Sharing Evidence to Inform Treatment Decisions (SHARE-IT):

Generic tools for shared decision-making linked to evidence summaries and clinical practice guidelines

Anja Fog Heen

Department of Medicine

Gjøvik

Innlandet Hospital Trust

Department of Health Management and Health Economics

Institute of Health and Society

Oslo University

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-348-0333-8

Cover: UiO.
Print production: Graphic center, University of Oslo.

Table of contents

Acknowledgements

The ideas and research presented in this thesis would not be possible without my patients. I am deeply grateful for the privilege of the bedside, which has provided valuable insights into the real-world decision making of medicine. These insights have played a pivotal role in creating, refining, and rejecting ideas and research presented in this thesis.

I was brought into the world of evidence-based medicine, guidelines and shared decision-making by my internship supervisor, Dr. Linn Brandt. Together with Professor Per Olav Vandvik, my main PhD supervisor, both of you have throughout the years been innovative, enthusiastic, supportive, and never afraid of challenges or of thinking outside the box. Thank you for including me in MAGIC, for your collaborative and innovative personalities, and for your idealism to improve patient care.

A very special thank you goes to my main research partner, co-supervisor and friend, Professor Thomas Agoritsas. Your extraordinary ability to have clarity in the most complex situations is mesmerizing. Your wit, wisdom and support have been vital throughout the research journey presented in this thesis. I am deeply grateful for your insights, both within the realm of research and beyond.

Distinguished professor Gordon Guyatt, my co-supervisor, has offered profound insights into evidence-based medicine and GRADE. Thank you for consistently providing precise, direct, and thoughtful feedback with admirable patience and for sparking a joy of writing I did not know excised. Together with Dr. Vandvik, Dr. Brandt and Dr. Agoritsas, you have fostered a collaborative research and innovation environment in MAGIC that thrives, consisting of colleagues and friends that are nice and fun to work with.

A special thank you goes also to my dear friend and colleague, Professor Victor Montori. Your unwavering support throughout the years has been invaluable. Thank you for your commitment to careful and kind care, for your thinking and writing, which resonate with and inspire so many.

My work would not have been possible without the support of Innlandet Hospital Trust and through grants from the South-Eastern Norway Regional Health Authority. I especially want

to thank Dr. Øystein Stubhaug, Dr. Even Reinertsen and Dr. Ingvar Stokstad. Your clinical skills, leadership and always having the patient's best interest at heart have been of deep importance both in my clinical work and research.

A warm thank you goes also to Mette Hagen Farstad. Your ability to find solutions to any issue or problem fast and seemingly effortless has been crucial throughout the years and I am very grateful for all your help and support.

Furthermore, I also want to thank Professor Mette Kalager, Professor Michael Bretthauer, and the Clinical Effectiveness Research Group for generously sharing office space and for continuously inspire to critical thinking and research.

I also want to thank the former and current leaders of the Medical Department at Lovisenberg Diaconal Hospital, Dr. Gudmund Nordby and Dr. Bjørn Jardar Brandsæter for providing flexibility in combining research and clinical work in an inspiring work environment.

Many others have provided invaluable support throughout this journey. I would like to express my heartfelt gratitude to my parents and brothers, Tollef and Kristian, for their unwavering support, open discussions, and encouragement of critical thinking.

Last, I want to thank my dear husband Harald for always supporting me and standing by my side. Without you, my world of medicine would not exist. Thank you.

Anja Fog Heen
Oslo, July 2023

Abbreviations

AGREE II	Appraisal of Guidelines for Research and Evaluation
AHRQ	Agency for Healthcare Research and Quality
API	Application Programming Interfaces
BMJ RapidRecs	BMJ Rapid Recommendations
CASP	Qualitative Research Checklist of the Critical Appraisal Skills Programme
COMRADE	Combined outcome measure for risk communication and treatment decision making effectiveness
DECIDE	Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence
EBM	Evidence-based Medicine
EtD	Evidence to Decisions
GIN	Guideline International Network
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GRADE-CerQual	GRADE -Confidence in the Evidence from Reviews of Qualitative research
HTA	Health Technology Assessment
IPDAS	International Patient Decision Aid Standards
LIPUS	Low intensity pulsed ultrasound
MAGIC	MAGIC Evidence Ecosystem Foundation
MAGICapp	MAGIC Authoring and Publication platform
MATCH-IT	Making Alternative Treatment Choices Intuitive and Trustworthy
MOOSE	Meta-analyses Of Observational Studies in Epidemiology checklist for reporting
NEATS	National Guideline Clearinghouse Extent of Adherence to Trustworthy Standards instrument
NICE	National Institute for Health and Clinical Excellence
PICO	Population (P), Interventions (I), Comparisons (C) and outcomes (O)
PDF	Portable Document Format
RIGHT	Reporting Items for practice Guidelines in HealThcare
RCT	Randomized controlled trial
SAVR	Surgical aortic valve replacement
SDM	Shared Decision-Making
SHARE-IT	Sharing Evidence to Inform Treatment Decisions
SNAP-IT	Smooth National Adaptation and Presentation of guidelines to Improve Therapy
SoF	Summary of Findings
TAVI	Transcatheter aortic valve insertion
UCD-11	User-Centered Design 11-measure
WHO	World Health Organization

Thesis summary

Background: Clinicians rely on trustworthy guidelines for decision-making at the point of care, while encounter decision aids support shared decision-making (SDM). However, both guidelines and decision aids face challenges in production, underlying evidence, uptake, and practicality in clinical practice. They often ignore the burden of treatment and practical issues that patients' needs to consider when implementing an intervention into their daily life. Trustworthy guidelines need to incorporate patients' values and preferences for management options, but often fail to do so. These challenges warrant generic and coordinated solutions.

Aim: In the Sharing Evidence to Inform Treatment decisions (SHARE-IT) project, we aimed to: 1) develop a framework for the generic translation of Grading of Recommendations Assessment, Development and Evaluation (GRADE) evidence summaries and guidelines into encounter decision aids; 2) design and user-test a set of interactive, adaptable presentation formats of decision aids and integrate these in an authoring and publication platform (MAGICapp); 3) develop a framework of patient-important practical issues on management options, integrate this in MAGICapp and explore the feasibility of including of such issues in real-life production and publication of guidelines and encounter decision aids and; 4) conduct a systematic review on patients' values and preferences to inform a guideline panel creating recommendations.

Material and methods: We applied human-centred design principles, International Patient Decision Aid Standards (IPDAS) and GRADE methods to develop decision aid presentation formats linked to guidelines. Using grounded theory, we identified and categorized practical issues from HealthTalk.org registry and Option Grids to develop a generic framework and presentation format for practical issues.

We iteratively user-tested a developed decision aid prototype and practical issues framework in real-life consultations and conducted semi-structured interviews of participants. We performed qualitative content analysis and categorized the results using a revised Morville's framework of user-experience. We conducted a systematic review update on patients' values on aortic valve insertion (TAVI) and surgical aortic valve replacement (SAVR) for aortic stenosis using established methods for evidence synthesis, following a checklist for

systematic review update, using a validated search filter, and assessing the quality of evidence using GRADE and CerQual.

Results: Following feedback from stakeholders and a multidisciplinary team, we developed an initial multilayered framework for translating GRADE evidence summaries into encounter decision aids. Through user-testing in 28 real-life consultations, we iteratively developed digital, interactive, and adaptable encounter decision aid presentation formats and integrated these in MAGICapp.

Clinicians and patients found the tool to facilitate SDM and easy to use. However, patients wanted more information on the impact of interventions on their daily life and treatment burden, while clinicians lacked tools for supporting these conversations. Consequently, we developed a framework of 15 practical issues categories and integrated these in MAGICapp resulting in a framework of practical issues in decision aids and guidelines. Implementation in 15 BMJ Rapid Recommendations added 283 issues to 35 recommendations.

Evidence of suboptimal rigor on patients' values and preferences on TAVI and SAVR for aortic stenosis, conducted to inform a BMJ RapidRecs guideline panel, showed that participants were willing to accept a higher mortality risk than suggested by current evidence, preferred minimally invasive procedures, and found improvements in health-related quality of life domains as reasons to undergo treatment. This systematic review lacked patient-important outcomes and only addressed a few practical issues.

Conclusion: SHARE-IT encounter decision aids, developed within MAGIC, demonstrate feasibility of creating decision aids semi-automatically using digitally structured data from evidence summaries. These decision aids are enhanced by incorporating a framework of practical issues, thereby promoting holistic decision-making in clinical settings. By systematically incorporating practical issues alongside summary of findings tables and studies on patients' values and preferences in real-life guideline development, guideline developers are better equipped to provide recommendations that address the diverse needs of patients and clinicians. Further research is needed to explore how practical issues align within the evidence ecosystem and determine the optimal method for their inclusion.

Sammendrag

Bakgrunn: Troverdige retningslinjer er nødvendig for å kunne ta gode faglige beslutninger i klinikken. Samvalgsverktøy kan støtte samvalg mellom helsepersonell og pasienter. Både samvalgsverktøy og retningslinjer har imidlertid felles problemer knyttet til produksjon, underliggende kunnskapsgrunnlag, implementering og praktisk bruk i en klinisk hverdag. Både retningslinjer og samvalgsverktøy hensyntar ofte ikke behandlingsbyrde og praktiske forhold knyttet til pasienters implementering av behandling i egen hverdag. Inkorporering av pasienters verdier og preferanser knyttet til behandlingsalternativer er nødvendig i troverdige retningslinjer, men dette er ofte ikke tilfellet. Disse utfordringene nødvendiggjør generiske og koordinerte løsninger.

Mål: I prosjektet "Deling av evidens for å informere behandlingsbeslutninger" (SHARE-IT) var målet å: 1) utvikle et rammeverk for generisk oversetting av kunnskapsoppsummeringer og retningslinjer for utviklet med Grading of Recommendations Assessment, Development and Evaluation (GRADE) til samvalgsverktøy; 2) designe og brukerteste interaktive, adaptive presentasjonsformater for samvalgsverktøy og integrere disse i en publiseringsplattform (MAGICapp); 3) utvikle et rammeverk for praktiske forhold knyttet til behandling, integrere dette i MAGICapp og utforske gjennomførbarheten av å inkludere praktiske forhold i reell utvikling av retningslinjer og samvalgsverktøy; og 4) gjennomføre en systematisk oversikt om pasientenes verdier og preferanser med hensikt å informere et BMJ Rapid Recommendations som utvikler behandlingsanbefalinger.

Materiale og metode: Vi anvendte prinsipper for brukersentert design, internasjonale standarder for samvalgsverktøy (IPDAS) og GRADE-metoder for å utvikle presentasjonsformater for samvalgsverktøy knyttet til retningslinjer. Ved hjelp av databasert teoriutvikling identifiserte og kategoriserte vi praktiske forhold fra databasene HealthTalk.org og samvalgsverktøyene Option Grids for å utvikle et generisk rammeverk og presentasjonsformat for praktiske forhold knyttet til behandling.

Vi gjennomførte iterativ brukertesting av en utviklet prototype for samvalgsverktøy og et rammeverk for praktiske forhold i reelle konsultasjoner, og gjennomførte semistrukturerte intervjuer med deltakerne. Vi utførte kvalitativ innholdsanalyse og kategoriserte resultatene

ved hjelp av et revidert rammeverk (Morville) for brukeropplevelse. Vi utarbeidet en systematisk oversikt over pasienters verdier og preferanser knytte til behandling av aortastenose ved bruk av etablerte metoder for kunnskapsoppsummeringer. Vi brukte et validert søkefilter for pasienters verdier og preferanser og vurdere kvaliteten på dokumentasjonen ved hjelp av GRADE og CerQual.

Resultater: Etter tilbakemeldinger fra en tverrfaglig gruppe og sentrale interessenter utviklet vi et lagdelt rammeverk for å oversette GRADE-oppsummeringer til samvalgsverktøy. Gjennom brukertesting i 28 reelle konsultasjoner utviklet vi iterativt digitale, interaktive og adaptive presentasjonsformater for samvalgsverktøy og innlemmet disse i MAGICapp. Klinikere og pasienter fant verktøyet nyttig for samvalg og enkelt å bruke. Imidlertid ønsket pasientene mer informasjon om hvordan intervensjoner påvirker deres hverdagsliv og behandlingsbyrde, mens klinikere manglet verktøy for å understøtte disse samtalene. Vi utviklet derfor et rammeverk med 15 kategorier av praktiske forhold og integrerte disse i MAGICapp, resulterende i et rammeverk for praktiske forhold i samvalgsverktøy og retningslinjer. Implementeringen i 15 BMJ Rapid Recommendations førte til 283 praktiske forhold i 35 anbefalinger.

Pasienters verdier og preferanser for behandling av aortastenose viste at deltakerne var villige til å akseptere en høyere mortalitetsrisiko enn faktisk mortalitetsrisiko, foretrakk minimalt invasive prosedyrer og betraktet forbedring i livskvalitet som grunn til å gjennomgå behandling. Tillit til dokumentasjonen var lav. Denne systematiske oversikten manglet utfallsmål viktige for pasienter og adresserte bare noen praktiske forhold knyttet til behandling.

Konklusjon: Samvalgsverktøy utviklet i SHARE-IT og MAGIC muliggjør halvautomatisk produksjon av samvalgverktøy fra GRADE-oppsummeringstabeller. Disse samvalgsverktøyene styrkes av et integrert rammeverk for praktiske forhold knyttet til behandling, noe som fremmer helhetlig, klinisk beslutningstaking. Utvikling av retningslinjer blir bedre rustet til å gi anbefalinger som imøtekommer ulike behov til pasienter og klinikere ved å systematisk inkludere praktiske forhold sammen med GRADE-oppsummeringstabeller og studier om pasienters verdier og preferanser. For å utforske hvordan praktiske forhold harmonerer med et økosystem av kunnskap og for å bestemme den best mulige metoden for inkludering praktiske forhold i økosystemet er det behov for ytterligere forskning.

Articles in the thesis

1. **Decision aids that really promote shared decision making: the pace quickens**
   Thomas Agoritsas, Anja Fog Heen, Linn Brandt, Pablo Alonso-Coello, Annette Kristiansen, Elie Akl, Ignacio Neumann, Kari Tikkinen, Trudy van der Weijden, Glyn Elwyn, Victor M. Montori, Gordon H. Guyatt, Per Olav Vandvik
   BMJ 2015 Feb 10; 350:g7624.

2. **Decision aids linked to evidence summaries and clinical practice guidelines: results from user-testing in clinical encounters**
   Anja Fog Heen, Per Olav Vandvik, Linn Brandt, Gordon Henry Guyatt, Shaun Treewek, Thomas Agoritsas
   BMC Medical Informatics and Decision Making 2021 Jun 29;21(1):202.

3. **A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines**
   Anja Fog Heen, Per Olav Vandvik, Linn Brandt, Victor M. Montori, Lyubov Lytvyn, Gordon Guyatt, Casey Quinlan, Thomas Agoritsas
   Journal of Clinical Epidemiology 2021 Jan;129:104-113. Epub 2020 Oct 10.

4. **Patient values and preferences on valve replacement for aortic stenosis: a systematic review**
   Anja Fog Heen, Lyubov Lytvyn, Michael Shapiro M, Gordon H. Guyatt, Reed Alexander Cunningham, Yuan Zhang, Veena Manja, Per Olav Vandvik, Thomas Agoritsas
   Heart. 2021 Aug;107(16):1289-1295. Epub 2021 Feb 9.

1. Background

1.1 MAGIC and SHARE-IT

A little over a decade ago, the idea of directly linking tools for shared decision-making (SDM) (i.e., encounter decision aids) to clinical practice guidelines, henceforth *guidelines*, came up in the Making GRADE the Irresistible Choice (MAGIC) research and innovation program, later to become the MAGIC Evidence Ecosystem Foundation (1). With my particular interest in SDM, this idea formed the foundation of my thesis and a journey of iterative research and innovations towards a new generation of tools for SDM to be used in the clinical encounter for supporting patient-centered care.

The overarching objective of MAGIC, initiated to solve core problems with guidelines, was to provide clinicians and patients with user-friendly decision support tools that would facilitate well-informed decisions at the point of care. MAGIC, firmly embedded within concepts of evidence-based medicine (EBM) and patient-centered care, adds digital solutions to standards and methods for trustworthy guidelines to enhance tools for decision support.

As can be deduced from the MAGIC acronym, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system and its specific methods for critical appraisal of research evidence and guideline development was considered appropriate to solve the identified problems with guidelines. In the absence of available platforms to create, disseminate and update guidelines in user-friendly and digital formats more easily, MAGIC developed an online authoring and publication platform (MAGICapp) that was launched in 2013 (see more in paragraph 1.7.1) (2).

An early insight in MAGIC was that most recommendations in trustworthy guidelines necessitate SDM, as most management choices are sensitive to patients' values and preferences (3-7). SDM is the collaborative process in which patients and clinicians make decisions together as partners, using best available evidence. This process involves discussing the available options, their associated benefits and harms, and considering the patient's values, preferences, and circumstances (7-9). Including patients in decisions that are relevant to them is not only an ethical imperative, but also a fundamental aspect of EBM and patient-centered care (3, 7, 10-12).

EBM has undergone a gradual shift in focus, recognizing the importance of guidance to inform medical decision-making rather than providing best current evidence alone. As a result, guidelines have become predominant tools, largely replacing systematic reviews as the primary sources for evidence-informed decisions (5, 13). With an increased emphasis on patient-centered care and SDM, the development of decision aids has emerged as tools to support patients and their clinicians in making well-informed decisions together (14, 15).

Inspired by pioneering work on encounter decision aids – at the time a new type of SDM tools to be used in the clinical encounter (16, 17), by our colleague Dr. Victor Montori and his team in the Knowledge and Evaluation Research Unit at the Mayo Clinic, Rochester, USA, we initiated the Sharing Evidence to Inform Treatment Decisions (SHARE-IT) project in 2012.

Concurrently, we realized that both guidelines and decision aids face similar challenges. The resources required to summarize evidence for trustworthy decision aids may be similar to those needed for producing a systematic review or a guideline. Their production is time-consuming, they are often not based on the best available evidence, often rapidly become outdated, their uptake is highly variable, and they tend to be unwieldy in the constraints of real-life clinical practice (5, 18-22). These challenges call for generic and coordinated solutions.

This thesis reports the findings of the SHARE-IT project and our development of generic encounter decision aids linked to evidence summaries and guidelines in digitally structured formats, through MAGICapp. The revelation of burden of treatment and practical issues as core components of SDM tools has also become central to inform this thesis. Through the real-life implementation of our encounter decision aids and practical issues in guidelines, most notably The BMJ Rapid Recommendations, henceforth BMJ RapidRecs, the results also hold implications for the current development, publication and updating of guidelines within an enhanced evidence ecosystem.

