanswered, being consistent with findings by Dawes and Sampson. Del Fiol et al found that clinicians spent less than 3 minutes on answer retrieval, indicating that they selectively pursued queries with anticipated quick and easy answers.

5.2.3 From information seeking to clinical action

Barriers to efficient and effective dissemination and implementation have been extensively investigated.\textsuperscript{37,44,157–159} Ducharme et al.\textsuperscript{160} concluded that active dissemination with use of a simple actionable message represented a superior solution to the common ailment of clinicians’ fatigue. «…the intention-behavior gap results from two main problems that can be addressed, failing to get started and getting derailed.» As briefly outlined in the introduction, Kastner et al. found that guideline implementability in broad terms comes down to conscientious creation of guideline content (e.g. by considering feasibility) and effectively communicating that content to the end users (e.g. by use of a simple, clear, persuasive message in multiple formats).\textsuperscript{50}

Recommendations are not always easily identified, include insufficient information to be readily applied, are phrased in broad and general terms and omit to facilitate shared decision making with the patient.\textsuperscript{159,161} As stated multiple times, a clear and straightforward presentation of guideline content is key to helping clinicians act correctly on the recommendations.\textsuperscript{147,162,163,164} The 2011
IOM report even includes it as a quality mark of guidelines and states that they should “consider important patient subgroups and patient preferences….provide a clear explanation of the logical relationships between alternative care options and health outcomes, and provide ratings of both the quality of evidence and the strength of the recommendations.“ The informational needs of clinicians are inherently complex and the usefulness of recommendations results from their relevance and validity, in addition to the amount of work needed to access the recommendation.\(^{147}\) They need to provide flexibility,\(^{10}\) clarity and facilitate tailored application at the point of care.\(^{165}\)

Several attempts at finding effective ways of presenting summaries, recommendations and uncertainty have been performed,\(^{6,163,166–170}\) but to the best of our knowledge there is still scarce evidence on satisfactory ways of presenting not just singular elements of a guideline, but guidelines presented as complete packages. We therefore aimed to specifically explore ways of communicating graded evidence and recommendations to clinicians.

5.2.4 Dissemination through multilayered guideline presentation formats

A guideline format needs to be feasible to produce (for guideline developers), acceptable to stakeholders (e.g. publishers and clinicians) and publishable across different media outlets.

Given the present vast technological opportunities it seems unreasonable to keep producing lengthy textbook-like guidelines aimed at isolated PDF publications with little leeway for adapting its use to different needs. Our multilayered presentation format is translatable to any electronic platform (the web, tablets, smartphones, directly linked in the electronic medical record), providing clinicians with actionable recommendations up front with underlying information buried in deeper layers. Quick and easy access (“EBM at 2 a.m.”), with a format that can to some degree be tailored to individual needs, can potentially facilitate dissemination and uptake of guidelines.\(^{44,171}\) Even though there is thus fare insufficient evidence to support the notion that an electronic guideline format increases uptake,\(^{172}\) it seems unwise to keep writing guidelines in a stationary text-processing based format given the vast opportunities for multi-use of one-off input of data.

During the developmental phase of the multilayered presentation format, we realized that there is no one solution that fits all end users. Different clinical situations will trigger varying informational requirements, individual clinicians will have differing levels of numeracy etc. By providing end users with layering of information, provided both as text summaries, tables and other graphical representations we aim at meeting most health care professionals’ baseline levels of knowledge and informational needs.
We started out mainly concerned with communicating the recommendation itself clearly and the rationale behind it, i.e. looking back in time to what the evidence said and what the guideline makers discussed. However, during usability testing the participants expressed nearly equal concern about how this information could be used in a patient discussion, i.e. looking forward in time. One could thus conceptualize the multilayered format as being a communication tool that rests in between these two spaces and caters to both of them, rather than just being a clear communication from the guideline makers. Within the broader MAGIC program we are developing semi-automatically generated decision aids, directly linked to individual recommendations, to be used during a clinical encounter to facilitate a conversation between the clinician and the patient.119

5.2.5 Communicating concepts of uncertainty and trustworthy guidelines

It is recognized that clinicians are little better equipped than the general public at understanding risks and probabilities.166,168,170,173–175 Reyna et al.168 theorize that most people’s immediate comprehension and judgments are based on the “fuzzy gist of information” as opposed to focusing on the precise numbers. Acknowledging this preferred mode of memory processing Reyna et al. suggest trying to explain numbers verbally by providing a context and stressing understanding of the bottom-line (e.g. by stating if 20/1000 is a high baseline risk or not and if a reduction by treatment is large or not). They suggest that information be displayed visually.

Reyna et al. state that providing exact numbers on risk might however be counter productive, especially in cases of primary prevention, as the relative small nominators defer patients from being compliant with recommended treatment. However, McGettigan et al.176 discovered that «interpretation of clinical information and subsequent treatment decisions were more likely to concur across specialties when data were presented in numeric, as opposed to verbal, terms.» They furthermore found that people are subconsciously influenced by the framing of the probabilities; observing that people are risk averse when considering gains, but risk seeking when considering losses. They conclude that «framing effects could be reduced or eliminated by providing information on the reciprocal relation between the gains and losses». The best way of communicating risk is even more important as framing bias affect not only clinicians, but also patients.173

We judge that clinicians (and patients) should be well informed on baseline risks to be able to make appropriate priorities when faced with multiple choices and the highly probable risk of polypharmacy in multimorbid patients.177 Our survey and randomized trial elicited positive
feedback on the perceived usability of being provided absolute risk estimates. For now, given the paucity of conclusive evidence, we have opted to stick with absolute frequencies, keeping both the denominator and the class reference constant, and omit confidence intervals, to minimize the risk of confusion.\textsuperscript{173,126,178–181} Due to the wide variability in end users’ numeracy, we advice guideline authors to also add a verbal gist when phrasing the key information. Such gists are combined with absolute estimates of effects in the MAGICapp summary of findings (SoF) tables.\textsuperscript{182} They will be included in the interactive SoF tables,\textsuperscript{183} recently adopted as preferred presentation formats in systematic reviews developed by the Cochrane collaboration.