1.2 EBM and the evolution of GRADE

When conceived in 1992, EBM represented a new paradigm for medical decision-making. It was originally defined as the conscientious, explicit, and judicious use of current best evidence when making decisions about the care of individual patients (11). This recognizes

that relevant scientific evidence, clinical judgement, and patients' values and preferences are key components (Figure 1). By incorporating these elements, EBM transformed the practice of medicine, grounding it in scientific rigor and empiricism (10, 23), while also highlighting the importance of clinical expertise and patients' preferences as integral parts of the decision-making process (10).

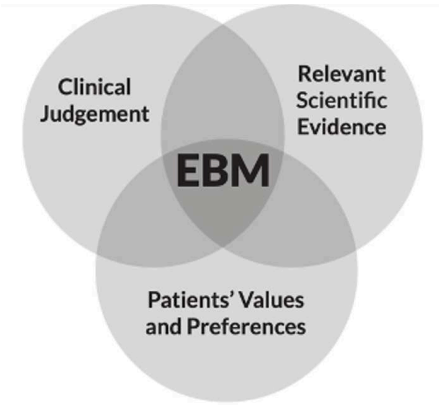


Figure 1. Key components clinical decision-making in evidence-based medicine (11)

Although critical appraisal of research evidence has always been at the core of EBM (11), the field has witnessed significant advancements in methods and tools for decision support over its 30-year evolution. These methods and tools aim to facilitate the answering of clinical questions concerning diagnosis, prognosis and treatment, for health care professionals, patients, and policy-makers (13). Structured "PICO" formats have proven effective in formulating such questions, defining the population (P), interventions (I), comparisons (C) and patient-important outcomes (O) (24).

The process of critical appraisal has evolved from using checklists to adopting more systematic and transparent approaches to appraise the body of evidence, particularly through systematic reviews (25, 26). Among these approaches, GRADE emerged as the predominant system since its introduction in 2004. GRADE also includes guidance for how to move from evidence to decisions, such as recommendations in guidelines (27, 28). GRADE maintains a clear distinction between the certainty of the evidence (from critical appraisal) and the strength of recommendations (applicable to guidelines) (29-31).

GRADE is most advanced for management issues, with detailed guidance for systematic reviews and meta-analysis of randomized trials as well as for recommendation development

in guidelines (25, 26, 30, 31). However, GRADE methods are increasingly developed for questions concerning diagnosis (32, 33), prognosis (34) and values and preferences (35-37). GRADE is widely adopted and used by more than 110 organizations worldwide (25, 29, 38).

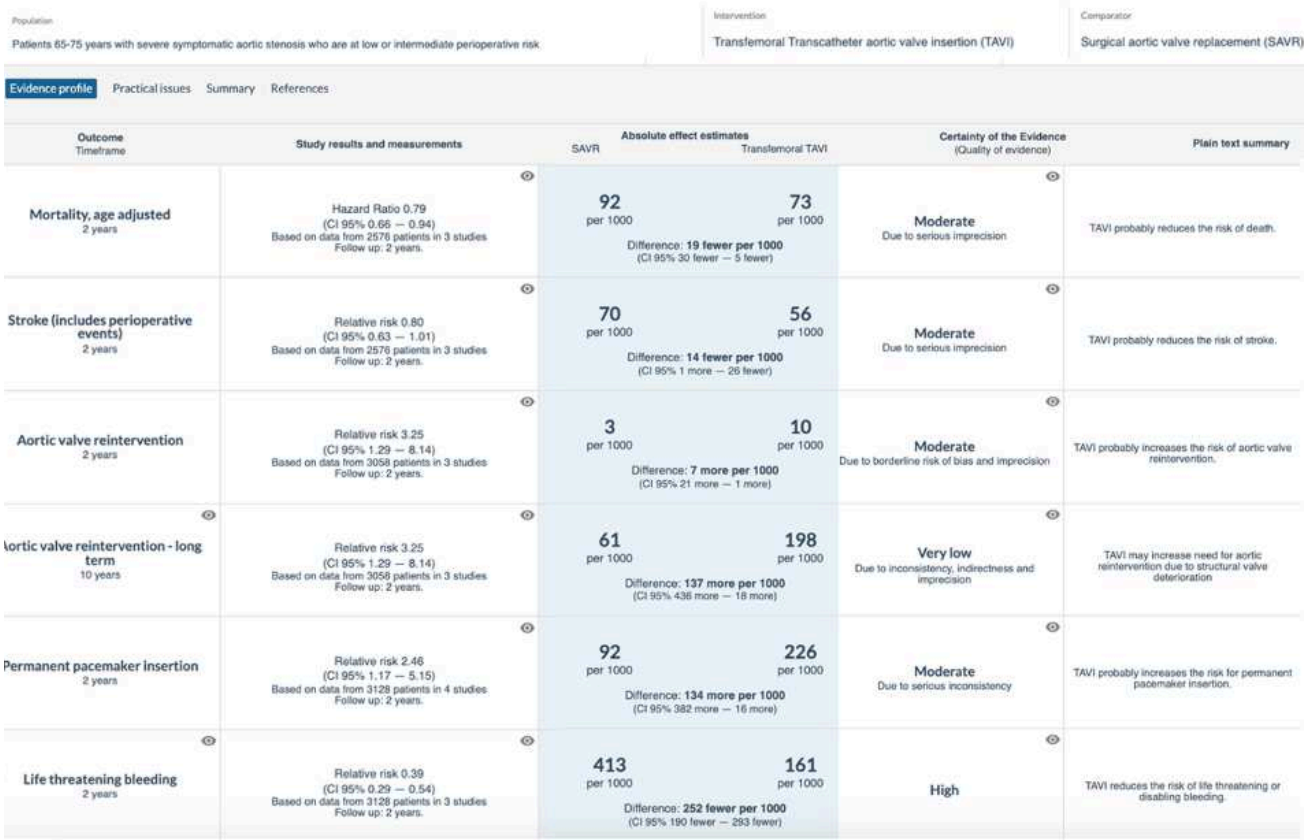
1.2.1 GRADE evidence summaries and certainty of evidence

The critical appraisal of research evidence using GRADE is informed by the body of evidence derived from systematic reviews, based on PICO questions (24). When rating the certainty of the body of evidence, GRADE specifies four categories of certainty (high, moderate, low, and very low). The term "certainty of evidence" is synonymous with "quality of evidence" and encompasses an overall rating for each patient-important outcome across studies.

Several factors can reduce the quality of evidence, including limitations in study design or execution, imprecision, inconsistency of results, indirectness of evidence, and publication bias. Conversely, factors that increase the quality of evidence include a large magnitude of effect, that all plausible confounding would reduce the demonstrating effect or increase the effect if no effect was observed, and a dose-response gradient (29, 39). In the context of making recommendations, the certainty of evidence reflects the confidence that the effect estimates are adequate to support a particular decision or recommendation (29, 40).

Evidence summaries, typically in the form of evidence profiles or Summary of Findings (SoF) tables, are key outputs from critical appraisal using GRADE. These summaries provide relative and absolute effects as well as certainty of the evidence for patient-important outcomes. Additionally, they include a plain-text summary of the findings (41). Research conducted to improve the user experience of summarized evidence show that SoF tables improve comprehension of systematic reviews (42, 43). These summaries encompass many essential components necessary for developing encounter decision aids, which aligns with the objective of the SHARE-IT project.

Figure 2 shows a SoF table for severe symptomatic aortic stenosis with structured presentation of study results, absolute effect estimates, certainty of evidence, and plain text summary on patient-important outcomes of the treatment alternatives (transcatheter aortic valve insertion (TAVI) versus surgical aortic valve replacement (SAVR) (44).



Population
Patients 65-75 years with severe symptomatic aortic stenosis who are at low or intermediate perioperative risk

Intervention
Transfemoral Transcatheter aortic valve insertion (TAVI)

Comparator
Surgical aortic valve replacement (SAVR)

Evidence profile | Practical issues | Summary | References

Outcome Timeframe	Study results and measurements	Absolute effect estimates SAVR	 Transfemoral TAVI	Certainty of the Evidence (Quality of evidence)	Plain text summary
Mortality, age adjusted 2 years	Hazard Ratio 0.79 (CI 95% 0.66 — 0.94) Based on data from 2576 patients in 3 studies Follow up: 2 years.	92 per 1000	73 per 1000 Difference: **19 fewer per 1000** (CI 95% 30 fewer — 5 fewer)	Moderate Due to serious imprecision	TAVI probably reduces the risk of death.
Stroke (includes perioperative events) 2 years	Relative risk 0.80 (CI 95% 0.63 — 1.01) Based on data from 2576 patients in 3 studies Follow up: 2 years.	70 per 1000	56 per 1000 Difference: **14 fewer per 1000** (CI 95% 1 more — 26 fewer)	Moderate Due to serious imprecision	TAVI probably reduces the risk of stroke.
Aortic valve reintervention 2 years	Relative risk 3.25 (CI 95% 1.29 — 8.14) Based on data from 3058 patients in 3 studies Follow up: 2 years.	3 per 1000	10 per 1000 Difference: **7 more per 1000** (CI 95% 21 more — 1 more)	Moderate Due to borderline risk of bias and imprecision	TAVI probably increases the risk of aortic valve reintervention.
\ortic valve reintervention - long term 10 years	Relative risk 3.25 (CI 95% 1.29 — 8.14) Based on data from 3058 patients in 3 studies Follow up: 2 years.	61 per 1000	198 per 1000 Difference: **137 more per 1000** (CI 95% 436 more — 18 more)	Very low Due to inconsistency, indirectness and imprecision	TAVI may increase need for aortic reintervention due to structural valve deterioration
Permanent pacemaker insertion 2 years	Relative risk 2.46 (CI 95% 1.17 — 5.15) Based on data from 3128 patients in 4 studies Follow up: 2 years.	92 per 1000	226 per 1000 Difference: **134 more per 1000** (CI 95% 382 more — 16 more)	Moderate Due to serious inconsistency	TAVI probably increases the risk for permanent pacemaker insertion.
Life threatening bleeding 2 years	Relative risk 0.39 (CI 95% 0.29 — 0.54) Based on data from 3128 patients in 3 studies Follow up: 2 years.	413 per 1000	161 per 1000 Difference: **252 fewer per 1000** (CI 95% 190 fewer — 293 fewer)	High	TAVI reduces the risk of life threatening or disabling bleeding.

Figure 2. GRADE SoF table showing benefits and harms of TAVI and SAVR for patients with severe, symptomatic aortic stenosis who are at low or intermediate perioperative risk

1.2.2 GRADE for qualitative evidence synthesis

The Confidence in the Evidence from Reviews of Qualitative research (GRADE-CERQual) method has been developed "to support the use of findings from qualitative evidence syntheses in decision-making". It complements GRADE for quantitative evidence synthesis.

GRADE-CERQual provides a framework for assessing the confidence in evidence from qualitative evidence syntheses based on four components: methodological limitations, coherence, adequacy of data, and relevance. This approach enables transparent assessments of the confidence in results and supports the use of findings from qualitative evidence syntheses in decision-making processes. The results are often presented in SoF tables (45, 46). SoF tables have proven to be valuable in making evidence more useful and accessible (42, 43, 47).

1.3 Clinical practice guidelines

In today's healthcare, trustworthy guidelines play a ubiquitous role as clinicians rely on them to support decision-making at the point of care (5). Clinical practice guidelines are defined as "statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options". This definition comes from a comprehensive report: *Clinical Practice Guidelines We Can Trust*, published by the Institute of Medicine (now the National Academy of Medicine) in 2011 (5).

This report proposed eight standards for the development and identification of trustworthy guidelines with the goal of enhancing the quality of care and patient outcomes. These standards include: 1) emphasizing transparency of development and funding; 2) detailed disclosure and management of conflict of interest; 3) composition of the guideline development group; 4) use of systematic reviews to inform guideline recommendations; 5) establishing evidence foundations and rating strength of guideline recommendations; 6) articulation of recommendations; 7) external review; and 8) updating (5). Similar standards have also been issued by organizations such as Guideline International Network (GIN) (48), National Institute for Health and Clinical Excellence (NICE) (49), and the World Health Organization (WHO) (50, 51).

These standards, which have gained acceptance from key guideline organizations (51), are used both for the development of guidelines and for assessing the trustworthiness of published guidelines. Several validated tools, such as the National Guideline Clearinghouse Extent of Adherence to Trustworthy Standards instrument (NEATS), Appraisal of Guidelines for Research and Evaluation (AGREE II) and Reporting Items for practice Guidelines in HealThcare (RIGHT), are used for this purpose (52-54).

1.3.1 GRADE methods for guidelines

To meet these standards, guidelines should include all relevant factors required to move from evidence to decisions in a systematic and transparent process (5, 27, 28, 55). For this purpose, GRADE provides advanced guidance through the Evidence to Decision framework, (EtD). The process of moving from evidence to recommendations includes assessing the balance

between benefits and harms, certainty of the evidence, values and preferences, and other considerations such as resources, acceptability, feasibility, and equity (27, 28, 55).

The extent to which resources, acceptability, feasibility, and equity are explicitly considered depends on the focus of the guideline. While these factors are clearly relevant in from a healthcare system perspective (e.g., national guideline for public health care), guidelines that take an individual patient perspective (e.g., BMJ RapidRecs) typically do not fully incorporate these factors when moving from evidence to recommendations (27, 28, 44).

1.3.2 Patient' values and preferences in guidelines

To develop trustworthy recommendations in guidelines, guideline panels need to consider and incorporate patients' values and preferences regarding management options and the factors relevant to decision-making (5, 27, 28, 36). In this context, values and preferences can be described as "the relative importance patients place on outcomes" for management decisions (36). They encompass patient perspectives, beliefs, expectations, and goals related to their health and overall well-being. This involves patients weighing the potential benefits, harms, costs, and burdens associated with different options (56).

Incorporating patients' values and preferences in guidelines is driven by several factors. Firstly, it aligns with the overarching ethical imperative of patient-centered care and autonomy, emphasizing that patient values should guide clinical decisions (3, 6, 12, 57). Secondly, recommendations that align with patient values and preferences may be more easily accepted and implemented by those intended to benefit from the guidelines (6, 36, 58, 59). Thirdly, contributions of patients and the public can help identify patient-important outcomes, assess the meaningfulness of findings, and evaluate how recommendations interact with patients' values. Finally, recommendations that consider patients' preferences can better facilitate SDM (6, 30, 59, 60).

The inclusion of patients' values and preferences in guidelines can be achieved through high-quality systematic reviews of studies focusing on patients' values and preferences, as well as active participation of patients and the public in the guideline development process (5, 27, 30, 36, 58, 61). However, this is far from routine when guidelines are developed. Armstrong and Bloom reported that only 8% of the 101 organizations reviewed required patient and public

involvement in guideline development groups (36, 61) and Kung et al. found in a large sample that only a minority of guidelines (16%) included patients in their panels (62).

With this as a backdrop, The BMJ RapidRecs process and methods have emphasized patient involvement through formulation of the PICO questions, panel composition, conduction of systematic reviews on patients' values and preferences, and prioritizing input from patients in guideline meetings (63). There is no consensus on the optimal method of how to do this, although it is currently being investigated (64).

1.3.3 Strength of recommendations

According to GRADE, recommendations can be classified as either strong or weak (or conditional) in favor for or against an intervention. Guideline developers typically issue strong recommendations for an intervention when they are confident that the benefits clearly outweigh harms and burdens. This implies that most patients are believed to prefer following the recommended course of action, and the recommendation can be adopted as a policy in most situations.

Conversely, weak recommendations are issued when there is a fine balance or equipoise between benefits and harms, when the certainty in evidence is low, or when there is a large variability in patients' values and preferences (25, 26, 30, 31, 55). The corresponding decisions are sometimes called preference-sensitive decisions, highlighting the importance of considering patient preferences in the decision-making process (4).

Exceptions to the GRADE guidance for strong recommendations and certainty of evidence are five seldom-occurring situations in which a strong recommendation is warranted despite low certainty in the evidence (31, 65).

1.4 Shared decision-making (SDM)

1.4.1 Evolution of SDM

The development of welfare societies in the post-second world war period allowed for increased patients' rights. Involvement of patients in medical decision-making, especially from the 1970s, has largely emphasized the principle of autonomy. What the competent, informed patient wants trumps paternalistic decision-making (decisions believed to be in the

patient's best interest, but without their consent) in most cases (4, 8, 66). These principles underpin an overarching aim of SDM: to respect patient autonomy and promote patient engagement (3, 67), to avoid the causation of harm, to balance benefits and harms, and to distribute benefits, risks and costs fairly (12, 68).

Internationally, SDM has increasingly been seen as an ethical imperative and a hallmark of good clinical practice (8, 67, 69). This development has also affected laws and regulations. In Norway, it led to the strengthening of patient rights through the Patient Rights Act of 1999, amended in 2012 to the Patient and User Rights Act. Elements from the law from 1999 included the requirement for informed consent and the right to have an individual care plan.

Furthermore, the law clearly describes that the patient has the right to participate in medical decisions concerning available and justifiable options, and that information about those options should be given to patients in an understandable and individualized way (70). Although SDM is highly advocated for through ethical imperatives, international and national policies, and legislations, in health insurance programs, and in research, SDM is far from being routine at the point of care (8, 12, 67, 71).

1.4.2 Definitions and models for SDM

The "sharing of decision making" was coined in 1972 by Veatch in a paper exploring which physician-patient roles that foster the most ethical relationships (72, 73), and the term SDM was first used by Katz in year 1984 (74). Despite many attempts, SDM is not consistently defined (75). We have elected to use a commonly applied definition: "The process where patients and clinicians make decisions together, as partners, using best available evidence, discussing the options, their benefits and harms, and together considering the patient's values, preferences, and circumstances" (7-9). This definition has the merit of highlighting the conversational aspect of SDM.

Different models of SDM have been developed, several of which share similarities (8, 66, 73, 75-79). These models aim to support deliberations between a patient and a clinician, and rest on an understanding that decisions should be influenced by exploring and respecting what matters most to patients as individuals, through meaningful dialogue (8, 77). A systematic review from 2019 analyzed 40 different models for SDM and found the following components to be the most prominent elements: describe treatment options; make the

decision; include patient preferences; tailor information; deliberate; create choice awareness and learn about the patient (76). These elements are also identified by others to be key ingredients of SDM (15). In addition, linguistic evaluations of SDM definitions have shown that they convey a process characterized by a clinician who speaks while the patient mostly listens, and the patient is invited to contribute only if the clinician so chooses to do so, contradicting the notions of SDM (75).