One of the trustworthy standards is to communicate the confidence in effect estimates and the strength of the recommendation.\textsuperscript{15} No conclusive evidence exists on how to best convey this vital information.\textsuperscript{42,167,184,185} During usability testing we discovered that these complex methodological issues elicited confusion, uncertainty and a certain degree of animosity in the participants. We found that participants’ perceived understanding improved when we tidied up the phrasing by using precise and clear plain language to define methodological factors (e.g. “quality of the evidence” rather than “confidence in the magnitude of the effects”). After numerous rounds of brainstorming, sketching and lastly testing of different symbols, we found no solution that surpassed the commonly used labels of strong and weak recommendations. Thus far, we can only conclude that conceptual understanding of the recommendation strength and quality of evidence remains suboptimal, even in guideline developers.\textsuperscript{186} Keeping the strong/weak labeling necessitates educating guideline developers and clinicians in the practical implications and differences between strong and weak recommendations, aiming at minimizing animosity and incorrect application of the recommendations.\textsuperscript{187,188} Within the MAGICapp user interface, we have added a short explanatory legend directly underneath the strength of the recommendation with a link to a comprehensive explanation (www.magicapp.org). The effect on clinicians’ understanding by providing this additional explanation has yet to be fully evaluated.

5.2.6 Adaptation of trustworthy guidelines

ADAPTE is the leading adaptation framework, and has been adopted by several guideline organizations. It facilitates adherence to a stringent development methodology and provides transparency to the process. A survey performed by GIN in 2010 found overall positive feedback on the ADAPTE process. Reported anticipated barriers included complexity of the process, lack of source guidelines of sufficient high quality and difficulties in adapting to context of use.\textsuperscript{57} Since then, several developers have applied their own modified version of ADAPTE, indicating that the framework might be perceived as somewhat cumbersome and/or lacking in certain areas.
Our five-step approach for adaptation is constructed to increase efficiency by reducing complexity in the setting of an already fully developed guideline, preferably using the GRADE system. As the taxonomy for modifications is devised around GRADE, a pre-requisite is available expertise in the application of GRADE and capacity within the responsible organization to provide basic training in guideline development. Guideline organizations that hold these skills may benefit from a greater emphasis on a structured and transparent adaptation of individual recommendations through use of the taxonomy.

5.2.7 The rationale of adaptation - an empty promise?
Adaptation holds several potential benefits, major ones being resource-saving and enhanced uptake by making recommendations more acceptable and applicable.

Several barriers can potentially hinder an efficient adaptation, making organizations likely to opt for de novo development. As previously mentioned; one major issue is the lack of source guidelines of sufficient quality and currency. One can’t make a silk purse of a sow’s ear, i.e. that it is fruitless adapting a low quality guideline and hope to end up with a superior end result. Proprietary guidelines, lacking expertise in guideline methodology, need for training, adaptation being no less complex than de novo development and insufficient a priori planning and follow through can curb an efficient process. Lastly, application of differing grading systems across source guidelines makes it cumbersome to align the classification in the adapted guideline product.

De novo guideline development is, when performed with stringent adherence to trustworthy standards, a resource demanding undertaking. The notion of not duplicating someone else’s hard work has an alluring appeal. It seems logical that basing a guideline on one or a few source guidelines would restrict resource consumption (time, personnel, money). Real-life experiences thus far have not been able to neither confirm nor disprove this potential. Directly comparing resource use is problematic due to insufficient reporting and variables such as magnitude and scope of the adapted guideline. That being said, with sufficient planning we would have been able to complete the Norwegian adaptation within one year’s time - half the time spent on AT9 – with the help of less than half the personnel. We achieved efficiency mostly by not having to retrieve, screen and rate the available evidence, confirming that access to a trustworthy source guideline with SoF tables simplified adaptation.
ADAPTE defines acceptability as “the extent to which the users are likely to adopt a recommendation, based on internal qualities such as clarity, comprehensiveness, and logical reasoning and on external factors such as the burden imposed on the process and system of care, patient and providers attitudes and beliefs, and patients needs, expectations, and preferences.” Applicability meanwhile is defined “as the extent to which the users are able to put a recommendation into practice, based on internal qualities such as a clearly defined eligible patient population that matches the population to which the intervention is targeted in the local setting and external factors such as the availability of the necessary knowledge, skills, provider time, staff, equipment, and other resources.” Even though both ADAPTE and our five-step adaptation process mainly focus on the development of a guideline, and far less on implementation, contextual adaptation is considered a necessary step in knowledge translation, and there are indications that tailoring impacts on guideline uptake.

Recent major guideline efforts in Estonia, Belgium, Egypt and Kazakhstan demonstrate the feasibility of adaptation. The more resource intensive efforts made by the Estonians had the added benefit of building national expertise in guideline development – an asset they can utilize in future endeavors. In Egypt and Belgium, they did demonstrate that time spent on each adaptation was inversely correlated to capacity building over time. The Belgians on their side showed that having high quality source guidelines lightened the workload, and ended up doing major adaptations to only 10% of around 750 adapted guidelines to comply with local resource availability. Furthermore they all demonstrated that having a clearly defined protocol and available tools facilitated the process.

As we have discussed above, the problem of low quality source guidelines might be improving. The increasing use of GRADE by guideline organizations worldwide lends promise to the problem of differing grading systems across guidelines. Even if the different source guidelines are applying different or no grading systems, one can still make use of their reference lists and save time on search and screening.

5.2.8 Lessons learned from making the Norwegian antithrombotic guideline
The case study we performed in Norway was in many ways a cumbersome endeavor and if nothing else we have demonstrated that it is possible to successfully adapt a large guideline despite limited resources and several unforeseen barriers. Furthermore, we demonstrated that an extensive guideline, traversing several different specialties and involving 11 different main authors, could be written entirely in the multilayered format.
During the project we have experienced the advantages, limitations and challenges of guideline adaptation and dissemination. The guideline group had to change publisher three times. As the preliminary deliberations and discussions with the journal of the Norwegian Medical Association (Tidsskriftet) fell through, the original plan of having them procure an electronic publication platform had to be revised. The guideline group joined efforts with the Directorate of Health to communicate updated evidence on antithrombotic therapy to health care professionals. In the fall of 2012, the ACCP contacted the guideline group, requiring a licensing agreement to use and adapt AT9 before work could continue. The Directorate was unable to enter into such an agreement with a non-governmental agency, and the guideline group and the Directorate parted ways. The predicament of lacking a formal publisher pushed forward the long awaited establishment of the Norwegian Society of Thrombosis and Hemostasis (NSTH). Furthermore, the research group found no better solution than to design and program our own authoring and publication platform, i.e. MAGICapp, to be able to publish the guideline as intended. We spent the delay caused by mandatory deliberations with ACCP concerning the licensing agreement making crucial improvements on the usability of the publication platform. In the end we published the adapted guideline online as initially planned, having NSTH as official publisher. We advertised the publication in Tidsskriftet202–206 and other major national newspapers and medical journals and involved stakeholders and opinion leaders both during the development (as panelists) and during peer review. NSTH has since 2013 been formally included as an interdisciplinary specialty association in the Norwegian Medical Association and held its first yearly conference in Oslo in 2015. Among featured topics were plenaries on the antithrombotic guideline and updated recommendations.

The experience with the Norwegian adaptation demonstrated the value of a solid leadership steering the guideline development and providing guidance until consensus. A close collaboration between content and methods experts allowed rapid resolution of complex problems and, therefore, limited time loss. We experienced first hand the detrimental effects of insufficient a priori planning and mapping of copyright issues.

We have experienced a limited uptake and use of the Norwegian antithrombotic guideline following publication in November 2013. Data so far (as of October 2015) show that the guideline has been accessed on average 800 times per month, with 84% of viewings done by computer. We briefly discuss reasons for a limited uptake.
Up until the current date we have made systematic and ad hoc investigations into reasons for the limited use of the adapted antithrombotic guideline, and mark them according to Morville’s user experience facets (in parenthesis). Firstly, most physicians in Norway seem unaware of and/or struggle to find the guideline (findability). The new publication platform MAGICapp, programmed with use of state of the art software, is nearly impossible to access on older web browsers, such as Internet Explorer ≤ 8 (accessibility). Unfortunately, at the time of publication most hospitals in Norway utilized older versions of Internet Explorer, while the user interface on mobile devices was hampered by generous spacing between data points, basically limiting screen view to single recommendations (usability).

As part of a research and innovation collaboration between the MAGIC organization and Hospital Innlandet Trust, we performed usability testing with a handful of hospital physicians during their everyday clinical work. As opposed to the initial usability testing performed as part of the DECIDE project, this approach was superior as it better mimicked real life situations and was not limited to testing mock-ups. We asked the physician to access the Norwegian antithrombotic guideline to find answers to a clinical question of their choice. Feedback was collected using semi-structured interviews similar to those applied during the development of the multilayered format.

End users found the presentation format and the user interface of the website unfamiliar and unintuitive. They were hampered by navigational difficulties, as they either had to scroll through all 249 recommendations or navigate by use of a long table of contents. The software would sometimes lag in loading new data, having the detrimental effect of having the physician reading a recommendation NOT applicable to their clinical query (findability, usability). Some weren’t able to access the underlying layers of the recommendation, not seeing that they could click on the recommendation box (findability, usability). We observed language barriers in the form of an overly academic and cumbersome phrasing of the headings and recommendations (understandability). They struggled to navigate to the practical information and could not apply the recommendations without crucial information on risk scores, dosages etc. (findability, usability, understandability, usefulness).

To address these issues, we have made several improvements. Performance speed has been increased, and with the help of an interaction designer the user interface on both mobile and web viewing is more intuitive. Previously, users wishing to access the guideline through the announced web link (www.nsth.no) had to navigate via MAGICapp’s main page listing all available guidelines. Now they are routed directly to the guideline, where we have also added a search engine, rephrased
several of the headings to a more everyday clinical language (e.g. «venous thromboembolic disease» was changed to «deep vein thrombosis and pulmonary embolism») and added either hyperlinks to risk scores or guiding text in the headings and recommendations. The guideline can be downloaded and stored on smartphones and tablets, accessed from the home screen. The effect of these measures has yet to be assessed.

In addition to this, we do recognize the probable need for applying specific implementation strategies to address individual recommendations introducing treatment strategies that collide with current treatment standards, such as is recommended by both Flottorp and Kastner.\textsuperscript{50,51}

We have formally evaluated the adaptation framework and presentation format. We have not, however, systematically monitored knowledge use or evaluated their outcomes.\textsuperscript{40} Ducharme et al.\textsuperscript{160} found that «The omission of the assessment step is believed to explain the low success rate of a variety of KT interventions, which hovers around 10%.” We concede that there is compelling evidence that dissemination alone is unlikely to succeed at changing clinical practice.\textsuperscript{207}

5.2.9 In summary
In a time where the internet has revolutionized access to digital information anywhere, anytime, on every device,\textsuperscript{208} and “users can create and interact with information in ways that give classic theories of dissemination a new twist”,\textsuperscript{42} guidelines are lagging behind.\textsuperscript{48} This thesis aimed to address challenges that span scientific and practical domains, related to dissemination and adaptation of guidelines developed according to recent standards and systems for trustworthy guidelines while making use of technological advances with digitally structured data in new tools for authoring and publication. National or local contextualization and accessible user interfaces of guidelines could yield reliable, relevant and readable recommendations that might promote dissemination of guidelines.\textsuperscript{51,156,209}

Our studies on new strategies for improved dissemination and adaptation of guidelines represent emerging solutions used by an increasing number of guidelines at the international level in a time of advances in definitions, standards, systems and tools for trustworthy guidelines. Our combined research and innovation activities - performed in parallel through MAGIC - also highlight remaining challenges to address with the ultimate goal of 1) allowing clinicians to find, understand and appropriately use recommendations tailored to their information needs in busy clinical practice and 2) allowing guideline authors to maximize efficiency in the creation and adaptation of trustworthy guidelines.
6 CONCLUSION

I. We developed and, through usability testing, iteratively improved a digital, multilayered guideline presentation format, structured according to standards for trustworthy guidelines. We found that most participants thought the format to be useful, but struggled with conceptual understanding of the underlying guideline methodology (paper 1).

II. We demonstrated that physicians clearly preferred a novel multilayered presentation format to the standard format. The multilayered format was perceived as understandable and useful as the standard format (paper 2).

III. We developed a five-step adaptation process customized to trustworthy guideline standards, and added a taxonomy tool devised according to the GRADE system to facilitate a systematic and transparent assessment and communication of individually modified recommendations (paper 3).

IV. We demonstrated the feasibility of the digitally structured, multilayered format and adaptation process by applying both strategies when adapting a large guideline for antithrombotic prevention and therapy to the Norwegian setting (paper 4).
7 WHAT WE HAVE LEARNED AND WHERE DO WE GO

We have designed, usability tested and developed a functioning publication solution for trustworthy recommendations. The authoring and publication platform serves the dual end purpose of being a testing ground for all MAGIC related research endeavors and supplying clinicians with transparent recommendations directly linked to the underlying evidence assessments, risk estimates and individual references.

We have demonstrated the feasibility of a pragmatic and compressed adaptation framework. We have furthermore highlighted the crucial importance of maintaining transparency, also during adaptation, and demonstrated that this has to a large extent been lacking thus far. We have yet to demonstrate that national tailoring or new presentations improves adoption of recommendations.

We have outlined methodological limitations with our studies that reflect challenges in designing and conducting knowledge translation research and call for improved methods to assess key concepts such as understanding of guideline content and real life use of guidelines in practice.