A very practical model is the 3-step model for SDM in clinical practice, developed and later updated by Elwyn et al. (8, 66). The first step is the *team talk* which places "emphasis on the need to provide support to patients when they are made aware of choices, and to elicit their goals as a means of guiding decision-making processes" (66). It is during this team talk that clinicians can emphasize "choice awareness" (80). *Option talk*, the second step, refers to discussing options and comparing alternatives. The third step, *decision talk* refers to supporting patients to explore preferences and arrive at a decision that reflects their values and preferences, guided by the expertise and experience of health professionals (66).

Another similar model for SDM is the "SHARE Approach", a five-step process for SDM, developed by the US Agency for Healthcare Research and Quality (AHRQ). It includes exploring and comparing benefits, harms, and risks of each option (77).

A different model is the *interprofessional SDM model* designed to broaden the perspective of SDM beyond the patient-physician dyad and addresses different levels of SDM within a healthcare system (81).

SDM can also be viewed as a versatile decision-making approach that encompasses various forms aimed at addressing patients' problems in collaboration with clinicians. The framework of *purposeful SDM* provides different forms of SDM for each kind of problematic situation: weighing alternatives and selecting options, negotiating conflicts, solving problems, and developing existential insights (78, 79).

1.4.3 Barriers and facilitators for implementation of SDM

The slow uptake and implementation of SDM in clinical practice are driven by multiple barriers across individual, organizational, and societal levels (22, 69, 71, 82-88). At the societal level, cultural influences on what illness represents, constructions of risk, decision-

making processes, and the involvement of stakeholders in treatment decisions play a major role (86, 87). Organizational characteristics such as leadership, culture, teamwork, resources, priorities, and workflows also influence the implementation of SDM. Additionally, at a system level, the culture of healthcare delivery, education, incentives, policies, and guidelines have been shown to influence implementation of SDM (85). Barriers specific to hospital settings, such as noisy and busy environments and lack of private settings, are also identified to be barriers for SDM (84).

Clinicians may lack the necessary skills in clinical communication required for SDM (7, 84, 89). disagree with the concept of SDM, or disagree with the idea of asking patients about their preferred role in decision-making, believe that SDM is not possible due to time constraints (69, 82, 84, 90), or believe SDM is already implemented (71, 84). They may also find SDM not applicable due to patient characteristics (e.g., reduced cognitive function) or other clinical factors (e.g., acutely ill patients) (71, 82-84).

Patients-related factors, such as the lack of knowledge of risks, insufficient informational capacity, and a belief that they should not disagree with their clinician, have been found to be barriers to SDM (84). Patients may also experience *decisional conflict*; defined as "personal uncertainty about which course of action to take when choice among competing options involves risk, regret, or challenge to personal life values" (91). The use of decision aids has been shown to improve the quality of decisions and reduce decisional conflict (92, 93).

Known facilitators for the implementation of SDM include the motivation of clinicians, the perception that SDM will have positive impact on the clinical process or patient outcomes, and improved communication skills and training (82). Facilitators for patients in SDM implementation include having a positive and trusting relationship with their clinician, involvement of informal caregivers, previous positive experiences with SDM, and having sufficient informational capacity (82-84).

1.4.4 The connection between SDM and EBM

The lack of appreciation for the connection between EBM and SDM is generally not explicitly stated as a barrier for SDM (7). SDM and EBM are often considered separate domains, as evident from how they are taught; SDM is typically included in communication courses, while EBM is taught within epidemiology programs (94). From perspective of EBM,

SDM is described as a “mechanism by which evidence can be explicitly brought into the consultation and discussed with the patient” (7).

However, this translation of evidence has often been overlooked (69), possibly because the first publication of the EBM paradigm in 1992 did not explicitly address SDM and the importance of patients' values and preferences (10). Conversely, from the perspective of SDM, any tools that support it should be firmly grounded in the best synthesis of the current body of evidence. Recognizing this inherent link between EBM and SDM, there have been calls for the joint production of evidence summaries and tools that support SDM (6, 59), which constitutes a central rationale behind the SHARE-IT project.

1.5 Tools for SDM

SDM can be facilitated and complemented using tools to support patients to reach a decision that is congruent with their values and preferences. Such decision support tools are typically called decision aids and come in different formats (e.g., web-based tools, pamphlets, apps, or videos (14). They differ from other health educational materials by making the decision being considered explicit and providing a personalized focus on options and outcomes (benefits, harm, and uncertainties) to prepare for and support decision-making (14, 93, 95).

Different types of decision aids exist. Some can be used independently from the clinical encounter and aim to assist patients in making their own decisions by providing relevant information, improving knowledge, and encouraging engagement in decision-making. These are often referred to as *patient decision aids*. In contrast, *encounter decision aids* are purposefully designed to facilitate collaborative conversations at the point of care, supporting clinicians and patients to make decisions together (74, 96). While most decision aids are designed for specific conditions, treatments, or tests, there are also generic decision aid formats available to support the process of SDM (97).

1.5.1 IPDAS and the development of decision aids

As for systematic reviews and guidelines, there is a need to ensure that decision aids are of high quality. In this regard, the International Patient Decision Aid Standards (IPDAS) Collaboration has developed a set of standards and principles for decision aid content,

development, implementation, and evaluation (95, 98-101). These standards play a central role in informing the development of the SHARE-IT decision aids.

In 2005, IPDAS Collaboration established 12 quality domains with a checklist of 74 criteria to be used to enhance the content, development, implementation, and evaluation of decision aids based on evidence synthesis (95, 98). The 12 domains were: 1) systematic development process; 2) providing information about options; 3) presenting probabilities; 4) clarifying and expressing values; 5) using patient stories; 6) guiding/coaching; 7) disclosing conflicts of interest; 8) providing internet access; 9) balanced presentation of options; 10) using plain language; 11) basing information on up to date evidence, and 12) establishing effectiveness (95).

In 2012, a major update of the standards was performed to further improve decision aids, also adding explicit guidance on how to summarize and present research evidence on benefits and harms. The last evidence update of IPDAS was published in 2021 with a specific focus on identifying recommendations for criteria changes. A broader consensus process will be issued before any of the IPDAS criteria are changed (100).

In the 2012 update, IPDAS explicitly refer to GRADE as an opportunity to enhance decision aids to reflect best current research evidence (18, 102). The standards also identified key features of the development process of decision aids, beginning with scoping and design, followed by iterative testing with patients and health professionals, testing in real-world settings before producing the final version for dissemination or further research (101, 103).

A review from 2021 of the design and development processes of 283 decision aids concluded that decision aid developers increasingly have embraced principles of user-centered design, but often underreport clinician and patient involvement in the development process (101). In addition, another review showed that a third of examined decision aids had no scientific references, and most did not report on evidence appraisal or on the quality of evidence. Similarly, an evaluation from 2018 showed that less than half of organizations developing decision aids had documented their approaches for evidence summarization (104).

IPDAS recommends that information about options should be presented in a neutral, unbiased, and complete way, both in terms of content, format, and display for users to "process this information without bias" (105).

Key attributes in IPDAS on providing a balanced presentation of options are: 1) to present all reasonable options; 2) to include all information necessary to make a choice; 3) to make it easy to compare options without emphasizing one option; and 4) to ensure that the information is trustworthy and based on scientific evidence.

A "side-by-side" display of information is regarded to be most balanced way to present information (105). A balanced presentation is important to avoid framing effects that can influence the decision-making process, and central to this is how probabilities and risks are presented and communicated (106).

An expert consensus group have identified key issues in risk communication informing IPDAS, based on a broad review of evidence on related to how to best convey numeric information. Overarching principles include the use of numerical risk formats, presentation of options and outcomes in the same risk formats and testing of formats in the population to whom the risk applies. Other central issues are how to present the chance that an event will occur and its uncertainty and relation to time, the visual formats such as use of symbols, labels and colors, consideration of context and skills of both patients and clinicians when presenting estimates and the use of risk calculators (107-109).

The standards recommend that decision aids include a process to support patients to clarify their values. Listing the pros and cons of a decision was found to be the most common method used when summarizing evidence, although there is no agreement on a single best practice for how this is done (110). The 2021 update of the underlying evidence of IPDAS is clear that patient decision aids should include an explicit value clarification method, as this decreases decisional conflict and increases value congruence (111).

1.5.2 Effects of decision aids

The most recent Cochrane systematic review of the effects of decision aids in people facing treatment or screening decisions identified 105 studies involving 31.043 participants (14). Decision aids have been applied across many clinical topics including prevention and

medication decisions (112-117), different surgical procedures (118, 119), and screening decisions (120, 121).

Decision aids have been demonstrated to make patients feel more knowledgeable, better informed, and clearer about their values. Furthermore, patients using decision aids probably have a more active role in decision-making and more accurate risk-perceptions. Low-quality evidence show that they may achieve decisions that are consistent with their values.

Notably, these are pooled results that include both decision aids used to prepare for a consultation and decision aids used during a clinical encounter. Subgroup analyses with comparison of patient decision aids with encounter decision aids found similar improvements regarding knowledge and accurate risk perception (14), later to be confirmed a systematic review of the impact of encounter decision aids (122).

Contrary to popular beliefs, there is currently insufficient robust evidence to support the claim that decision aids improve medical adherence or health outcomes. This has been investigated in various clinical scenarios (112-117).

Limited focus has been on clinicians' experience using decision aids (90). Both patients and clinicians often have inaccurate expectations about the benefits and harms of interventions, with patients overestimating benefits and underestimating harms, while clinicians may have inaccuracies in both directions (123, 124). Therefore, decision aids can play a crucial role in providing a more accurate understanding of the possible impact of a decision, benefiting both clinicians and patients. Additionally, interventions that target both patients and clinicians, rather than focusing on one group alone, are likely be more effective in improving the adoption of SDM (125, 126).

1.6 Burden of treatment and practical issues related to decision-making

Both IPDAS and GRADE highlights the importance of assessing all factors important for decision-making, in the context of developing guidelines or decision aids (24, 95, 103). Nevertheless, most focus has been on evidence synthesis and presentation of typical benefits and harms, rather than including burden of treatment and other relevant issues also important for decision-making.

Burden of treatment can be defined as the "workload of health care and its impact on patient functioning and well-being" (127, 128), although various definitions exists (129). It takes into account everything patients do to take care of their health (130), and when excessive, constitute an onerous burden (131-133), often invisible to physicians and other health professionals (129, 134).

Burden of treatment has been described as a multidimensional concept, consisting of both subjective and objective elements (135). To assess burden of treatment alongside traditional medical outcomes, Tarlov et al. published in 1989 a framework that assessed a broad array of patient-reported outcomes (e.g., physical, mental, social/ role functioning, structure, and process of care). This allowed physicians to inform patients about trade-offs with different treatment options (136). Patients with chronic conditions found the framework useful when reflecting on importance of outcomes (137).

Eton et al. have later developed a more specific framework of burden of treatment, categorizing burden of treatment into different themes (work patients must do to care for their health, challenges/stressors that exacerbate felt burden and impacts of burden). This can be used for conversations at the point of care as well as a self-report measure for patients to study and analyze burden of treatment (128, 138).

In addition, Tran et al. have developed a rich taxonomy of burden of treatment separating healthcare tasks, factors that exacerbate the burden of treatment, and consequences of healthcare tasks imposed on patients in their daily lives. Contrary to the framework developed by Eton (128, 138), this taxonomy also includes the *consequences* of burden of treatment (139).

In the work presented in this thesis, we have chosen to use the broader term "patient-important practical issues", henceforth *practical issues*, to describe issues patients face when implementing interventions and their subsequent impact on their daily lives. Examples of such issues can include coordination of care, required tests or office visits, lifestyle changes, as well as effects and restrictions on social activities, diet, work, or travel. These issues can come as additions to the burden of illness and are for many patients unstainable due to

imbalance between high health care workloads and demands and their individual capacity (133, 140).

The lack of consideration of practical issues, patient context (83, 141, 142), and burden of treatment (130, 132, 135, 139, 143) can result in suboptimal clinical decisions. The results of these suboptimal decisions can be accentuated by the inaccurate expectations of benefits and harms of interventions both by patients and clinicians (123, 124). This led us to explore systematic inclusion of practical issues in decision aids and guidelines.

1.7 MAGICapp in the evidence ecosystem

1.7.1 MAGICapp

With the aim of addressing key issues with guidelines - based on globally accepted standards, (5, 48-51) methods (GRADE) (26, 36, 45, 55) and making use of digitally structured data – the MAGIC founders realized in 2010 that no platform was available to facilitate the authoring, publication and updating of guidelines and evidence summaries. This insight led to the development of MAGICapp, which was subsequently enhanced with the integration of decision aids produced from guidelines as key components, as reported in the SHARE-IT project and this thesis.

All MAGICapp development includes combined research, innovation, and technology development, performed in parallel processes. Of particular importance for SHARE-IT was the DECIDE project (see paragraph 3.2.1) as well as the "Smooth National Adaptation and Presentation of guidelines to Improve Therapy" (SNAP-IT) project that resulted in multilayered guideline formats directly informing the development of encounter decision aids (47, 144).

Figure 3 illustrates how MAGICapp works as an authoring and publication platform for evidence summaries, guidelines, and encounter decision aids. The platform allows authors to write, publish and dynamically update content in a highly structured digital fashion (2). MAGICapp allows automated publication of the output in multilayered formats on all devices. By accessing pertinent information (e.g., recommendations) in the top layer, users can easily explore deeper layers with more detailed information to fully inform their decision-

making (47, 144). Digitally structured data in such formats allow more efficient dissemination within an enhanced evidence ecosystem, as described below.

Over the past decade, the use of MAGICapp has expanded while the platform has been continuously improved. In 2023, more than 48 000 authors have an account to create content in MAGICapp and more than 60 organizations worldwide have a license to publish guidelines. All content published in MAGICapp is freely accessible to end-users such as clinicians and patients.

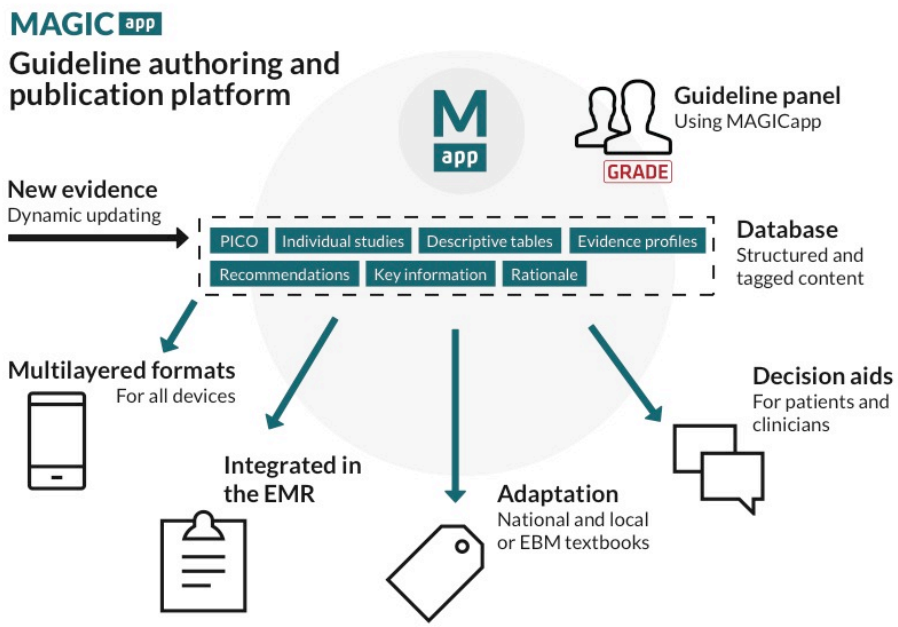


Figure 3. MAGICapp features on the authoring and publication side

1.7.2 A digital and trustworthy evidence ecosystem

An early insight in MAGIC was that any efforts to improve guidelines, evidence summaries, and encounter decision aids are futile if they do not make an impact on the actual care. This resulted in the idea and development of a digital and trustworthy evidence ecosystem, as a conceptual framework to connect tools for decision support to upstream evidence production and downstream implementation into policy and practice. Problems with siloed actors (figure 4a) resulted in some proposed overarching solutions (figure 4b), driven by core requirements.

Through actors working together, explicitly agreeing on standards, methods, processes and use of platforms, digitally structured data can ideally flow from its production to documented

improvements in delivered care. This framework is one example of the various models for evidence ecosystems and learning health systems, all aiming to address similar challenges but with slightly different approaches (145-147).