Suggestions for future research

During the past years of being a part of the EBM world and working not only on the Norwegian Antithrombotic Guideline, but also consulting other guideline groups on GRADE and guideline development, we have experienced first hand the challenges facing guideline developers attempting at producing trustworthy guidelines. There seems to be some general frustration at the amount of work needed to develop trustworthy, evidence-based guidelines. Mapping and validating the minimum set of resources, education and tools necessary in a digitized setting, across different guideline groups, can facilitate future streamlined development processes for guideline organizations.

De novo development and adaptation should share the same backbone, effectively affecting future research endeavors to improve and crystallize both methodologies. Future research should continue to address issues of how to best solve the vexing problem of opaque and antiquated guidelines.

Further research is needed on how to best convey uncertainty. We will most likely find that clinicians will have differing preferences, but formal evaluations of correct understanding is
necessary. The SHARE-IT project (part of the MAGIC program) is qualitatively testing different ways of displaying absolute risk estimates in a patient-clinician encounter.\textsuperscript{119,210} How to develop and communicate recommendations on multimorbidity\textsuperscript{211} and multiple treatment comparisons remain highly current challenges.

On a much broader note - knowledge translation is a massive field, and has until recently been only partly addressed by the MAGIC program. A new project going by the working title of the \textit{Evidence Ecosystem} is underway. The ecosystem pertains to the lifecycle of evidence, from its production (primary research), synthesis (systematic reviews), dissemination (guidelines), implementation and ultimately impact on health-related behavior and outcomes, further informing future research. MAGIC’s evidence ecosystem aims to develop, explore and validate efficient methods and effective communicative tools for each of these life stages, utilizing the inherent assets of digitally structured and linked data. Addressing the knowledge to practice gap in such a concrete and holistic way seems a valid way to go.
REFERENCES


77. Espirito Santo A, Choquette A. Experience of adapting and implementing an evidence-based


144.  Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful


158. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs’


194. Bygrave H, Saranchuk P, Makakole L, Ford N. Feasibility and benefits of scaling up...


Appendices
## Adaptation frameworks

<table>
<thead>
<tr>
<th>Reference</th>
<th>Country</th>
<th>Study details</th>
<th>Description of adaptation framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ADAPTE Collaboration 2009</td>
<td>International</td>
<td>Resource Toolkit</td>
<td>3 phases with 24 steps.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• SET UP: Determine if adaptation is feasible, assemble organizing committee, select topic, identify necessary skills and resources, complete set-up tasks, write protocol.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• ADAPTATION: Determine PIPOH, search for CPGs, screen CPGs, reduce number of CPGs, assess quality, currency, content, consistency, acceptability/applicability of recommendations, review assessments, select between CPGs/recommendations to adapt, prepare adapted document (transparent and user friendly).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• FINALIZATION: External review, consult endorsement bodies, consult developers of source CPGs, acknowledge source documents.</td>
</tr>
<tr>
<td>Other adaptation frameworks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abulkahir 2010</td>
<td>International</td>
<td>NCCN (National Comprehensive Cancer Network) adaptation methodology applied in the Middle East and North Africa (MENA region)</td>
<td>Regional experts develop and update local editions of the CPGs in collaboration with NCCN Guideline Panel representatives. A regional chair coordinates the development with support and advice from NCCN staff and leadership. Main principles include:</td>
</tr>
<tr>
<td>Carlson 2014</td>
<td></td>
<td></td>
<td>• <em>Core work performed at an in-region scientific symposium attended by NCCN representatives,</em> also reviewing current status in the region and providing education for regional oncologists.</td>
</tr>
<tr>
<td>Icli 2010</td>
<td></td>
<td></td>
<td>• Multidisciplinary and multinational CPG groups are recruited to include important stakeholders.</td>
</tr>
<tr>
<td>Jazieh 2010</td>
<td></td>
<td></td>
<td>• <em>The CPG group reviews and drafts adapted recommendations prior to the symposium.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• <em>A set format is utilized to communicate rationale for modifications and updated references.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• NCCN panel representatives participate in the CPG groups, the review and determination of whether the proposed adaptations are appropriate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• <em>Final adapted and translated CPG is verified by NCCN through scientific translators and then published online exclusively at NCCN.org.</em></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Location</td>
<td>CPG Adapted</td>
<td>Methodology</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| Andersen 2014 | US | ASCO (American Society of Clinical Oncology) | - The methodological review is performed by methods experts using the AGREE tool.  
- *ASCO staff updates the literature search*  
- Content review is performed by a multidisciplinary ad hoc panel led by two co-chairs  
- A recommendation matrix is made if there is more than one source CPG.  
- The ASCO Clinical Practice Guidelines Committee (CPGC) reviews and approves the manuscript |
| Bower 2014 | US | adapted CPGs from Canada and the US on depression and fatigue in patients with cancer |
| Bayley 2014 | International | Methods devised by an international panel of experts on cognitive rehabilitation after traumatic brain injury (INCOG). Multiple CPGs adapted. | - Prioritize topic  
- Multidisciplinary group of experts  
- Search of CPGs  
- AGREE appraisal  
- Recommendation matrix  
- Selection & development of recommendations. *Discussed at day long meeting using the matrix*  
- *Decision algorithm*  
- Updated literature search added to extensive reference library  
- Finalize recommendation grading & algorithm. All team members had to agree.  
- *Prioritization of recommendations for implementation.* |
| Chadban 2012 Langham 2014 | Australia and New Zealand | KHA-CARI (Kidney Health Australia Caring for Australians with Renal Impairment) adapting CPGs from KDIGO (Kidney Disease, Improving Global Outcomes) | A simplified 4-step process with focus on assessing currency, consistency, applicability as preparation for drafting adapted CPG.  

**Step 1: Assess guideline currency.**  
- Review search strategy. Update or develop new recommendations as needed based on new evidence.  

**Step 2: Assess guideline consistency.**  
- *Rate evidence quality according to GRADE.*  
- Evaluate consistency between the selected evidence, its summary, interpretation and the recommendations.  
- Assess coherence between the recommendations.  

**Step 3: Assess applicability.**  
- Does the population match?  
- Is the intervention and/or equipment available?  
- Are there any national legal, organizational or resource based barriers to implementation?  
- Is the recommendation compatible with individual... |
Step 4: Prepare an adapted guideline document with recommendations and suggestions reflecting assessments made in Steps 1 to 3.