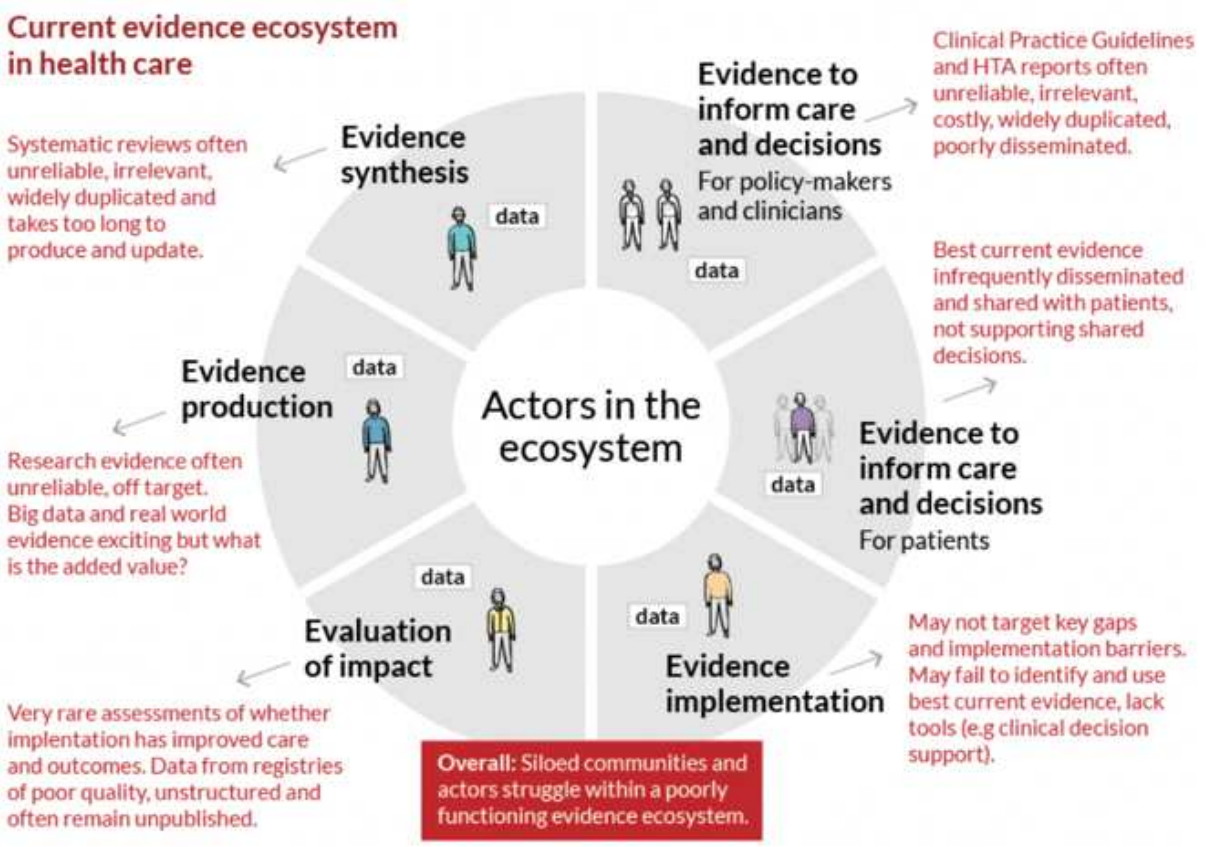


Figure 4a: Problems in a broken evidence ecosystem with siloed actors (HTA - Health Technology Assessment)

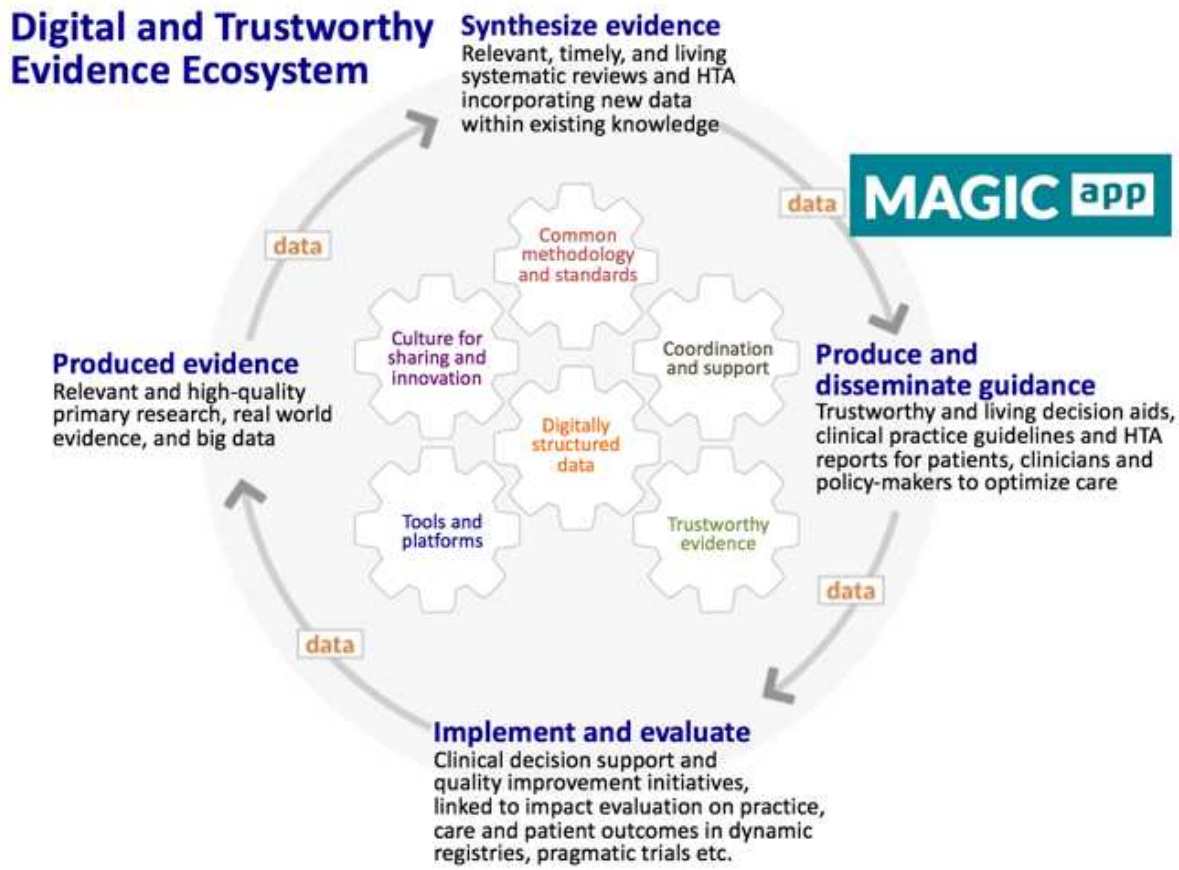


Figure 4b: Digital and trustworthy evidence ecosystem with MAGICapp as core platform (HTA - Health Technology Assessment)

One key driver for enhancing the evidence ecosystem is the use of digitally structured data. With the individual guideline recommendations, evidence summaries, and encounter decision aids in MAGICapp representing the entities (as opposed to PDF formats or unstructured text), digitally structured data can be exported and imported between platforms. This is made possible through Application Programming Interfaces (APIs) and other technological solutions (e.g., widgets and links). Such export and import make processes more efficient and data easier available in multiple end-products.

Digitally structured data allows MAGICapp to connect to other key platforms in the evidence ecosystem, to let data flow more efficiently. For example, data in platforms for evidence synthesis (e.g., Revman and Covidence) can go directly into SoF-tables in MAGICapp and digital outputs from MAGICapp (e.g., recommendations, evidence summaries and encounter decision aids) can be integrated in the electronic health record as decision support systems to further ease implementation efforts.

The ideas, innovations, and the development of the evidence ecosystem have progressed significantly in recent years and have accelerated with the COVID-19 pandemic through the new concept of "living evidence" (148, 149). Here, systematic reviews and guidelines can be dynamically updated and incorporate new evidence more quickly as it becomes available.

For example, living guidelines published through MAGICapp (148, 150) have demonstrated that enhanced dissemination is possible through websites, pathways, journal publications, and local protocols. However, living guidelines are not the norm, and only available for a limited set of recommendations with several challenges and method questions that are not fully answered (151, 152).

2. Aims and objectives

The aim of the SHARE IT project was to develop, test, and evaluate alternative presentation formats for generic encounter decision aids linked to recommendations in guidelines, based on GRADE methods and digitally structured data, in an online authoring and publication platform.

This includes the following objectives:

1. Develop a framework for the generic translation of GRADE evidence summaries and guidelines into encounter decision aids (article 1 and 2)
2. Design and user-test a set of interactive, adaptable, presentation formats of encounter decision aids and integrate these in MAGICapp (article 1 and 2)
3. Develop a framework of patient-important practical issues on management options, integrate this in MAGICapp and explore the feasibility of including of such issues in real-life production and publication of guidelines and decision aids (article 3)
4. Conduct a systematic review on patients' values and preferences to inform a guideline panel creating recommendations (article 4)

3. Materials and methods

3.1 General overview of methods and related research projects

The SHARE-IT project applies principles from human-centered design and utilizes a combination of quantitative and qualitative research methods. The studies are partially conducted within several related research projects and initiatives, described in more detail in the following paragraphs. The material and methods are presented separately when the research projects differ in design and execution, and they are reported only once when the same methods are used in different studies.

Figure 5 gives an overview of overarching and specific methods and related research projects and initiatives related to each paper.

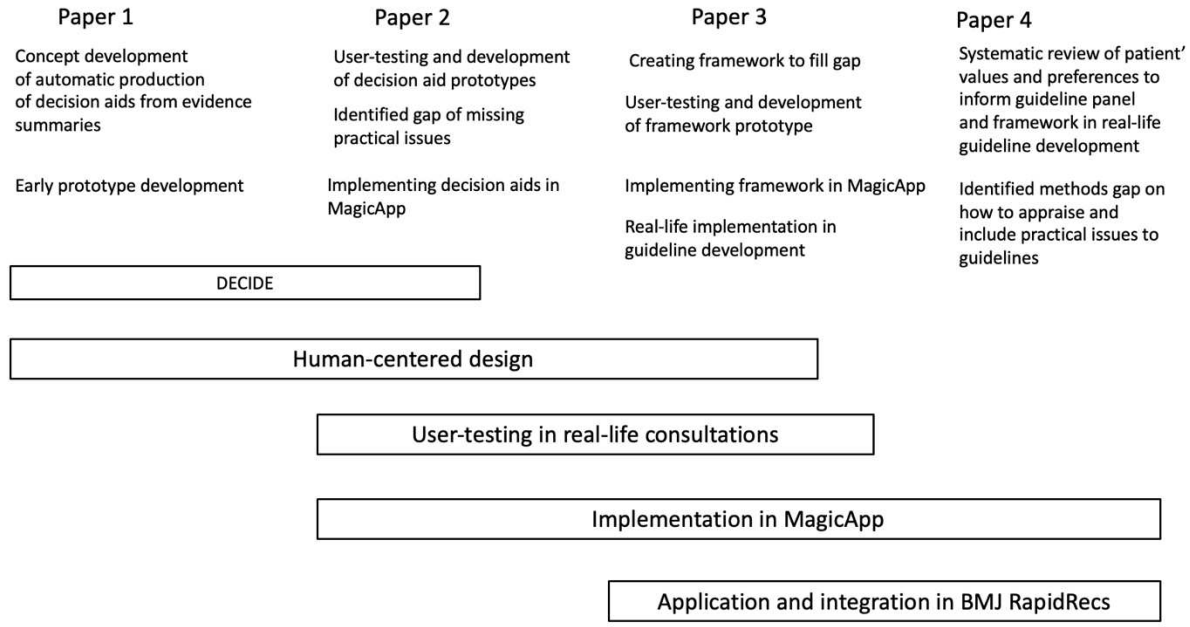


Figure 5. Overview of projects and central principles in papers

3.1.1 Human-centered design and Morville's honeycomb

The user experience is at the core of human-centered design, also known as user-centered design (153), and is central in the research presented in this thesis. User-centeredness is described as a multidimensional concept of four distinct dimensions (154):

1) User focus, which means designing and developing the system around the needs and capabilities of users.
2) Work-centeredness, which involves designing and developing the system based of users' workflow and tasks.
3) User involvement or participation, which entails users in the design and development process allowing them to participate actively.
4) System personalization, which refers to individualizing the system for or by individual users.

The user experience can be influenced by various factors, which can be explored and categorized by using Peter Morville's conceptual honeycomb framework of user experience. This framework consists of different facets: usefulness, usability, desirability, findability, accessibility, credibility, and value. Each facet illuminates different aspects of the experience and the distributed nature of factors influencing a user experience (43, 155).

In our study, we used a revised version of the honeycomb framework of user experience, adapted from the original version (Figure 6) (43).

Figure 6. Modified Morville's model for testing the experience of users (43)

The adapted version was developed based on research that focused on how a design approach could support evidence-informed practice, regarded more specific to our use. This adapted version differs from the original honeycomb by adding *understandability* and

identification/*affiliation* while removing the facet *valuable* from the original version (43). A description of each facet can be found in table 1.

Facet	Description
Accessibility	Are there physical barriers to gaining access, also for people with handicaps, like color blindness, fonts, wi-fi?
Usability	How easy and satisfying is this product to use? The right click, the right screen?
Understandability	Does this person comprehend correctly both *what kind* of product this is, and comprehend the content correctly? Is this person's subjective experience of whether they understand in line with their actual (correct or incorrect) understanding?
Credibility	Is the product/content experienced as trustworthy?
Usefulness	Does this product have practical value for this person? Is it relevant for the decision process?
Desirability	Is the product something this person wants? Has a positive emotional response to?
Identification	Does this person identify with the product, on a personal or a social level? Or is it alienating, experienced as being not designed for "someone like me"?
Findability	Can this person locate the product or the content that they are looking for?

Table 1. Description of Morville's user experience facets

3.2 Related research projects

3.2.1 DECIDE

The Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE) project was a multi-national research endeavor launched by the GRADE working group and co-funded by the European Union (156). The research reported in the first and second paper in this thesis was conducted within DECIDE project. After the project ended in 2014, we continued to conduct user-testing and develop the encounter decision aids.

The objective of the DECIDE project was to explore methods for effective communication of evidence-based recommendations to key stakeholders, including healthcare professionals, policymakers, and managers, and the public. The DECIDE consortium did not place objectives or methodology within a particular knowledge transfer or implementation framework. Instead, the project aimed to provide empirical support for various communication strategies, particularly focusing on how research evidence is presented to optimize access and use of information contained within guidelines (157). The project also addressed key communication features supported by current evidence to improve implementation (34).

3.2.2 BMJ Rapid Recommendations

The BMJ RapidRecs is a collaboration between MAGIC and The BMJ, embedded in the evidence ecosystem, demonstrating how evidence synthesis and guidelines can be efficiently created, published, and globally updated. The aim is to translate emerging research into user-friendly and trustworthy recommendations, evidence summaries, and encounter decision aids in a timely and transparent process, minimizing bias, and adopting a patient-centered perspective (158). This is done by identifying potentially practice-changing evidence, conducting systematic reviews on benefits and harms of the intervention, prognosis, and the values and preferences of patients, using GRADE methods.

Guideline panels in the BMJ RapidRecs include patient partners, including those with lived experience of the topic, front-line clinicians, clinicians with research expertise, and methodological experts in health research methodology and guideline development. Potential conflict of interests, both financially and intellectual, are managed with utmost prudence.

Systematic review authors are included in the panel to the extent necessary to ensure optimal communication regarding evidence assessment and recommendation development. The guideline panel provides trustworthy guideline recommendations, encounter decision aids, infographics displaying pertinent information, and key practical issues associated with the developed recommendations. These are published in MAGICapp and in the BMJ after peer review (63).

The research presented in article four and parts of the research in article three in this thesis was conducted under the BMJ RapidRecs umbrella.

3.3 Decision aid framework and tool development

The first paper in the thesis provides an overview of the challenges associated with current decision aids. It also describes the conceptual and initial development of a framework for translating GRADE evidence summaries and guidelines into encounter decision aids, linked to trustworthy guidelines. Additionally, it presents the design of an early prototype for encounter decision aids with interactive and adaptable presentation formats (159).

The second paper focuses on the iterative development and user-testing of the encounter decision aids, as well as their integration into MAGICapp (160).

For the development of the SHARE-IT encounter decision aids, we applied different standards, frameworks and methods, including IPDAS (99), GRADE methods (25-31, 35-37, 39), the Ottawa Decision Support Framework (92, 93), and the 3-step model for SDM in clinical practice developed by Elwyn et al. (8, 66). The decision aid was developed by the SHARE-IT research team through a series of steps, including development of a framework, template, prototype, and iterations leading to the final decision aid tool (Figure 7). These processes are described in more detail in the following paragraphs.

Figure 7. Development process of SHARE-IT encounter decision aids tools

3.3.1 Development of an initial encounter decision aid template

We developed an initial template for our encounter decision aids, drawing inspiration from decision aid cards developed by pioneer Dr. Montori and his team in the Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, USA. The template allows patients to choose the key outcomes that matter most to them, and in which order they want to discuss them (16, 17).

We framed our encounter decision aid on the 3-step model for SDM developed by Elwyn et al. described in paragraph 1.4.2 (8). Additionally, we incorporated elements from the Ottawa Decision Support Framework, which combines various decision-making theories such as expected utility theory, prospect theory, and conflict theory (92, 93).

We also considered the ecological rationality theory, which recognizes that decisions and heuristics are made within the context of limited time and computational ability (161). This theory is particularly relevant when designing tools for SDM linked to evidence summaries, as the number of outcomes, multiple treatment alternatives, estimates, and uncertainty can exceed patients' and clinicians' computational abilities.

We followed a modified "mobile first" approach (162), collaborating closely with an interaction designer who was familiar with GRADE and human-centered design, to sketch and create the initial template. We used an online tool for visualization of data (163) and Blueprint software, allowing us to customize the first template to a tablet screen (e.g., an iPad) (164).

3.3.2 Stakeholder feedback and brainstorming

To move from the developed template to a decision aid prototype and a conceptual framework, we brainstormed with a multidisciplinary team of experts in SDM, experts in GRADE and other methodologists, patients, front-line clinicians, developers, and designers recruited for the DECIDE project. We conducted three face-to-face meetings with stakeholders in DECIDE (Canada 2012, Italy 2013, and Peru 2013), clinicians and experts in SDM, guideline developers, and designers. The experts were presented with the initial template and subsequent decision aid prototype for brainstorming, discussion, and feedback. The group met physically at relevant conferences for brainstorming sessions and had regular virtual meetings in between.

Patients and clinicians who tested the early template provided verbal feedback directly to the research group. Stakeholders contributed feedback, suggestions, and improvements to the developed template, framework, and prototype. Brainstorming sessions included structured discussions on key elements of the tool and framework.

3.4 Development and user-testing of the encounter decision aid prototype

We conducted formal user-testing of the initial prototype in clinical encounters to further refine the encounter decision aids. The findings and feedback were analyzed by the multidisciplinary group, and the prototype was modified accordingly. These steps were iteratively repeated to enhance flexibility, optimize the content and interface, and ensure patient and clinician input throughout the process, thereby improving the tool's acceptability (17, 102).

3.4.1 Materials for user-testing of the encounter decision aid prototype

We created decision aid prototypes for 22 different clinical scenarios, including decisions related to antithrombotic therapy and cancer treatment. We selected recommendations from published guidelines with evidence summaries developed using the GRADE framework to populate the estimates sections in the encounter decision aids (165, 166). We developed ten encounter decision aids on antithrombotic therapy based on a Norwegian guideline under development by the Norwegian Society for Thrombosis and Hemostasis with support from MAGIC (167).

Additionally, we elected a clinical question regarding the duration of tamoxifen treatment to prevent recurrence of breast cancer and created a GRADE evidence summary also based on a high-quality systematic review in collaboration with field experts (168). The selected recommendations were based on comparisons between two available options. All encounter decision aids reflected decisions that were considered particularly sensitive to patients' values and preferences, typically accompanying weak recommendations according to GRADE (25). The encounter decision aids were available in English and Norwegian.

Participants and setting

We performed user-testing of the encounter decision aids in real-life clinical encounters in secondary and tertiary health care facilities in Norway (Innlandet Hospital Trust, Gjøvik, Oslo University Hospital), United Kingdom (Ninewells Hospital, Dundee) and Canada (McMaster University Hospital and Hamilton General hospital, Ontario). We recruited a convenience sample of physicians with a variable level of clinical experience, knowledge of risk communication, and familiarity of the clinical topic covered by the encounter decision aids.

We recruited patients through the participating physicians as part of the either outpatient clinic visits or acute hospital inpatient admissions.

Instructions to participants and data collection

The participating clinicians were instructed on how to use the encounter decision aid by a study team member. The instruction on how to use the tool was brief (5-10 minutes) and standardized to increase the applicability of the tools, and to be able to identify and differentiate between intuitive features and features that required adaptation of refinement in a subsequent design iteration. One of the study members directly observed the clinical encounter, focusing on interactions involving the use of the decision aid, and moments where the patients had the opportunity to ask questions about their management. We audio-recorded the consultations, transcribed them, and when the encounter happened in Norwegian, translated the transcripts to English by a professional translator.

Directly after the consultation, we complemented the direct observations by a think-aloud session with patients and clinicians separately, focusing the actual experience using the tool. In the think-aloud session, we used a developed semi-structured interview guide with questions meant to elicit specific feedback on their experience and on the format and usability of the tool in the encounter.

We also collected suggestions for improvement of the tool and ended the interview by the 20-item COMRADE (combined outcome measure for risk communication and treatment decision-making effectiveness) instrument. This provides a quantitative assessment of *risk communication* and *confidence in the decision* using a 5-point scale.(169, 170) For use in Norwegian, COMRADE was back translated by a bilingual speaker to assure congruency with the English version.

Data analysis

We analyzed the transcripts and interviews using Hsieh and Shannon's conventional and directed content analysis. First, we performed conventional content analysis by reading all transcripts to obtain a sense of all user-testings. Then, we searched for units of meaning and condensing text, independently and in duplicate to derive codes from the data. We then compared and added codes to the results, developing categories and subcategories based on linked and related codes (171). We searched for barriers, problems, and facilitating elements

or characteristics of the tool that influenced the user experience and the process of SDM.

Secondly, we analyzed the transcripts and interviews using directed content analysis by using Morville's honeycomb of user experience as predetermined codes using a detailed coding instruction of the different facets of user experience (155, 171). Each extracted unit of meaning were analyzed and according the categories of findability, usefulness, usability, understandability, credibility, desirability, and accessibility (155). Finally, we also rated each element as regards the quality of the experience using the following categories: positive feedback, neutral experience, suggestions for improvement, minor frustration, and major frustration.

3.5 Development and application of a framework for practical issues

The third paper in this thesis describes the development of a generic framework of patient-important practical issues on management options. It also describes how we integrated this framework in MAGICapp and applied it in real-life production and publication of guidelines and encounter decision aids in the BMJ RapidRecs (172). We performed user-testing of the framework as an integrated part of the user-testing described in paragraph 3.4.

3.5.1 Materials for framework development

We chose a purposeful sample of two data sources covering a large and varied set of health conditions to identify generic categories of practical issues. Both samples applied a rigorous and trustworthy methodology in identifying patient experience and their most frequently asked questions, including practical issues.

The Health talk registry (www.healthtalk.org), from the Health Experiences Research Group and the University of Oxford, UK, is large sample of patient experience collected through thorough methodology using focus groups and standardized interviews on a large set of health conditions (173). The registry data is based on a systematic collection of interviews with patients regarding their condition and experience, analyzed by an experienced and trained researcher and discussed in an advisory panel including patients, health professionals and researchers. The findings are set in the context of the latest clinical evidence or current best practice (173, 174).

Option Grids, which were produced, at the time of analysis, at the Dartmouth Center for Health Care Delivery, USA, and Cardiff University, UK. Option Grids constitute at that time a specific example of decision aids that aimed to include patient experience in the form of frequently asked questions, which were elicited using a standardized methodology (19, 175). Option Grids come in the form of one-page decision aids, focused on patients frequently asked questions. Their development systematically included patient partners (175).

3.5.2 Methods for framework development and integration in guidelines and SDM tools

Grounded theory

Using Strauss and Corbin's grounded theory approach (176) with an iterative study design, we collected, coded, and categorized the data available in our sources as of March 2014. We collected the data in single words, phrases, or paragraphs dependent on context and processed into a Microsoft Excel database. Two researchers conducted iterative coding and comparison in parallel with constant comparative analysis. We added specific codes to all data considered to be practical issues relevant to management of options on either therapeutic or diagnostic alternatives, excluding issues that were solely about experiencing a health condition.

We then compared and grouped codes with similar content together, beginning the process of conceptualizing generic practical issues categories. Using an inductive approach with constant comparative analysis, we compared codes, refining them, and grouping initial categories into broader categories. We performed axial coding to explore and define connections between categories and among categories and data (177).

We continued this iterative inductive approach until we arrived at a final set of related generic categories forming a framework of practical issues. An interactional designer developed presentation formats and pictographs for each category in close cooperation with the study team. These pictographs and presentation formats were improved through feedback from researchers, patients, and clinicians before and through user-testing.

Integration of framework in guidelines and decision aids

We integrated the final practical issues framework to be a part of MAGICapp to allow practical issues to appear together with GRADE SoF tables and in the encounter decision aids (Figure 8).

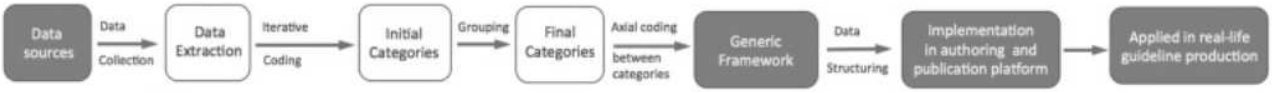


Figure 8. Method for practical issues framework development and integration

3.5.3 Feasibility of using the practical issues framework in guideline development

We used the developed framework in the production of 15 different guidelines within the context of BMJ RapidRecs as a pragmatic approach to assess feasibility of the framework in real-life guideline development. We searched in relevant patient experience databases, online evidence textbooks, tools for SDM, research of patients' values and preferences (178-180) for relevant practical issues and treatment burden for each management option within the guidelines.

We presented these results to the guideline panel for review and refinement. We classified the issues into the relevant categories of the practical issues framework in close collaboration with patient partners and front-line clinicians in the guideline panel. Issues regarded as especially pertinent for decision-making or issues with a large variability in values and preferences were presented to the full guideline panel to ensure all elements relevant for decision-making was included when appraising the body of evidence. Finally, we collected feedback from the guideline panel on the process of including practical issues in the development of the recommendations and interviewed patient partners regarding the usefulness of the framework to identify practical issues.

3.6 Systematic review on patients' values and preferences

The fourth article in this thesis is a systematic review addressing patients' values and preferences, conducted to inform a BMJ RapidRecs guideline update on transcatheter aortic valve insertion (TAVI) and surgical aortic valve replacement (SAVR) for aortic stenosis (44). Results from this review allowed us to explore how patients' values and preferences and a populated practical issues framework could directly inform the guideline panel developing a set of recommendations.

3.6.1 Methods and materials for the systematic review

We applied established methods for high quality evidence synthesis, here with both quantitative and qualitative data to be expected (181, 182). As the review was an update of an

earlier conducted review (178), we followed a checklist for systematic review updates (183). We updated the objective, defined criteria for inclusion and exclusion of studies and performed a systematic literature search using a validated methodological search filter for retrieving studies on patients' views and preferences (184).

We followed the Meta-analyses Of Observational Studies in Epidemiology (MOOSE) checklist for reporting (185). We searched in relevant databases using a combination of keywords and subject headings for "aortic stenosis" and "valve replacement" without language and publication restrictions. We also performed a grey literature search via conference abstracts and searched in the reference lists of eligible studies.

Study selection

We included qualitative and quantitative studies with participants ≥ 18 years with aortic stenosis whose values, preferences, and practical issues related to the decision to undergo TAVI or SAVR were elicited. We excluded studies that transformed quality of life measures into utility values and studies reporting health-related quality of life as these were assessed in the associated systematic review of treatment effectiveness informing the BMJ RapidRecs (186). Studies that did not report original data or were case reports and -series were also excluded.

Data collection

Two authors conducted calibration exercises before individually screening titles and abstracts using prespecified criteria. We, independently and in duplicate, reviewed full-text articles of potentially relevant studies and resolved disagreements by discussion or consultation with a third reviewer. Participant demographics, clinical characteristics, methods used to elicit values and preferences, and findings were extracted. We corresponded with two authors of included studies for further information and contacted authors of two abstracts that were ultimately excluded.

Analysis and synthesis

We first extracted data from the included qualitative and quantitative studies. The extraction form was informed by the patient-important outcomes from the BMJ RapidRecs guideline panel meetings and the developed practical issues framework. We inductively added new outcomes codes to the pre-exciting outcomes and codes. Extracted quantitative results were

qualified into narrative statements to prepare for integrated synthesis (187). We then conducted thematic analysis and synthesis on qualitative study results, following a method developed by Thomas and Harden and used for systematic reviews of people's perspectives and experiences (182).

In addition, direct patient quotes from primary studies were extracted and labelled as direct quotes. We then coded the extracted data to its meaning and content, deriving codes directly from the data. Codes were then compared within and across studies to check for consistency of coding. In parallel, we compared, discussed, and refined these codes. We grouped codes together when relevant, and refined codes to develop descriptive themes across studies. The descriptive themes were combined with the converted quantitative results to formulate the final analytical themes pertaining to patients' values, preferences, and practical considerations concerning the decision between TAVI or SAVR.

Study quality and certainty of evidence

We assessed individual study quality using the Qualitative Research Checklist of the Critical Appraisal Skills Programme (CASP) (188) for qualitative studies and risk of bias for quantitative studies using an instrument developed by Zhang et al.(35). We assessed the overall certainty of the evidence using GRADE for quantitative findings (35, 37), and CERQual for qualitative findings (189).

We rated certainty of evidence as high, moderate, low, or very low for each finding. Findings started at high certainty and rated it down if there were concerns in one or more domains (190). For CERQual, certainty could be rated down for methodological limitations, coherence, adequacy, and relevance (189). For GRADE, certainty could be rated down for risk of bias, inconsistency, indirectness, imprecision, and publication bias (35, 37).

4. Summary of results

The results presented in this thesis began with the development of a framework for the generic production of GRADE evidence summaries and guidelines into encounter decision aids. We used these results to design and user-test a set of digital, interactive, adaptable presentation formats of encounter decision aids and integrated these in MAGICapp.

Results from user-testing of the encounter decision aids with patients and clinicians led us to explore the inclusion of practical issues in encounter decision aids. Based on these results, we developed a framework of such practical issues and integrated this in MAGICapp to complement GRADE evidence summaries and encounter decision aids. We tested the real-life production and publication of encounter decision aids and practical issues framework within the context of guideline development through BMJ RapidRecs. An integrated part of this guideline development was to conduct a systematic review on patients' values and preferences to inform the guideline panel when developing recommendations.

The results are presented in more detail in the following paragraphs, using figures from the BMJ RapidRecs to show how the papers are related.

4.1 Encounter decision aid framework development

We built the initial template of the encounter decision aid using insights from research and development of an interactive and adaptive multilayered presentation formats for guidelines conducted within the DECIDE project and MAGIC (47, 144). The first layer of the initial template (figure 9) had general information about encounter decision aids, risk group selection, and the possibility to select different treatments options related to the relevant recommendation. The second layer consisted of options to select graphical presentation of estimates related to the different outcomes (figure 9).

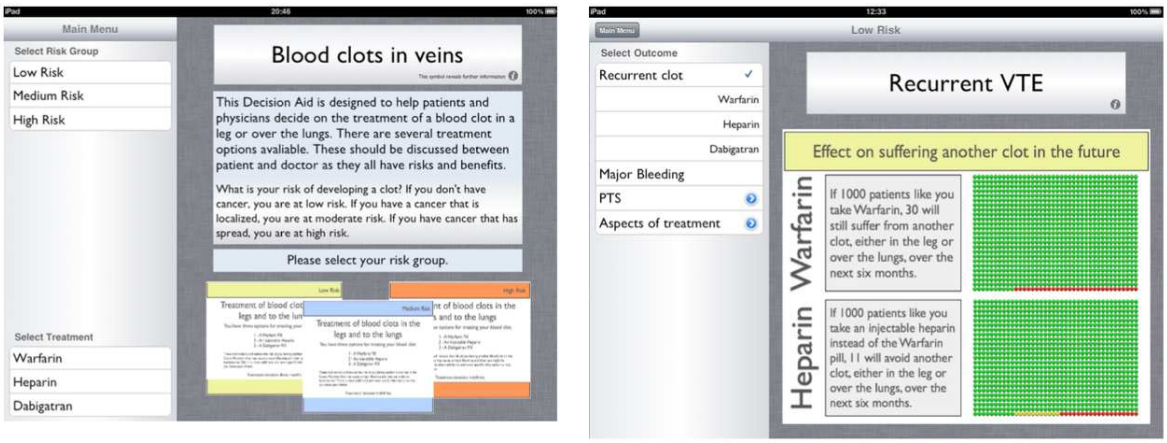


Figure 9. Early developed multilayered decision aid template (left: first layer, right: second layer)

Extensive feedback from stakeholder meetings resulted in iterative development of the tool on how to communicate risk and uncertainty; navigation within the tool; how to facilitate the use of the tool in the clinical encounter and outside of the encounter; and the inclusion of burden of treatment. We developed the template to include a bar graph presentation and a numeric presentation of the estimates in addition to a graphical presentation related to each outcome (Figure 10). We changed the colors of what is considered benefits and harms from the signal colors red and green to more neutral coloring. Certainty of the evidence and impact on daily life and cost were added to the tool.

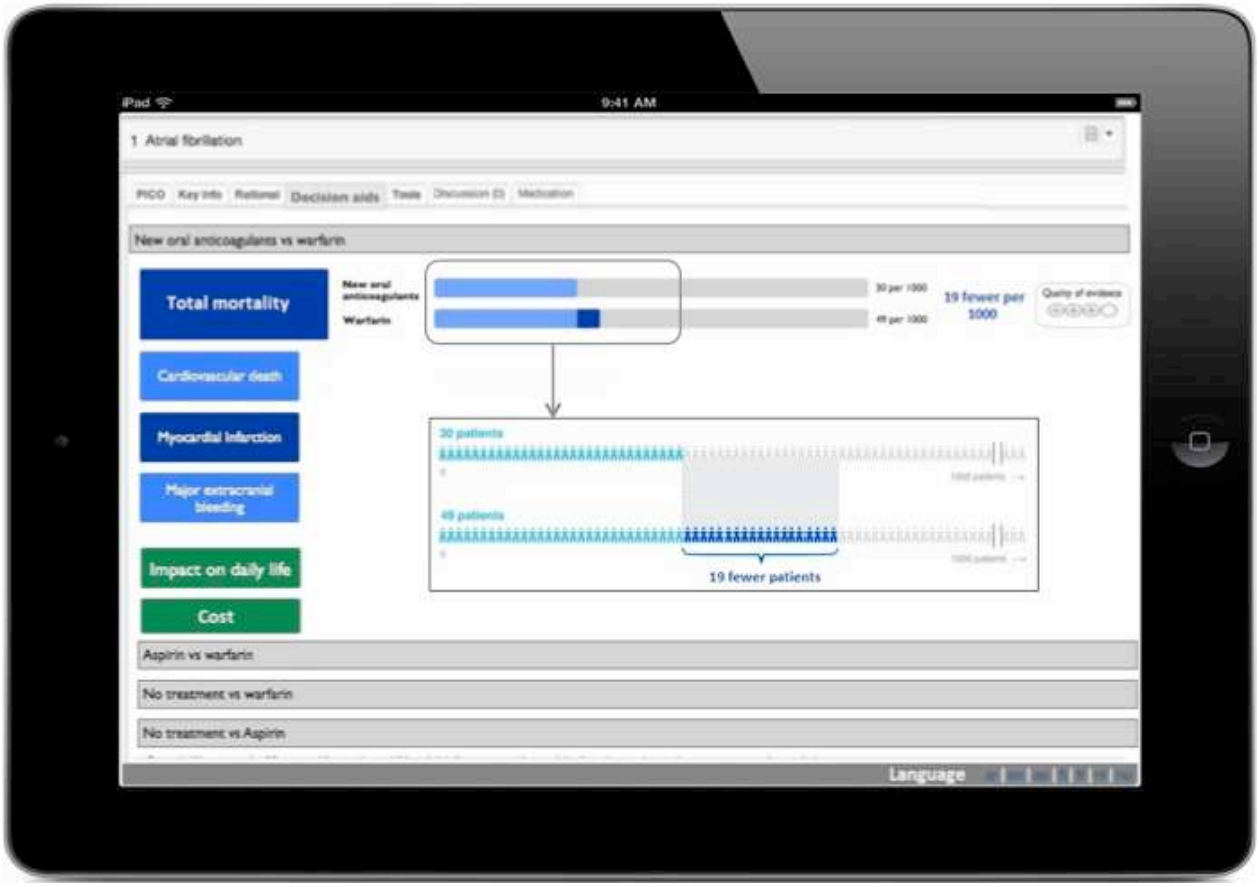


Figure 10. Early decision aid template

4.2 Encounter decision aid user-testing results

We used the developed template as a matrix for development of 74 different encounter decision aids prototypes through four major rounds of iterations. We user-tested the encounter decision aids in 28 real-life consultations with the following topics: anticoagulation treatment for atrial fibrillation, treatment for deep vein thrombosis, pulmonary embolisms, prophylactic anticoagulation, treatment in pregnancy, tamoxifen treatment for breast cancer, and aspirin

treatment as primary prophylaxis for cardiovascular disease. The median age of patients was 53 years (19-90 years) and 64% were women.

Through conventional content analysis, we found a large variability in what was perceived as enough or too little or too much information in the tool. This resulted in the possibility to edit names and description of outcomes, the number of presented outcomes and their order, independent of the underlying recommendation and descriptions made here. Several patients called for a possibility to take notes or copies of the decision aid to take home to discuss with family and caregivers. Based on this, we developed a print-friendly version of the decision aid. Feedback was also given on readability and access, font colors or size or contrast. This resulted in an off-line version and change in design and colors.

Many observations (43%) related to ways to use the tool in consultations, while 32% were expressions of positive feedback, and 12% suggestions for improvement. User-testing did not reveal any showstoppers.

Results from directed content analysis using Morville's facets of usability and rating the quality of the experience showed that most reported issues involved understandability, usefulness, and usability.

Combined results from conventional and directed content analysis showed that most patients and physicians perceived that the tool was useful and provided necessary information, improved value clarification and SDM. The tool was perceived quiet, but some found the first time use awkward.

We identified barriers and issues related to the different facets and used these to develop the tool in subsequent iterations. To improve usability of the tool to follow the natural flow of a conversation and to raise choice awareness, we added and further developed supportive sentences in the different layers of the tool. "What aspect would you like to discuss next? Choose and compare" above the relevant outcomes was added in the second layer and "Among a 1000 patients like you, on average with [*intervention*] in the third layer (Figure 11 and 12). We also developed a short digital, education module to show how the tool is intended to be used (191).