<table>
<thead>
<tr>
<th>Chakraborty 2014</th>
<th>Australia</th>
<th>Cancer Council of Australia and the Aboriginal and Torres Strait Islander community performed formal evaluation of ADAPTE as part of a CPG development, adapting several source guidelines on lung cancer</th>
<th>General comments / suggestions:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>The ADAPTE offers minimal guidance about the costs or time involved in guideline adaptation versus de novo development.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Although the ADAPTE does specify skills and organizational requirements, these descriptions do not offer sufficient detail; the complexity of the process was a challenge and required additional methodological expertise.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The expertise and time required to undertake each step are not within the capacity of guideline development group members. These tasks are better undertaken by a project team who presents the guideline development group with recommendations for ratification at each meeting.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aspects of the ADAPTE assessment module can be included within the search and screen module to enable earlier assessment of guideline quality and thereby reduce the number of guidelines that require detailed assessment.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The recommendation matrix can be combined with an evaluation of applicability.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In order to assess the level of evidence supporting each recommendation, the rating scales used by each of the shortlisted guidelines were translated to the National Health and Medical Research Council definitions of evidence.</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Country</td>
<td>Institution/Program</td>
<td>Methodology</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Cooley 2013</td>
<td>US</td>
<td>Dana-Farber Cancer Institute and Boston Medical Center</td>
<td>7-step process:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>developed a CPG on symptom management in an outpatient thoracic oncology setting</td>
<td>Step 1: Identify expert panel members for each target symptom group.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Step 2: Formulate research questions. Review and synthesize the literature <em>(performed by core research team)</em>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Step 3: Convene the panels to translate CPGs into computable algorithms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Step 4: Internal review of drafts by another panel.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Step 5: Organize a multidisciplinary consensus panel meeting to discuss the evidence-based algorithms.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Step 6: Revise algorithms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Step 7: Review and approve revised computable algorithms.</td>
</tr>
<tr>
<td>Davino-Ramaya 2012</td>
<td>US</td>
<td>Kaiser Permanente’s (KP) national guideline program</td>
<td>CPG development is overseen by a Quality Committee.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Identify Guideline Development Team (GDT) and staff. <em>Each group includes stakeholders from all KP regions, clinical experts, evidence analysts and methodologists.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• <em>Solicit suggestions for external CPGs from the GDT or other subject matter experts.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Define health questions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Search and screen CPGs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Assess the identified CPGs using the AGREE II tool, with an emphasis on «Rigor of Development».</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Decide whether to adopt certain recommendations or the entire CPG.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Draft modifications tailored to KP.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• <em>The Lead Team drafts a guideline report for GDT consideration, consisting of the CPG assessment and recommendations.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• External review.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Plan for updates and implementation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• CPG production and dissemination.</td>
</tr>
</tbody>
</table>
| Harstall 2011 | Canada | The CPG developer is somewhat unclear, but is probably the Alberta Ambassador Program. A CPG on low back pain was adapted from multiple source CPGs | Stage 1 Set Up:  
- Form an Advisory Committee for project oversight.  
- *Assess current knowledge by performing a survey and identify knowledge gap.*  
- Formulate research questions and delineate target audience.  
- *Collaborate with Health Technology Assessment Group/organization.*  

Stage 2 Adaptation:  
- Perform a systematic literature search.  
- Assess the CPG using the AGREE tool with help from HTA researchers & methodologists. Develop evidence inventory tables with source recommendations.  
- Arrange monthly meetings for deliberations.  

Stage 3 Finalization:  
- Review of CPG by provincial CPG programs.  
- Seek endorsement by patient advocacy groups, professional colleges and associations. Establish plan for dissemination and updating. |
| Kristiansen 2014 | Norway | The Norwegian Society for Thrombosis and Hemostasis adapted an international CPG on antithrombotic prevention and therapy | A 5-step adaptation process based on the GRADE methodology.  

Step 1 Planning:  
- Establish editorial committee to oversee CPG development.  
- Recruit CPG panel members.  
- Choose topics to adapt.  
- *Arrange start up conference.*  
- Plan implementation strategies.  

Step 2: Assess source guideline.  
- Initial individual assessment of recommendations by help of partly prefilled recommendation matrix including *adaptation taxonomy.*  
- Report conflicts of interest.  

Step 3: Adaptation.  
- *Draft CPG reviewed by the editorial committee in duplicate by one content and one methods expert.*  
- External review.  

Step 4: Publication.  

Step 5: Evaluation of process and plan for future updates. |
| Poltawski 2013 | UK | Unclear CPG developer of a CPG on community-based exercise programs after stroke, based on multiple source CPGs | STRIDES (STroke Rehabilitation Intervention-Development Evidence Synthesis). A framework that used Intervention Mapping to develop a multifaceted exercise program. Interventions are characterized in terms of target outcomes, proximal objectives, methods and practical strategies.  
- Define health question.  
- Search and screen CPGs. *Quality assess individual recommendations.*  
- Select recommendations (from subthemes grouped according to Mapping categories).  
- Create a customized guideline with a synthesis of recommendations by program aims, proximal objectives and strategies. |
|---|---|---|---|
| Rimmer 2014 | US | Unclear CPG developer. A CPG on obesity from the CDC (Center for Disease Control and Prevention) was adapted to people with disabilities | GRAID (Guidelines, Recommendations, Adaptations Including Disability).  
Phase 1 Set-Up:  
- Identify CPGs.  
- Convene expert panels.  
- Assess the quality of recommendations.  
Phase 2 Development of GRAIDs:  
- Perform a scoping review of resources/research that support CPG applicability.  
- *Assemble focus groups with stakeholders.*  
- Convene expert working group online meetings. Each group reviews the compiled literature for quality, clarity and appropriateness.  
- Develop draft GRAIDs to be reviewed by several working groups.  
Phase 3 Finalization:  
- *Hold annual expert panel consensus meeting to assess GRAIDs using the AGREE & GRAID criteria. (appropriateness, usability, acceptability).*  
- Stakeholder review.  
- External peer review.  
- Feedback from developers of source CPGs.  
- Update. |

CPG = Clinical practice guideline. PIPOH = Population, Intervention, Professionals, Outcomes, Health care setting. HTA = Health Technology Assessment.
Interview guide paper I

DECIDE WP1 User testing: Health care recommendations

Test person no.:  
Place:  
Date:  
Interviewer/notetaker:  

1. Checklist

For facilitator, bring:
- Printed copy of clinical scenarios
- Different formats of health care recommendations, both digital version on iPad and a paper back-up in case the test subject is uncomfortable using the digital demo.
- Blank iPad screen on paper for the participant to sketch on.

For observer/note taker, bring:
- Paper and pen to take notes
- Tape recorder

2. Introduction

> Go through the written information they have already received
- What we are doing
- Who is participating, why we invited you
- How the test is conducted
- What happens to the data/recording
- Rights to quit or retract recording
- Questions?