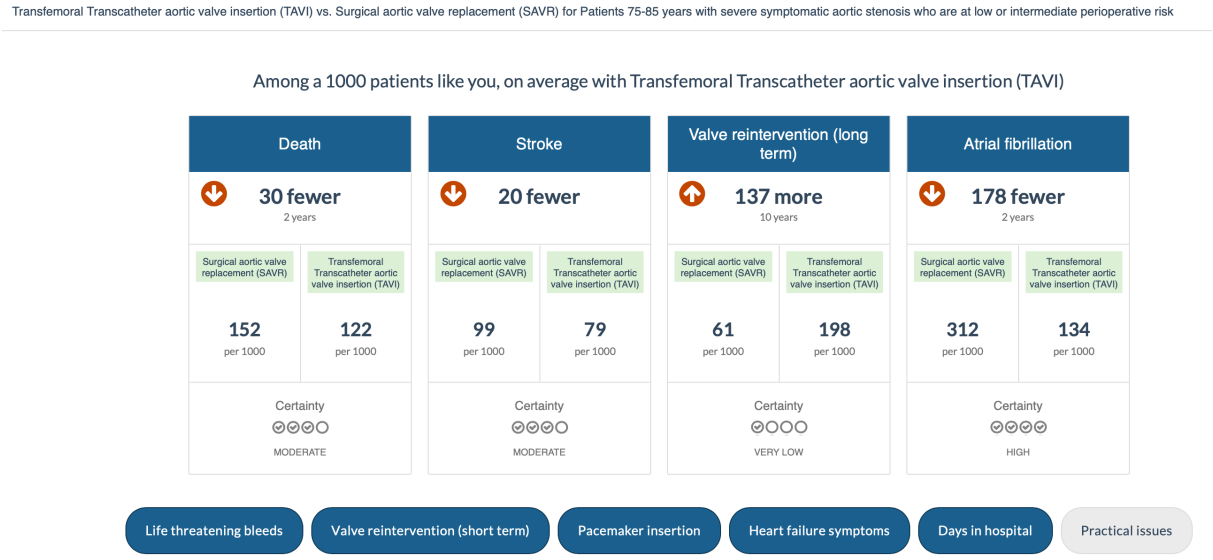


Figure 11. Comparisons of outcomes between SAVR and TAVI in the second layer of the decision aid

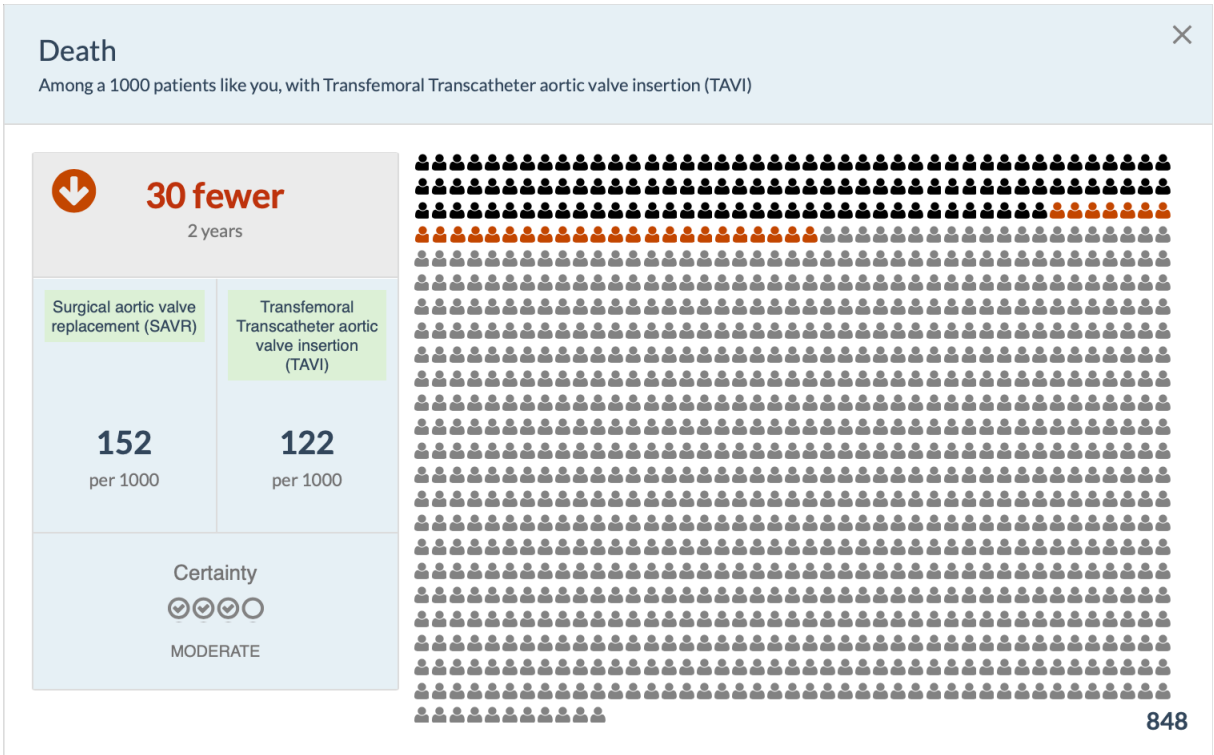


Figure 12. Third layer decision aid comparing long term valve reintervention with SAVR and TAVI

We developed a prototype for presentation of continuous outcomes planned to be tested with a decision aid for medical treatment of lung cancer. This presentation format was not tested partially due to concerns related to quality of the underlying guideline and lack of resources.

We then integrated the prototype in MAGICapp. This resulted in automatically generated encounter decision aids based on estimates from GRADE SoF tables, allowing these to be published and dynamically updated within the context of guidelines. The integration with MAGICapp results in patients and clinicians having access to the underlying evidence as well as guideline recommendations at the point of care (Figure 13). Finally, we created a wizard to allow customization of the encounter decision aids (e.g., change wording of outcomes and interventions) and the possibility to generate encounter decision aids widgets so that decision aids can be showed outside of the MAGICapp platform.

Although not reported in the article and not conducted as formal research, we conducted a screening of all available encounter decision aids in MAGICapp by September 2018 to obtain an overview of the translation of evidence summaries into encounter decision aids. The purpose of this screening was to determine if the translation of evidence summaries resulted in an acceptable output. We categorized the translations as either of acceptable or low quality. Out of the 85 publicly available guidelines, we found a total of 1355 encounter decision aids, out of which 592 had an acceptable translation of the evidence summary. Examples of low-quality translation included encounter decision aids that lacked any estimates or used scales and rankings that were not presented in an understandable format.

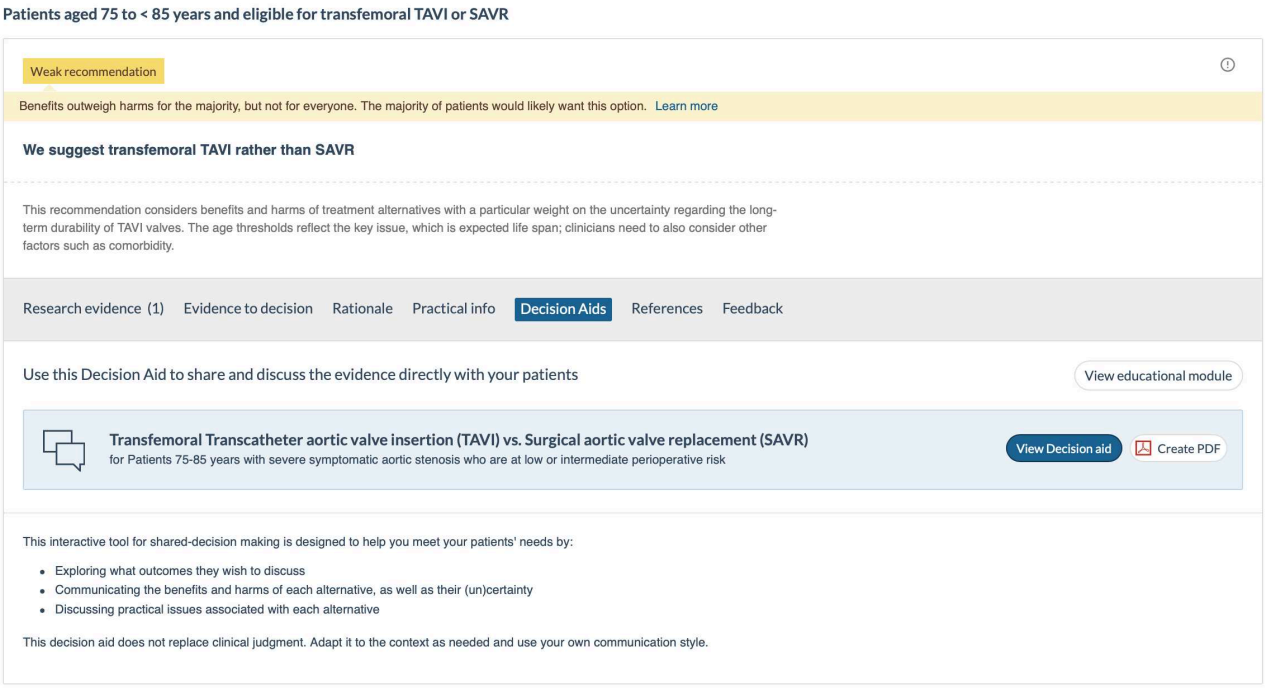


Figure 13. Encounter decision aids integrated in MAGICapp

In the 20 of the 28 consultations (response rate 72.7%), we assessed risk communication and confidence in the decision with COMRADE, with a median of 1 ("strongly agree") in both domains.

4.3 A new framework for practical issues to inform SDM tools and guidelines

User-testing, reported in the previous paragraphs, uncovered that many patients wanted more diverse and comprehensive information regarding how different management options could impact their daily life, going beyond the probabilities of medical outcomes. The early decision aid template provided some information on cost but had limited details on how each management option could specifically impact their daily life.

Patients frequently asked clinicians about these issues and expressed the need for more information on how interventions would affect their daily lives. This included details such as recommended follow-up intervals for visits and tests, as well as potential precautions related to their diet, travel, and activities. However, clinicians lacked the necessary tools to facilitate these discussions. Consequently, we undertook a systematic exploration of this area and developed a framework for practical issues to be integrated into encounter decision aids and subsequently incorporated in evidence summaries.

Development of generic categories of practical issues

We gathered a comprehensive dataset, including videos, transcripts, and text from 29 Option Grids and 297 themes from the Health Talk Registry. This dataset covered a wide range of clinical settings, conditions, treatments, and tests. We added specific codes to all data considered to be practical issues of management options and subsequently combined these codes to create 42 categories.

This process can be exemplified using the extracted data "Missed school days" and "Faced difficulties when making notes or passing exams" first being coded as "Education" and "Managed to continue to work" and "Sick leave typically lasted several months" first coded as "Work", and "Chose and occupation that is compatible with their condition" and "Opportunities for advancement had been limited" first coded as "Career". Through constant comparison and axial coding, we combined the codes "Education", "Work" and "Career" in the final category of "Work and education" (Figure 14). Consequently, the 42 categories were condensed into a final version of 15 generic categories, establishing a framework of practical issues (Table 2).

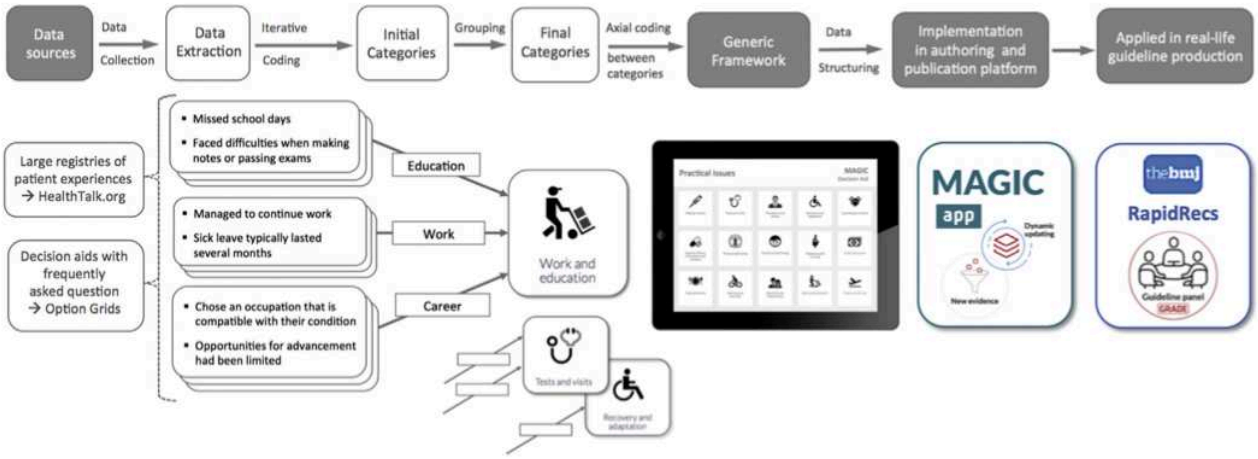


Figure 14. Outline of the development of the generic practical issues framework, integration in MAGICapp, and application in BMJ RapidRecs (172)

We grouped categories with similar content together in the framework, resulting in five practical issues related to care, five practical issues related to daily life, and five practical issues related to miscellaneous issues.

Care-related	medication routine	tests and visits	procedure and device	recovery and adaptation	coordination of care
Daily-life	food and drinks	exercise and activities	social life and relationships	work and education	travel and driving
Miscellaneous	adverse effects and antidote	physical well-being	emotional well-being	pregnancy and nursing	costs and access

Table 2. Developed framework of practical issues

Presentation format of the practical issues framework

Based on insights from earlier and parallel research (47, 144), we organized the framework with associated pictographs in a layered grid, exploring different multilayered presentation formats. The final version of the framework presents the 15 categories in a grid format, with each category name accompanied by a corresponding pictograph to facilitate inclusion in real conversations (Figure 15).

By clicking on the pictographs, users can access the second layer where a narrative description of the practical issues of the intervention and comparator is displayed as a superimposed card Categories that do not have underlying content are visually distinguished by being greyed out. (Figure 16).

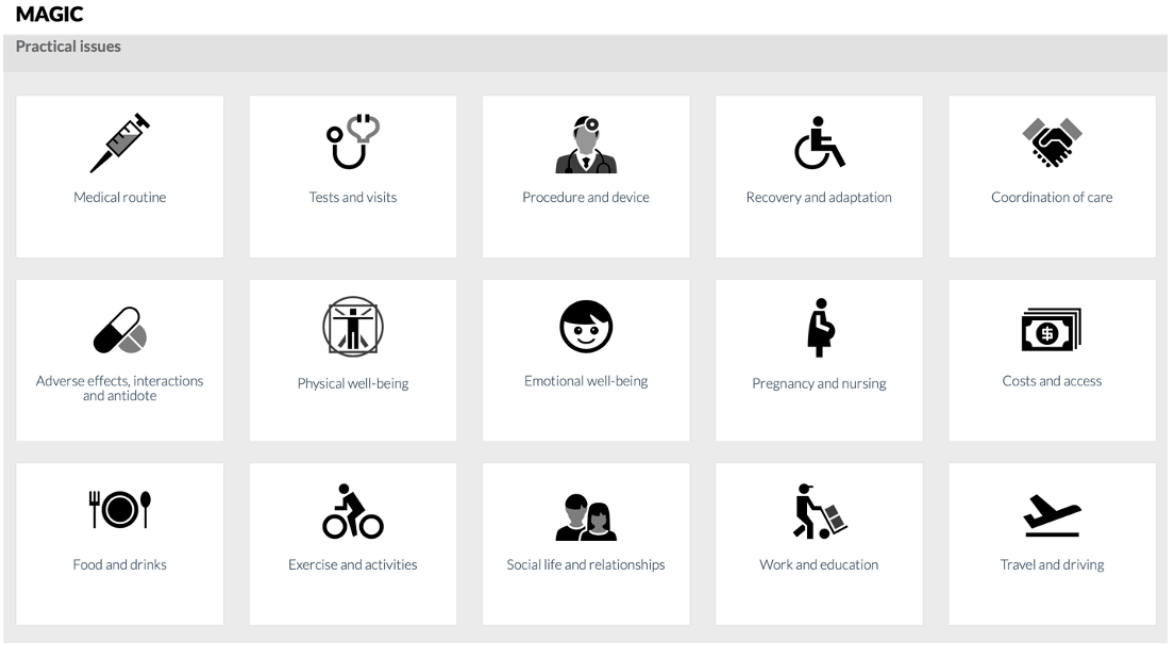


Figure 15. Presentation format of the practical issues framework

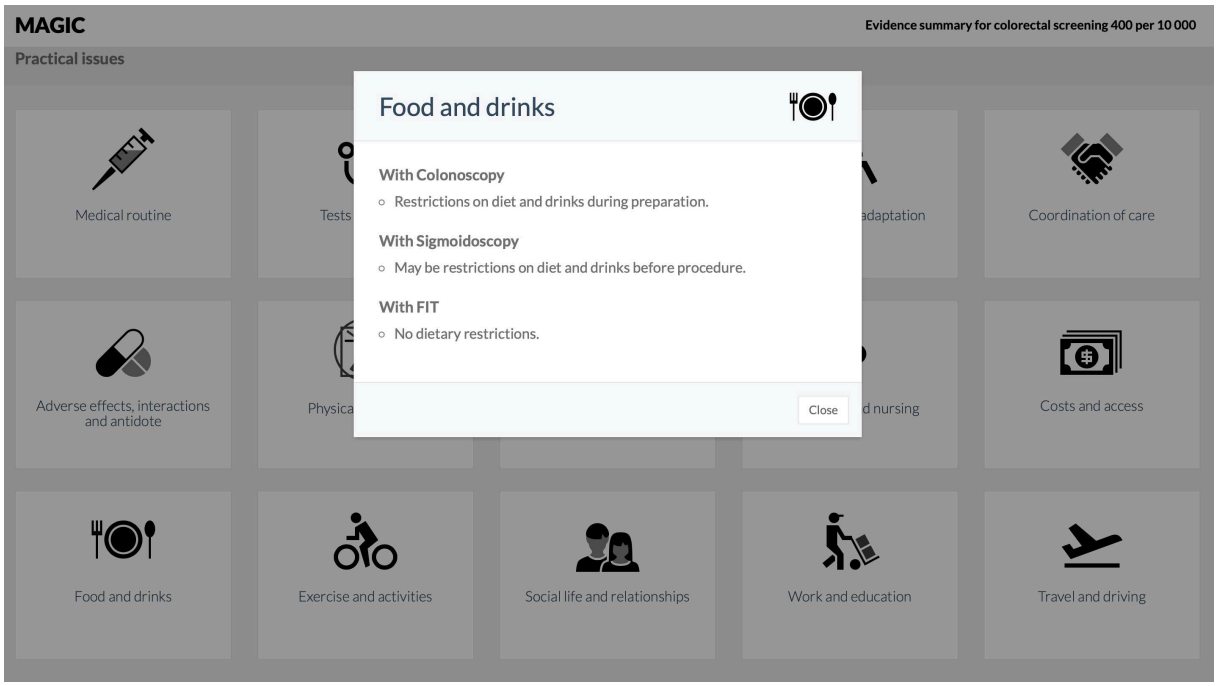


Figure 16. Populated "Food and drinks" category of the practical issues framework

User-testing of the framework

We user-tested the framework as an integrated part of the described user-testing of encounter decision aids in 28 real-life clinical consultations. The results of user-testing showed that most patients chose to explore the practical issues framework and found the navigation between the

numerical outcomes and practical issues to be seamless. The pictographs were described as intuitive, and the framework itself was easy to understand. Several of the patients and clinicians described that the amount of information provided was appropriate and did not overwhelm them.

In some encounters, categories without content sparked conversations between patients and clinicians. Additionally, many patients expressed that using the tool prompted them to consider issues that they would otherwise not have likely discussed during the encounter. Several clinicians were surprised by the patients' interest in practical issues and how these factors influenced the decision-making process.

Feasibility of using the framework in guideline development

We implemented the framework in the encounter decision aids prototype and later in MAGICapp. This allowed us to include and consider practical issues in parallel with traditional patient-important outcomes when creating and assessing evidence summaries in guideline development.

The framework was included in the development of BMJ RapidRecs from 2016 and onwards, covering treatment, test, and screening recommendations. Between 2016 to 2019, a total of 283 different practical issues were added to 35 recommendations in 15 different BMJ RapidRecs. The most frequently used categories were "procedure and device" and "adverse effects, interactions, and antidotes", while the least frequent were "social life and relationships", "coordination of care", and "travel and driving".

Guideline panel meetings and separate meetings with patient partners were conducted to gather feedback and suggestions for improvement on both the framework and the process of incorporating practical issues. A central identified challenge was the search and appraisal of evidence for practical issues. The deliberations between front-line clinicians and patient partners provided rich data that aided the search, appraisal, and determination of the importance of different practical issues.

A central part of BMJ RapidRecs is the corresponding infographics, which provide the gist of the evidence and include key practical issues derived from the populated practical issues framework.

4.4 Patient values and preferences on valve replacement for aortic stenosis: a systematic review.

In our systematic review, we identified eight eligible studies after reviewing 1230 unique titles and abstracts, consisting of two quantitative and six qualitative studies. These studies were conducted in Canada, Norway, Sweden, and USA and included in total 1096 patients. Two studies disclosed receiving funding from a manufacturer of TAVI valves. Certainty assessed using GRADE and GRADE-CERQual ranged from low to very low for all findings, as detailed in the article and the supplementary file (178).

The studies provided limited evidence regarding the explicit balance between benefits and harms with TAVI and SAVR. None of the studies reported patient preferences for choosing between TAVI versus SAVR. Instead, the studies focused on a selection of attributes in isolation without a comprehensive assessment of the beneficial and adverse outcomes associated with the different treatment options.

We found great variability on values and preferences on patient-imported outcomes, as defined by the corresponding BMJ RapidRecs guideline panel. All studies addressed mortality, and the results showed considerable variation in patients' willingness to accept the risk of mortality. Improvements in quality-of-life domains was central in most studies as a reason to undergo treatment. Patients often described improvements in quality-of-life domains as engaging in specific activities, maintaining, or regaining independence, or fulfilling obligations towards family and friends.

Regarding the long-term durability of valves used in TAVI and the need and timing for valve reintervention, uncertainty remains. No study directly addressed how patients valued valve failure or the timing of reintervention. One study examined preferences on valve durability, showing considerable variability in patients' willingness to accept a shorter duration of the effectiveness of TAVI. The certainty of this evidence was assessed as very low.

Several studies addressed invasiveness, length of hospital stay, and the cost of the procedure as concerns and factors influencing decision-making. Additionally, patients considered medical, functional, and social factors when deciding on treatment. Rehabilitation after TAVI

or SAVR, an important theme highlighted by patient partners in the guideline panel, was not addressed in any of the studies.

We used the results of this review to inform the corresponding BMJ RapidRecs guideline panel on TAVI and SAVR for severe, symptomatic aortic stenosis. Furthermore, we combined these results with results from systematic searches for practical issues relevant to the decision between TAVI and SAVR. We discussed these issues with patient partners and front-line clinicians before presenting the practical issues for the BMJ RapidRecs guideline panel for feedback and appraisal.

The practical issues were organized within the developed practical issues framework and integrated into the corresponding encounter decision aids and to the BMJ RapidRecs and complemented the evidence summary in MAGICapp. Key practical issues will be included in the infographic accompanying the BMJ RapidRecs publication, although the release of this guideline update has been postponed due to the pandemic.

4.5 Ethical considerations

4.5.1 Regional ethic committees and institutional review boards

The user-testing of the decision aid prototype underwent review by the Regional Committees for Medical and Health Research Ethics in Norway. Initially, there was uncertainty about whether it was necessary to present the project to their committee. However, they concluded that the project fell within their mandate for assessment. They granted approval for the study to be conducted without any ethical or other objections (Ref. nr.: 2013/1630).

Additionally, the study was also presented to the institutional review boards at Oslo University Hospital and Innlandet Hospital Trust. These review boards found no ethical or other concerns for the conduction of the research.

In the Scottish study, approval was obtained from the National Research Ethics Service Committee South West—Frenchay (15/SW/0127). Similarly, the Canadian study received approval from the Hamilton Integrated Research Ethics board (Ref. 13-373).

To ensure compliance with current data regulations, all audio recordings and patient consent forms were appropriately stored. We have not recorded any patient data such as names, birth dates (except birth year), or other medical conditions that could be used to identify patients.

4.5.2 Ethical considerations concerning user-testing

During user-testing of the prototype in consultations to inform real-life medical decisions, we obtained written consent from participating patients and clinicians after providing them with oral and written information about the study. The user-testing involved encounter decision aids that directly influenced real- life medical decisions. A key ethical consideration was to ensure that the information regarding interventions in the encounter decision aids was both trustworthy and relevant for both patients and their physicians.

In one scenario, we developed a SoF table based on a high-quality systematic review. We used these estimates to populate the decision aid and assessed content for the practical issues framework (168). To maintain clinical validity and trustworthiness of the evidence presented, we conducted this process in close collaboration with topic experts.

For the remaining encounter decision aids, we presented effect estimates, uncertainties, and practical issues from a guideline. This guideline was adapted and published through MAGICapp (47, 165, 192, 193) We assessed this guideline to be trustworthy and representative of the best current evidence available.

5. Discussion

The studies within this thesis represent a journey of iterative research and innovation towards a new generation of tools for SDM to support patient-centered care in clinical encounters. These studies reflect the principle of evidence-based medicine (EBM) – that the evidence alone never fully informs a decision at the point of care but also incorporate patient preferences. By connecting patient-centered care with guidelines, they represent an advance from *care for patients like this* of guideline recommendations to *care for this patient*– the true essence of the clinical task. In this development, SHARE-IT plays a vital role as a technological enabler, ensuring that EBM is effectively practiced when invoked.

I will first discuss the methodological considerations of the SHARE-IT development and design process. I will do this by applying the User-Centered Design 11-measure (UCD-11) (194). Secondly, I will discuss the strengths and limitations of user-testing of the developed encounter decision aids and the practical issues framework, as well as the methods used to analyze these results. Thirdly, I will discuss the strengths and limitations of the real-life implementation of practical issues in BMJ RapidRecs. This will be followed by a discussion of the implication of our innovations and findings, concluding with reflections around future areas of research.

5.1 Methodological considerations

5.1.1 The development process of encounter decision aids

Figure 17. UCD-11 measure (194)

UCD-11, published in 2021, is a specific tool to assess the development and design *process* of personal health tools. It does not aim to measure the quality of the end product or the quality of involvement of users (194) (Figure 17). I will, in the following paragraph, assess the development process of the SHARE-IT encounter decision aids using this measure, focusing on strengths and limitations of our development process.

I will comment on items 1-5 together as they all address end user involvement. I will focus on clinicians, guideline developers, and patients. In addition, I will address items 6-7 together as they relate to iterative cycle development and 8-11 together as they all address health professionals and expert involvement.

Clinician and guideline developer involvement

To understand the needs of end users when using a decision aid, and to improve the tool to fit the clinical encounter, we involved both clinicians and guideline developers early in the development process (item 1) and in the design and development of the prototype, working to

improve the tool to fit the clinical encounter (item 2). Stakeholders, including clinicians and guideline developers, evaluated and gave feedback on the prototypes (item 3). We also asked clinicians about feedback and suggestions to improve the decision aid (item 4) and observed clinicians using the tool together with patients (item 5). We used both observations and feedback to further develop the tool (item 5). We asked clinicians about their opinion of the tool through semi-structured interviews (item 8). We consulted clinicians before developing the prototype (item 9), between iterations (item 10), and we involved a multidisciplinary expert panel in the decision aid development process (item 11).

Overall, the direct and extensive involvement of clinicians throughout the development process represent major strength of our method. Studies have demonstrated that a strong focus on and involvement of end users increase the quality of the developed tool and its ease of implementation (101, 195). This involvement differs from the development of many other decision aids. Typically, decision aids prioritize feedback from patients over clinicians, even when clinicians are users of the tool. In contrast, *patient* decision aids often lack clinician involvement as users (see paragraph 1.5) (196).

We focused less on guideline developers in the development process. In part because we expected the translation of GRADE evidence into encounter decision aids would be covered in other parallel research projects, we underestimated the need to involve guideline makers (47). This underestimation became evident during user-testing when some of the GRADE evidence summaries did not translate to encounter decision aids of acceptable quality. Problems included medical jargon in the evidence summaries, insufficient lay language in the encounter decision aids, and the inclusion of many outcomes less relevant to the decisions aid in the evidence summaries.

This limitation is significant given our aim of providing a generic translation of GRADE evidence summaries and guidelines into encounter decision aids. However, this shows the value of developing encounter decision aids in parallel to guideline recommendations, to ensure that both are using language that is useful at the point of care, where both guidelines and encounter decision aids are hoping to have impact.

Patients

We did not involve patients sufficiently early in the development process (item 1). Although we brainstormed with a multidisciplinary team that included patients (see paragraph 3.32), we did not include patients in the subsequent stakeholder meetings. Secondly, we did not directly involve patients in refining the prototype (item 2). Instead of involving patients early in the development process, we used semi-structured interviews with patients and analysis of user-testing to develop and refine the prototype. During user-testing, patients actively participated in evaluating the encounter decision aids (item 3) through user-testing and semi-structured interviews. We sought their opinion on the tool (item 4) and observed how they used the decision aid together with clinicians (item 5).

This limited involvement of patients early in the development process is a clear limitation of our method and resulted in a delayed discovery of barriers. Examples of problems that probably would have been discovered earlier include the use of medical jargon in the tool, Wi-Fi issues, and the need for scrolling when using the decision aid. The extensive involvement of patients later in the process, similar to clinicians, is a major strength to our approach. Additionally, user-testing in a variety of real-life clinical encounters represents a significant strength that provides insights into how the tools can be implemented in routine clinical settings outside of a controlled research environment (101).

Iterations

The four major iterations the instrument underwent (item 6) represent another strength of the development process. We reported changes between each cycle and the rationale behind each change (item 7). Due to limited design resources, we developed the different parts of the encounter decision aids (e.g., certainty, supportive sentences, card layout) in parallel (item 7). This resulted in less defined iterative cycles and represents a limitation of our approach.

5.1.2 Strengths and limitations of user-testing settings

We conducted all the user-testing in Western countries. User-testing conducted in other settings may have yielded different results. Cultural differences related to SDM, visual presentations such as the use of symbols and colors in the tool, and technical issues such as how other script directions would affect the presentation formats might have contributed to

different results. This represents a limitation of our approach and potentially reduces the external validity of our findings.

Because we user-tested the tool in real-life encounters, we cannot be certain that the patients' satisfaction with the treatment decision confounds the results regarding how the tool was perceived.

We created a semi-structured interview guide (see appendix) to: 1) ensure the rigor and consistency of the interview setting; 2) reduce the risk of confounding; and 3) ensure that the interviews focused on the experience of using the tool. This structured approach to the interview setting represents a strength and supports the external validity of our findings.

The presentation of the COMRADE results represents another limitation. Due to time constraints either for the patient or clinician after the clinical encounter, we did not obtain COMRADE results from six (not eight, as is inadvertently presented in the published paper) of the user-testings. This may have resulted in a selection bias, as these patients may have had a different experience using the decision aid than those answering COMRADE. We have, however, no indications the responder and non-responder groups differ in ways that could indicate systematic error.

In addition, several of the forms had non-differentiation results. The researcher read the questionnaire aloud to the patient. Potential reasons for non-differentiation may be due to time constraints or a desire to quickly answer the survey which could result in a courtesy bias leading to underreporting of negative evaluations of the tool. Due to lack of results from several of the user-testings and non-differentiation, we did not emphasize the COMRADE results.

We included clinicians with variation in age, gender, clinical experience, and knowledge about SDM. This demonstrates that the tool is not restricted to only experienced clinicians or experts in SDM. This diversity increases the external validity of our findings.

Several of the recruited clinicians were colleagues of members of the study team. They were not a part of the research team and had no prior knowledge of the tool before participating in user-testing. We may have unconsciously recruited clinicians more positive to the tool than a

wider and selected sample. However, because we recruited them based on convenience and availability, the risk of selection bias is low. We cannot exclude the potential of acquiescence bias as the collegial bond between some of the interviewed clinicians and the interviewing researcher may have influenced the semi-structured interviews, e.g., clinicians censoring comments and being overly positive about the tool. However, as the majority of the results are derived from direct observations of clinical encounters, we have confidence that the risk of acquiescence bias significantly impacting the results is low.

While developing the SHARE-IT encounter decision aids, we regularly reflected on our roles as both researchers and developers of a commercial service in the marketplace (i.e., MAGICapp) and how this could influence user-testing, analysis of the results, and the prioritization of areas for further development. Although we have analyzed the results critically, our potential intellectual conflict of interest may have resulted in our analysis being overly positive or downplaying negative findings. However, the group of stakeholders functioned as an independent reference group for development prioritization and did not experience disagreements about tool development. We therefore believe the intellectual conflict of interest to be low. None of the researchers had any financial conflict of interests.

Many areas important for how the tool would work in real-life are connected to the structure of a clinical encounter. We therefore regularly discussed the nature and flow of an encounter within the team, especially with our designer and developers as they had less insight in the structure of clinical encounters than clinicians in the team. These discussions represent a strength in our method to create a decision aid that follows the natural flow of a clinical encounter.

5.1.3 Content analysis and Morville's honeycomb for user-experience

We analyzed user-testing data using both inductive and deductive approaches through conventional and directed content analysis. First, I will first discuss the use of Morville's honeycomb and issues specific to this. Second, I will discuss some of the strengths and limitations to the coding and analysis process. Finally, I will discuss why we chose to analyze data using multiple methods.

The adapted version of Morville's honeycomb of user-experience was developed based on research on how design can support evidence-informed practice (43). We used this deductive

method to explore compatibility between the user and the tool and how decision aid design could support conversations at the point of care. Given the interwovenness of the design of the decision aid and how it functions in real-life, the use of Morville as a clear strength to support the development of the tool.

Our team conducted content analysis iteratively, with constant comparison, in duplicate through detailed discussions between researchers. We discussed all data extensively across all user-testings. This extensive process, both in detail and depth, represents a clear strength to our method.

We experienced during coding that, using Morville's framework, one phenomenon could consist of more than one facet. We applied more than one category in these cases, then discussed these results extensively and decided on the category that we felt was most fitting. We therefore regard the risk of selection bias because of this to be low; firstly, because we discussed these in detail, and secondly, we did not design our studies to statistically represent a set of respondents. Therefore, we did not emphasize the frequencies of the Morville categories used.

We chose to use inductive conventional content analysis to explore issues that went beyond the user-experience of the tool. Rather than using software specific to content analysis, we instead used a spreadsheet, which provided us with greater flexibility in the analysis process, thus strengthening our approach. Both researchers who performed the analysis were also developers of the encounter decision aids.

To ensure a critical and transparent analysis of the results and avoid confirmation bias, we regularly discussed this aspect within the research team. Furthermore, we provided examples of the abstraction and interpretation process to enhance the credibility of our findings (197). Although there is a potential risk of intellectual conflict of interest, as described in paragraph 5.1.2, we consider this risk to be low.

Firstly, we used inductive conventional content analysis to explore issues that went beyond the user-experience of the tool. Critics often highlight coding in qualitative research as a method that can drain the data of its variety, richness, and individual character (198). In our case, the hermeneutical approach of inductive coding contributed to preserve the richness of

the data. Nonetheless, it is nearly impossible to avoid losing any context and variety in the process.

Secondly, we used Morville's framework as one analytical lens focused on phenomenological descriptions and issues directly related to user-experience of the decision aid. These complementary methods provided us with 1) an in-depth understanding of the clinical encounter, 2) insight into how to create an encounter decision aids formats that supports the natural flow of the encounter, and 3) uncovering the importance of practical issues. The use of these complementary methods supports the aim of our research.

5.1.4 Grounded theory

To develop the practical issues framework, we used the grounded theory approach. We chose grounded theory due to its suitability for constructing theories through the collection and analysis of the data, its iterative and comparative elements, and its structured and flexible methodology (176, 199).

Considering our limited experience with grounded theory, we adopted the detailed methodology proposed by Corbin and Strauss. This methodology served as a guide for each step of the process, from coding to the development of the conceptual framework (176). We chose Corbin and Strauss' methodology over the of the less detailed and structured "classical" grounded theory originally established by Glaser and Strauss (199).

We deliberated on whether the development of a practical issues framework could be classified as development of a *theory* that predicts a future phenomenon, as is central to classical grounded theory. However, subsequent works on grounded theory state that this method also can be used for development of framework or conceptual schemas. This perspective made it suitable for our purpose (199).

Using grounded theory, we analyzed English-language material from UK and USA. We included only data systematically collected using standardized methodology. We did not include material coming directly from patients, e.g., from web forums or social media. The exclusion of non-systematically collected data may have introduced selection bias and limited our approach. Inclusion of data directly from patients, such as web forums, and in other languages, countries, and cultures would have enriched our data, provided deeper insights,

and may have produced different categories of practical issues. However, due to limited resources, we were unable to include such data.

5.1.5 Strengths and limitations of including practical issues in BMJ RapidRecs

Rather than a formal feasibility study for implementing the practical issues framework in real-life guideline development, we chose a more pragmatic approach. In the following paragraphs, I will discuss strengths and limitations associated with implementing practical issues in BMJ RapidRecs and the appraisal of practical issues content.

All BMJ RapidRecs included practical issues. Due to illness within the research team, we retrospectively explored their inclusion in the guideline development process. Ideally, we would have performed a formal feasibility study, and the absence of such a study represents a limitation to our approach.