> Turn on audiorecorder.
Background questions – 5 minutes

<table>
<thead>
<tr>
<th>A</th>
<th><strong>Ask:</strong> How many years of clinical experience do you have?</th>
<th>... ... Years of clinical experience <em>(including internship)</em></th>
</tr>
</thead>
</table>

**Ask:** What is your clinical discipline and current position?  
How much of the time do you spend on clinical duty?

**Clinical discipline:**
- □ Primary care
- □ Internal medicine, please specify is subspecialty:
- □ Surgery, please specify if subspecialty:
- □ Other, please specify:

**Current clinical position**
- □ Primary care physician, under training  □ Primary care physician, specialist
- □ Intern
- □ Resident
- □ Consultant

**Ask:** What is your training in health research methodology *(academic background)*?
- □ Never done a formal course in HRM
- □ Done 1 or more formal courses but no masters/ Ph.D degree
- □ I have a masters/ Ph.D degree in HRM

<table>
<thead>
<tr>
<th>B</th>
<th><strong>Ask:</strong> When faced with a clinical problem you don’t know the answer to, what do you most often do? (Check all that apply if more than one action)</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>Consult a senior colleague or specialist</td>
</tr>
<tr>
<td>□</td>
<td>Look for the answer in a traditional medical text book</td>
</tr>
<tr>
<td>□</td>
<td>Look for answer in local protocol, pharmacopeia</td>
</tr>
<tr>
<td>□</td>
<td>Consult a clinical guideline (including UpToDate and other EBM textbooks)</td>
</tr>
<tr>
<td>□</td>
<td>Other, please specify:</td>
</tr>
</tbody>
</table>

| C | **Ask:** How often do you on average consult clinical practice guidelines when encountering a problem? |
Seldom or never □
Monthly □
Weekly □
Daily □

D  **Ask:** Which guideline resources do you most often consult?

E  **Say:** Why do you prefer using these resources?

---

**Repeat instructions**

A short bit of repetition before we begin.

**No right or wrong answer**
You are not being tested, it is our material we are testing. There are no right or wrong answers to our questions. If you think something is easy or difficult, clear or confusing, if you understand or don’t understand, we just want to know about it.

**Think out loud**
Think out loud. For instance:

- What you are looking at, describe your experience of it.
- If you are unsure about anything
- If you are surprised by anything
- If there are things you don’t understand, just say “I don’t know what this means...”

**My role**
My role is to ask questions. But, since it is your opinion we are interested in, I will be otherwise saying as little as possible. If you have any questions not regarding navigational issues, I will try to answer them after the test.

---

**The recommendations with underlying information**

1  Let the participant select an appropriate clinical scenario with a question about therapy at the end.

  **Ask:** "Which of the following scenarios do you wish to look at?" (tick off for selected scenario)

  *Scenario 1 is about antidiabetic treatment of diabetes type II* □
  *Scenario 2 is about use of anticoagulation in atrial fibrillation* □
1a When the participant has read the selected scenario

**Ask:** What treatment would you normally give/not give if you had to make a decision now without consulting additional sources of information or the patient about his/her preferences?

1b **Say:** Please suggest information resources where you might prefer to search for answers to this clinical question.

---

**> Wait before showing the guideline mock-up, read first part of section 2:**

2 **First impressions and navigation**

> **Alternate between each user test session which “package solution” you show first; either what we call the gist-version or the longer-version.**

**Say:** I’m going to show a compilation of screen shots and I want you to imagine that this is a guideline you can use to answer the clinical problem presented to you. Because this is only a mock-up there are only some of the links that actually work. If you push a button that has no functioning link, I would like you to tell me what kind of information you were seeking.

I am most interested in the content and structure of the guidelines you will be looking at. I wish to ensure that they can be presented in a variety of interactive formats such as web pages on your desktop computer and in the electronic medical record.

When I give you the iPad, I want your first immediate impression, your spontaneous reaction to the screenshots. Don’t think, just tell me the first thing that comes into your head when you see it. After you have had a few minutes to orientate yourself, we will go through each screenshot in detail separately.

> **Now show the mock-up.**

**Ask:** What is your first reaction?

> Make a note of any major navigational issues that occur. Feedback on this will of course be limited by the low-tech nature of the mock-up).

---

3 **Recommendation:**

**Recognition/Initial understanding of the recommendation content/format:**

**Ask:**

- Can you read this and explain what it means to you, using your own words?
- How easy is this recommendation to understand?
- How comfortable and confident would you feel acting on this recommendation as it stands, if it were published by a source you considered reliable?
- Do you prefer having the alternative treatment options listed in the recommendation or not?

### 4 Ask: Do you want or need additional information to act on the recommendation?
If so, please specify what type of information.

### Additional information:
- Facilitate the participant if she wants access to deeper layers of information.
- Keep encouraging the participant to think aloud and to give her first impressions.
- Let the participant know that she should inform you when she would normally stop/feel she has enough information to act on the recommendation, but emphasize that you want her opinion on additional layers of information as well.
- To further encourage the participant’s motivation to seek additional information, consider asking more complex therapeutic questions provided in the clinical scenario.

#### 4a Rationale:

**Ask:** Is the recommendation rationale easily understandable? Explain in your own words what it means.

Additional guiding questions:
- Is it helpful to you?
- Any information you think is lacking?
- Any information you think is superfluous?
- Any suggestions on how to improve the presentation?

**Ask:** where would you like to see this Rationale in a guideline; after the recommendation, here in the key info section, somewhere else?

Ask: would you present it separated from the Key info section, as another tab.

#### 4b Key information:

**Say:** Please comment on the entire key information section first, then we would like you to comment of the individual components (see below)

Is this key information section easily understandable?

Additional guiding questions:
- Is it helpful to you?
- Any information you think is lacking?
- Any information you think is superfluous?
- Any suggestions on how to improve the presentation?
- How do you interpret the color-bars on the left-hand side?

<table>
<thead>
<tr>
<th>Benefits/harms</th>
<th>Explain in your own words what it means:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy to understand?</td>
</tr>
<tr>
<td></td>
<td>Helpful?</td>
</tr>
<tr>
<td></td>
<td>Anything lacking?</td>
</tr>
<tr>
<td></td>
<td>Anything superfluous?</td>
</tr>
<tr>
<td></td>
<td>Any suggestions on how to improve the presentation?</td>
</tr>
<tr>
<td></td>
<td>Comments on the pop-up window with the differing ways of presenting the effect estimates?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall confidence in effect- estimates</th>
<th>Explain in your own words what it means:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy to understand?</td>
</tr>
<tr>
<td></td>
<td>Helpful?</td>
</tr>
<tr>
<td></td>
<td>Anything lacking?</td>
</tr>
<tr>
<td></td>
<td>Anything superfluous?</td>
</tr>
<tr>
<td></td>
<td>Any suggestions</td>
</tr>
<tr>
<td></td>
<td>(Consider showing participant screenshot with written explanation of confidence in effect)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Values and preferences</th>
<th>Explain in your own words what it means:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy to understand?</td>
</tr>
<tr>
<td></td>
<td>Helpful?</td>
</tr>
<tr>
<td></td>
<td>Anything lacking?</td>
</tr>
<tr>
<td></td>
<td>Anything superfluous?</td>
</tr>
<tr>
<td></td>
<td>Any suggestions on how to improve the presentation?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resources</th>
<th>Explain in your own words what it means:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy to understand?</td>
</tr>
<tr>
<td></td>
<td>Helpful?</td>
</tr>
<tr>
<td></td>
<td>Anything lacking?</td>
</tr>
<tr>
<td></td>
<td>Anything superfluous?</td>
</tr>
<tr>
<td></td>
<td>Any suggestions on how to improve the presentation?</td>
</tr>
</tbody>
</table>

| Other |
Appendix 1: Interview guide User testing DECIDE Work Package 1

5

- The EtR-table

*Ask:* Is this table easily understandable? Explain in your own words what it tells you.

Additional guiding questions:
- Is it helpful to you?
- Any information you think is lacking?
- Any information you think is superfluous?
- Any other suggestions?

- The SoF-table

*Ask:* Is this table easily understandable? Explain in your own words what it tells you.

Additional guiding questions:
- How would you compare these two tables?
- Would you prefer the information to be combined differently?
- Do you find it useful to have both?
- Any other suggestions?

6

Understanding strong and weak recommendations

*Say:* We now have some more questions about strength of recommendations. Please have another look at the first recommendation page.

*Ask:* How do you understand statements of strong or weak recommendations?

*Say:* Please take a moment to read this text and tell me what you think.

> *Show the test subject GRADE definition of strong and weak recommendation* (available as a pop-up from the recommendation page by pushing the icons next to the recommendations)

*Ask:* Does this definition make sense to you? Easy to understand? Helpful? Other thoughts?

*Ask:* Do you have any other suggestions or ideas about how to best convey strength of recommendation?
Appendix 1: Interview guide User testing DECIDE Work Package 1

7 **Alternative presentation formats**

> Guide the user back to the “table of contents”. At the bottom right hand corner there is a button labeled “Alternative lay-out”.

**Say:** The next question concerns different ways of presenting the strength of the recommendation.

> Show the alternative first layer of recommendation; introducing weak and strong recommendations + a legend underneath

**Ask:** Would you prefer this way of presenting the strength of the recommendation?

8 **Presenting “package 2”. This will either be the gist-version or the longer-version, depending on which one you started the session with.**

**Say:** I have to emphasize that bundling the different components into two packages is strictly due to practical issues. You should consider the different components as lego-blocks that you can mix and match which ever way you choose. E.g if you prefer one key factor from this alternative presentation to be combined with another from the first presentation please state that.

> Go back again to the table of contents and push the button labeled “Alternative lay-out 2”.

**Ask:** Would you prefer that we present the strength of the recommendation in this way?

> Go back to the next layer.

**Ask:** Which version do you prefer, meaning do you prefer brief or more extensive information to be shown up front?

---

After having seen all of the information layers

9 **Content**

**Say:** We have now presented you with what we call a top layer of information in the guideline. The idea is that you very easily can find more comprehensive information, for example by one click on your smartphone or on the computer screen.

**Ask:** Would you want to see more information in the guideline? If so what kind of information would you want to see?

- What would you now do if you had to act on the recommendation?
| 10 | **Structure/order and formatting:**  
**Ask:** Would you organize the information layers you have seen differently? E.g. in which order would you prefer having the background information (= key information + rationale)?  
- In which order do you feel they could be more helpful?
- Would you change the visual design at all, such as the font or the colors? |
|---|---|
| 11a | **Additional comments and suggestions for improvement:**  
**Ask:** Were there any things that particularly confused or frustrated you, or things you didn’t like?  
- Specific things that you felt were missing?  
- Was there anything that you especially liked? |
| 11b | **Summing up understandability**  
**Ask:** Did you find the information generally easy or generally difficult to understand? |
| 11c | **Summing up usefulness**  
**Ask:** Do you think this way of formatting recommendations would be useful or not so useful for you and your colleagues if you were going to make a clinical decision? (why?) |
| 12 | **Participants suggested alternative presentations of recommendations** |
**Say:** Let’s say you could be in charge and could design the presentation of recommendations exactly how you wanted so that they were most useful for you and your colleagues in your daily practice.

**Ask:** What might an ideal design of recommendations be like? You can use as many sheets of paper as you like?

> Present the test subject with blank iPhone screens on paper and ask them to draw their ideas or concepts.

---

**Say:** Thank you very much – that’s all. But we also would like your feedback on how we might have organised this session better. Any suggestions for improving the user testing?
DECIDE
Developing and Evaluating Communication Strategies to Support Evidence-Based Practice in Clinical Guidelines

New standards for trustworthy guidelines

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

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

Q 1: My age is
1. 25-35
2. 36-45
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
• Risk for stroke? Anticoagulation as prophylaxis?
### CHADS2-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: +1
  - ≥ 65 years old: +1
  - > 75 years old: +1

- **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 stenosis)**
  - Yes: +1

- **Diabetes Mellitus?**
  - If Yes: +1

- **Female?**
  - Yes: +1

**Score**

**Baseline risk for stroke (no anticoagulation treatment)** with CHADS2-VASc score 2

This means an assumed risk of 2.2 strokes per 1000 patients over 1 year

### Treatment for Gabriel?

- **Diagnosis:** Atrial fibrillation
- **Moderate risk of stroke (CHADS2-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

### The traditional steps of evidence-based practice

**Anticoagulation for a patient like Gabriel!**

- **Implement in your practice**
- **Evaluate appropriateness for your individual patient**
- **Critically appraise the research evidence and your confidence in the effect estimates found**

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

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>Disagree</td>
<td>Somewhat Disagree</td>
<td>Somewhat Agree</td>
<td>Agree</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never, rarely</td>
<td>Not followed all steps, but done critical appraisal</td>
<td>Followed all steps weekly</td>
<td>Followed all steps, daily</td>
</tr>
</tbody>
</table>

### 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?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>It’s a strong recommendation</td>
<td>It’s moderate quality evidence</td>
<td>It’s a moderate recommendation</td>
<td>It’s a weak recommendation</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate quality evidence</td>
<td>Low quality evidence</td>
<td>Study design (based on a single Randomized study)</td>
<td>Study design (based on a smaller systematic review)</td>
</tr>
</tbody>
</table>
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 your hand) when you are ready to answer.
6. Then switch and the other group gets blindfolds.

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!

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 S 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.)

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!
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 rivaroxaban). Should experience in real world 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 | 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 rivaroxaban). Should experience in real world 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? |
|---------------------------|----------------|----------------|----------------|----------------|
| 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 |
| 1  | 2  | 3  | 4  |

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.

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.10: Which, if any, antithrombotic treatment would you consider appropriate for Gabriel? |
|---------------------------|----------------|----------------|
| NOAC (Dabigatran, rivaroxaban or apixaban) | Aspirin | Warfarin |
| 1  | 2  | 3  | 4  |

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

<table>
<thead>
<tr>
<th>NOAC (Dabigatran, rivaroxaban or apixaban)</th>
<th>Aspirin</th>
<th>Warfarin</th>
<th>No therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Summary of anticoagulant monotherapy: Anticoagulants with each of the newer agents (dabigatran, rivaroxaban, and apixaban) lead 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 risk of myocardial infarction with dabigatran and twice daily regimen (dabigatran and apixaban). Should experience in real world populations mirror the net clinical benefit found in randomized trials, our confidence in the superiority of these drugs will increase. (See “Similarities” above.)

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

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?

<table>
<thead>
<tr>
<th>Benefits and harms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Benefits
- 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 some evidence of fewer intracranial bleeds with dabigatran, resulting in a absolute risk reduction of 2.2 per 1,000 patients
  - The absolute risk, however, is generally very low: 5/1,000 with warfarin, 3/1,000 with dabigatran.

Practical consequences: Daily medication with all. Regular INR controls and dietary restrictions with warfarin.

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

<table>
<thead>
<tr>
<th>Format A</th>
<th>Format B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Format A: We recommend treatment with an anticoagulant (e.g., dabigatran, rivaroxaban, or apixaban) rather than warfarin.

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

**Broad consensus**

**Common methods for guidelines**

<table>
<thead>
<tr>
<th>GRADE defines strength of recommendation as:</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;The strength of a recommendation reflects the extent to which we can, across the range of patients for whom the recommendation is intended, be confident that desirable effects of a management strategy outweigh undesirable effects.&quot;</td>
</tr>
</tbody>
</table>

**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"

<table>
<thead>
<tr>
<th>Q18a: &quot;I fully understand the difference between strong and weak recommendations and the implications for clinical decision making&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Strongly Disagree</td>
</tr>
</tbody>
</table>

**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"

<table>
<thead>
<tr>
<th>Q18b: &quot;This explanation is necessary to understand the difference between strong and weak recommendations and the implications for clinical decision making&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Strongly Disagree</td>
</tr>
</tbody>
</table>

**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 is 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.

<table>
<thead>
<tr>
<th>Q19a: &quot;I fully understand the difference between the different categories of quality and the implications for clinical decision making&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Strongly Disagree</td>
</tr>
</tbody>
</table>

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?
**Choice of oral anticoagulation**

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

**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**

- **Neurovascular outcomes:** Reversibility of stroke in patients treated for 3 weeks.
- **Death and stroke:** Less significant difference.
- **Major bleeding:** Overall no relevant difference, but the number of fatal hemorrhages was reduced with dabigatran, resulting in a reduction of 2 per 1000 patients.
- **Minor bleeding:** Overall no significant difference. The incidence is adaptive, which may increase the risk compared to warfarin. The absolute risk, however, is greatly overestimated. 5-10/1000 with warfarin, 6-10/1000 with dabigatran.

**Treatment discontinuation:** Due to side effects: 5% interruption with warfarin, 2% with NOAC.

**Practical considerations:** Only medication with NOAC. Regular INR controls and dietary restrictions with warfarin.

**Quality of evidence**

- Moderate. The expected effects of NOAC compared with warfarin’s failure in systematic review with meta-analysis and individual patient data, which had confidence intervals for death and bleeding. Dabigatran was associated with an increase in major hemorrhage and treatment discontinuation in its relative subgroup analysis.

**Preference and values**

Studies on anticoagulant and values have shown that the average patient is expected to suffer 3 major bleeds in 10 years. These studies have been guiding our recommendations. Their views, however, demand lack of low quality and there’s high degree of variability in preferences. We therefore suggest that the decision regarding treatment options is made together with the patient.

**Resources**

Cost did not influence this recommendation.

---

**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**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Confidence Interval</th>
<th>Relative Effect</th>
<th>NOAC</th>
<th>Oral-Shaking Patients</th>
<th>Warfarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality (1 year)</td>
<td>High</td>
<td>9% (10% - 10%)</td>
<td>7% (7% - 9%)</td>
<td>6% (6% - 7%)</td>
<td>0.8 (0.8 - 0.8)</td>
</tr>
<tr>
<td>Major bleeding (1 year)</td>
<td>Moderate</td>
<td>3% (3% - 4%)</td>
<td>3% (3% - 4%)</td>
<td>3% (3% - 4%)</td>
<td>0.8 (0.8 - 0.8)</td>
</tr>
<tr>
<td>Major bleeding (5 years)</td>
<td>Moderate</td>
<td>6% (6% - 7%)</td>
<td>6% (6% - 7%)</td>
<td>6% (6% - 7%)</td>
<td>0.8 (0.8 - 0.8)</td>
</tr>
</tbody>
</table>

---

**Thank you for your participation**

For more information go to:

- [http://www.decide-collaboration.eu](http://www.decide-collaboration.eu)
- [http://www.gradeworkinggroup.org](http://www.gradeworkinggroup.org)
- [http://www.magicproject.org](http://www.magicproject.org)

If there is time left we can look at your results.
### Taxonomy table for step 2 and 3 of the adaptation

<table>
<thead>
<tr>
<th>Registration form for modifying recommendations in the chapter: VTE, Thrombophilia, Antithrombotic Therapy, and Pregnancy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction to the panelist:</strong> Please complete this form and return to your chapter editor by March 10th. For each recommendation: Please assess whether to exclude or include. For the included recommendations – assess the need for modifications, and if so how and why. Explain by using your own words, and/or you can apply the taxonomy (for details we refer to the project protocol). If you, after having read the whole chapter, see the need for new recommendations, please register these at the end of the document.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Do you wish to exclude? If yes, explain why and move on to the next rec. If you wish to keep the rec, move on to the next column.</th>
<th>Do you wish to modify? This pertains to your first evaluation, not the final decision within the group. If no – move on to the next rec.</th>
<th>If you believe there is a need to modify the rec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1. For pregnant patients, we recommend LMWH for the prevention and treatment of VTE, instead of UFH (Grade 1B)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Reason:</td>
<td>If yes, note any conflicts of interest and move on to the next column. None: Primary: Secondary: Financial:</td>
<td>Comment:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PICO Studies Rel. effect estimate Effect size Benefit Harm Values Cost Other</td>
</tr>
</tbody>
</table>