The BMJ RapidRecs core team and panel members gradually developed a familiarity with the concept of practical issues. This familiarity had an impact on how and when we integrated practical issues into panel meeting discussions, drawing from our experiences with previous BMJ RapidRecs. We started integrating practical issues earlier in the guideline process, aligning them with the results from corresponding systematic reviews of diagnosis and prognosis, and adopting a more iterative and structured approach.

Although we followed user-centered design principles, resource constraints and limited availability of the guideline panel prevented us from applying all steps of a user-centered methodology. Ideally, the research team would have conducted iterative user-testing when integrating the framework into the real-life production of guidelines. This lack of methodological rigor limits the conclusions we can draw about the implementation process. A more systematic approach would likely have identified barriers and challenges associated with the framework at an earlier stage.

The search for and appraisal of evidence and facts for different categories of practical issues posed methodological challenges. There is a lack of methods on how to include burden of treatment and practical issues in guidelines (200). Traditional search methods in repositories and databases often fail to retrieve relevant results. We encountered difficulties in identifying evidence-based sources to inform certain categories across all BMJ RapidRecs guidelines.

Many of the burdens and practical issues are not reported in published research but can be found in a wide variety of sources, such as patient experience databases, medication guides and online forums, and social media (129, 201-204). Additionally, the validated search filter for patient values and preferences did not encompass practical issues.

Therefore, we sought a more exploratory evidence retrieval without a defined search strategy. In the guideline development process, we supplemented the review findings with results from various sources, including textbooks, informational material, and input from frontline clinicians and patient partners. Due to our extensive process of finding and appraising information related to all relevant practical issues categories, the risk of selection bias is, however, low. Nevertheless, the extensive process needed and the absence of established methods to find and appraise practical issues evidence highlights the need for the development of such methods.

5.2 Implications of key findings

In the in the following paragraphs, I will discuss the implication of our findings and put them in context. I will first discuss the generic aspect of SHARE-IT encounter decision aids and their strengths and limitations. Secondly, I will discuss some aspects related to how encounter decision aids can support SDM. Thirdly, I will discuss studies of patients' values and preferences and their role in informing the development of guidelines. Lastly, I will put the development of a practical issues framework in context to other research and findings, describe how the framework answers unmet needs, and how it has implications for encounter decision aids and guidelines.

5.2.1 Generic encounter decisions aids linked to guidelines: Here to stay?

The idea to link decision aid and guideline development in MAGICapp came from a shared set of challenges in their production, updating, and uptake (5, 18-21, 205). Some decision aid developers have used existing frameworks and templates to make the development process more efficient (97, 206). However, we are the first to semi-automatically translate GRADE evidence summaries and guidelines into encounter decision aids.

The production of SHARE-IT encounter decision aids has proved to work in a variety of conditions and within the context of real-life guideline development. This is demonstrated by more than 250 publicly available guidelines through MAGICapp, containing more than 2000 encounter decision aids. The work presented in this thesis supports the hypothesis that digitally structured GRADE evidence summaries reduce the resources needed to create or update encounter decision aids.

However, some major limitations apply to SHARE-IT encounter decision aids: 1) the need for customization to make them fit for purpose, 2) the observed suboptimal quality of many published encounter decision aids, 3) failure thus far to fully address multiple comparisons of interventions, and 4) the lack of evidence regarding the extent to which these tools are used in practice.

I will in the following paragraphs describe in more detail these limitations and potential solutions.

Customization

Fully customized decision aids often undergo extensive development processes for each specific question and topic (16, 17, 207). We aimed to enhance the efficiency of the development process by generically translating evidence summaries and developing an interactive, adaptable decision aid presentation format. However, user-testing sometimes revealed a need to modify the presentation formats, such as the outputs of GRADE evidence summaries, to make the encounter decision aids better fit the clinical encounter.

Responding to this challenge, we developed several features in MAGICapp that made it possible to customize the encounter decision aids. These features include a wizard that assists authors in 1) capturing and presenting numerical outcomes, 2) allowing the possibility of adjusting the number of outcomes presented in the encounter decision aids independent of outcomes in the evidence summary or the order of those outcomes, and 3) facilitating the use of lay language instead of medical jargon. While these features may not fully accommodate what can be achieved with fully customized encounter decision aids, the need for customization is limited if the underlying GRADE evidence summary is of good quality and written with the end - supporting SDM at the point of care - in mind.

Suboptimal quality of many encounter decision aids published in MAGICapp

In our user-testing, we incorporated trustworthy and high-quality guidelines published in MAGICapp, which differ from many lower quality guidelines published in MAGICapp. However, the suboptimal quality of those other guidelines and the SoF tables associated with those guidelines results in the production of suboptimal encounter decision aids. This poses a significant limitation for the semi-automatically production of high-quality and trustworthy encounter decision aids derived from GRADE evidence summaries.

Various factors contribute to the suboptimal quality of guidelines and SoF tables, resulting in suboptimal encounter decision aids. The correct application of GRADE is not guaranteed using MAGICapp, and many systematic reviews fail to summarize evidence concerning all patient-important outcomes. In particular, inadequate reporting or omission of information often involves harms (208-210). Additionally, if the underlying evidence in MAGICapp is not updated, the corresponding decision aid may rely on outdated evidence.

Furthermore, relevant to all guidelines in MAGICapp, regardless of quality, is the current limitation of MAGICapp in supporting the appraisal of qualitative evidence. This limitation restricts the development of encounter decision aids based on mixed method synthesis.

Failure to include all important outcomes and to update can result in variable quality of guidelines developed through MAGICapp, which in turn affects the quality of encounter decision aids. Consequently, clinical decisions may be based on flawed evidence (5, 21, 95, 211). Although MAGICapp offers the advantage of providing support and guidance throughout the appraisal process, leading to the publication of recommendations, guidelines, or encounter decision aids, it does not guarantee high-quality guidelines.

5.2.2 Multiple comparisons

Research on SHARE-IT encounter decision aids is limited to pairwise comparisons of interventions. However, GRADE has provided guidance on creating SoF-tables from network meta-analysis results across multiple comparisons. GRADE has developed this guidance concurrently with the SHARE-IT project (212, 213). In addition, the SHARE-IT encounter decision aids have been further developed as part of the Making Alternative Treatment Choices Intuitive and Trustworthy (MATCH-IT) project, specifically designed to handle multiple interventions (214). This tool has been published as a part of several BMJ RapidRecs

(215-217) and has undergone user-testing with clinicians and guideline panels, yielding promising results (not published).

The current work did not evaluate the encounter decision aids emerging from use of MATCH-IT. Considering the complexity of the involved evidence, we are planning studies like those described in this thesis to assess how patients and clinicians can utilize this tool collaboratively.

5.2.3 Uptake of SHARE-IT encounter decision aids

Low uptake and impact in practice represents a general limitation of decision aids (22, 82). We do not have knowledge about the uptake of our encounter decision aids. The general low uptake of decision aids and the lack of knowledge regarding uptake of decision aids linked to guidelines warrants further research on implementation, which is beyond the scope of this thesis.

However, some issues are specific to the SHARE-IT encounter decision aids. These encounter decision aids can be difficult to find if you are not familiar with MAGICapp, and this difficulty represents a barrier to their uptake. So-called widgets, an "easy access button" that can be included in webpages or in other programs such as electronic health records, can facilitate their access. Several of the BMJ RapidRecs guidelines use this feature (e.g., (218)).

5.2.5 How can decision aids best support SDM?

Our findings raise some overarching questions about the purpose of tools designed to support SDM. We created the SHARE-IT encounter decision aids to support *conversations* between patients and their clinicians to make decisions. Its aim is to support decision-making that not only relies on evidence but also take into account the patient's context, values, and preferences, in order to avoid overlooking implications for each individual's situation (141) and help patients and clinicians form plans of care that fit (219).

Unlike many decision aids that necessitate patients to engage with algorithms prior to a clinical encounter, the SHARE-IT encounter decision aids promote values clarification during the clinical encounter. Such value clarification methods that present users with the implications of their expressed values may lead to better outcomes, albeit being currently

considered as an exploratory hypothesis (220). The inclusion of the practical issues framework facilitates value clarification and outlines the consequences associated with selecting a specific intervention. Therefore, utilizing the practical issues framework holds promise for enhancing outcomes.

Alternatively, these clarifications are less focused on values or preferences and more on how different approaches will impact the daily lives of patients or their caregivers. Therefore, the aim of these clarifications is not to address values directly, but rather to provide increased clarity regarding the implications of a decision.

Critics have raised concerns about the SHARE-IT encounter decision aids and BMJ RapidRecs infographics, suggesting that they potentially complicate the information processing required for decision-making. Instead of using encounter decision aids alone, it has been proposed that the encounter decision aids and infographics should be used alongside decision support tools that allow for the weighing of pre-selected criteria, resulting in a score that can be discussed at the point of care (221). This raises the question of whether complex medical decisions are better addressed using simple encounter decision aids like the ones developed in the SHARE-IT project or encounter decision aids that incorporate algorithmic decision support systems.

A comparison between an encounter decision aid addressing osteoporosis treatment and a validated decision support system has demonstrated the superiority of encounter decision aids in terms of increasing patient engagement, improving knowledge, and enhancing understanding of risk (222).

Medical decision-making often involves criteria that are not predefined or data that are unavailable within a decision support system. Content in both decision aids and decision support systems still relies on judgment on what should be included in those tools (207). As described in Charles at al.'s formative paper titled "Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango)" (223) the complexity of this decision-making process surpasses the current capability of incorporating all relevant elements into a model (224, 225), which holds true for both decision aids and decision support systems.

This complexity raises the question of how much information should be presented in a decision aid to best support shared decision-making (SDM) while also accommodating individual variations in the desired level of detail necessary for making a decision (16, 17, 226). To address this variation, the SHARE-IT encounter decision aids incorporate multi-layered presentation formats, allowing patients and clinicians to access more detailed information in deeper layers and giving patients the ability to actively select the outcomes they are interested in. Furthermore, guideline and decision aid developers have the flexibility to choose the number and order of outcomes independently of the underlying evidence profile. This design offers the advantage of reducing the risk of information overload while providing the opportunity for patients to access detailed information on each outcome or practical issue if desired.

Finally, the rapid developments in machine learning and artificial intelligence offer potential tools to support the prediction of patient choices (227, 228). These tools can enhance human decision-making by improving decision strategies. However, there are concerns that artificial intelligence may not adequately account for patients' unique circumstances and characteristics and raises a multitude of ethical questions and concerns regarding its role in medical decision making (229, 230). A more comprehensive discussion on this topic is beyond the scope of this thesis.

5.2.6 Incorporating patients' values and preferences in guidelines

In the following paragraphs, I will discuss the connection between the conducted systematic review of patients' values and preferences and its role in informing a guideline.

Differences in preferences between patients and clinicians, including the tradeoffs between outcomes, frequently lead to inaccurate assumptions about patients' values and preferences (231-234). This issue also extends to guideline panels (235, 236). In response, guidelines panels have made efforts to incorporate patients' values and preferences by including patient partners and conducting studies focused on understanding patients' values and preferences. The aim is to develop guidelines that do a better job of addressing outcomes and other issues that are important to patients (36, 237).

Our BMJ RapidRecs TAVI and SAVR guideline update illustrate the limitations of the current situation. We found that the meta-analyses and systematic reviews on diagnosis,

prognosis (not published) or the systematic review of values and preferences for this guideline did not include all central patient-important outcomes and issues for decision-making or the relative importance of outcomes.

We found that improvement in quality-of-life domains played a central role in patients' decision to undergo treatment. These findings are consistent with the study conducted by Col et al., which explored the factors that matter most to patients when selecting treatment for severe aortic stenosis. They found that "goals and features that patients value differ from those reported in clinical trials and vary substantially from one individual to another" (238).

Taken together, these findings highlight the significant variability in patients' values and preferences, the limitations of systematic reviews on patients' values and preferences in informing guideline development and that even the development of trustworthy and high-quality guidelines does not guarantee the inclusion of all patient-important outcomes and issues important for decision-making.

These observations prompt us to question the adequacy of current methods for informing guideline panels about patients' values and preferences. It raises the issue of whether searches should extend beyond research evidence found in citation databases to comprehensively capture all relevant aspects of patients' values and preferences. Furthermore, it emphasizes the importance of including patient partners in guideline panels to ensure the inclusion of all patient-important outcomes.

5.2.7 Including practical issues in encounter decision aids

Through user-testing, we discovered that patients wanted to discuss burdens of treatment and the practical aspects of integrating an intervention into their daily lives. However, even experienced clinicians often lacked insights into many of these practicalities and burdens. They were occasionally surprised by the significance of practical issues for their patients and how important they were in the patient's choice of intervention. This is in line with findings of others that conversations between patients and clinicians often fail to adequately address how interventions impact daily lives and routines (231, 239).

In addition to lack of insight of burden of treatment and practical issues, clinicians also lacked tools to facilitate conversations about these issues. Most encounter decision aids typically

address only one or two contextual factors and burden of treatment, most often daily routine on how to take a medication or treatment, examined in a systematic review (240).

In a response to unmet needs of patients and clinicians wanting to discuss practical issues, but lacking the tools to do so, we developed the practical issues framework, which other researchers have also advocated. May, Montori, and Mair highlighted the importance of tools that describe treatment burden while considering patient values, preferences, comorbidities and social circumstances (127). In addition, it has been suggested that the identification and communication of burden of treatment for individual patients and incorporating patient feedback can make treatment plans less disruptive (241, 242).

The practical issues framework provided clinicians and patients with such a tool that supports deliberations and conversations on a wide range of practical issues and treatment burdens. During user-testing, we observed that the practical issues framework, which presented only categories and pictographs without specific content, served as cues for patient to discuss relevant issues. This finding indicates that the SHARE-IT tool facilitates discussions between patients and clinicians regarding issues that are not predefined in the tool. Furthermore, this suggests that the SHARE-IT tool has the potential to accommodate for different SDM purposes (see paragraph 1.4.2) that vary depending on the problem that the patient is experiencing (78, 79).

Eton et al.'s burden of treatment framework complements the developed framework of practical issues (128). Although this framework was not published when we developed our framework, Eton and al. had previously published a preliminary framework derived using data from a single center (128). While Eton’s burden of treatment framework has a different structure and objective, primarily aiming to inform patient-reported measures, it aligns with our practical issues framework. This provides evidence of the external validity of our findings.

In summary, systematic inclusion of practical issues in encounter decision aids is warranted. It provides a structured framework populated with information often overlooked in the clinical encounter. It supports conversations between patients and clinicians to make decisions that not only make intellectual sense - that is, they are evidence-based - but also makes

practical sense - that is, they are feasible given the patient's capacity to enact them in their daily lives without undue burden - thereby resulting in care that fits.

5.2.8 Practical issues in guidelines: Opening a black box in guidelines?

As discussed in paragraph 5.2.7, there is a strong rationale for incorporating practical issues in encounter decision aids. In the following paragraphs, I will discuss the inclusion of practical issues in guideline development and within the evidence ecosystem.

Dobler et al. argue that explicitly including the burden of treatment in guidelines supports guidelines that better reflect outcomes and issues important to patients and enable patients to make informed decisions about interventions, based on their capacity and values. However, despite its importance, the inclusion of burden of treatment and practical issues is rarely done, and the methods to do so is lacking (200).

In the initial BMJ RapidRecs, we noticed that the concept of burden of treatment and practical issues were unfamiliar to many guideline panelists. Initially, practical issues were only discussed to populate practical issues in the decision aid. However, from an evidence ecosystem perspective, the concept of practical issues gradually became bidirectional in the BMJ RapidRecs.

In subsequent BMJ RapidRecs, the practical issues framework served as a structured tool to facilitate the discussion and incorporation of issues important for decision-making during evidence appraisal. It included patient partners and the rest of the guideline panel and complemented the information provided in the SoF tables. Pertinent practical issues were then included in the infographics. In some guidelines, practical issues and burden of treatment directly influenced the development of guidelines and the strength of recommendations, particularly when the burden of treatment outweighed a small treatment effect.

Two examples of guidelines where practical considerations directly influenced the strength of recommendations are the BMJ RapidRecs on Low Intensity Pulsed Ultrasound (LIPUS) for bone healing(243) and Arthroscopic surgery for degenerative knee arthritis and meniscal tears (244).

To further enhance the understanding of the various aspects of burden of treatment and practical issues within guideline panels, the burden of treatment taxonomy (139) can offer guideline panels a shared language and descriptive framework to systematically address treatment burden during panel discussions with the potential to include treatment burdens more systematically into guidelines.

Our findings suggest that the concept of practical issues, within the evidence ecosystem, informed decision-making based on a broader set of data and evidence than the current standards and methods allow. The closest match we found is within the GRADE EtD framework, which considers acceptability and feasibility as key factors when going from evidence to recommendations (27, 28). We believe that systematically incorporating practical issues aligns with these factors and offers a specific and useful approach to addressing burden of treatment and practical issues. This raises broader questions about where practical issues best fit within the evidence ecosystem.

We have shown that inclusion of practical issues in the evidence ecosystem has the potential to better reflect all issues relevant for decision-making. By developing methods that allows for evidence production of practical issues, followed by evidence synthesis and inclusion alongside systematic reviews, the resulting guideline developing process has the potential to better reflect the most pertinent issues relevant for decision-making at the point of care.

Furthermore, the acknowledgement of burden of treatment generally and of practical issues particularly can have an impact on strength of recommendation itself, perhaps one of the highest levels of bidirectional exchange between the encounter decision aid and the guideline.

A strong recommendation may reduce the interest of clinicians and patients in critically considering the viability of simply implementing the recommendation, especially when the implementation process primarily relies on patients and caregivers and is practically challenging. Thus, in some instances, it may be necessary to downgrade these recommendations to invite clinicians to actively support patients in their implementation where the capacity to do so exists. Where it does not, clinicians can then shift their focus to aiding patients and caregivers in accessing resources from the healthcare system or the community to facilitate implementation. This process can potentially transform a weak recommendation into one against the proposed course of action.

In sum, the concept of practical issues holds significance within the evidence ecosystem as it allows for the inclusion of a broader range of data and evidence in decision-making in all steps of the ecosystem. The systematic inclusion of burden of treatment and practical issues has the potential to unveil a "black box" in the development of trustworthy guidelines. At the point of care, their inclusion may contribute to the work of making care fit for each patient.

6. Conclusion

1. Through iterative user-testing in real clinical encounters, we successfully developed a framework for SDM at the point of care with interactive and adaptable presentation formats. This framework translates GRADE evidence summaries and guidelines into encounter decision aids, providing a useful and intuitive tool that supports SDM.

2. We integrated the developed encounter decision aids into MAGICapp, enabling the semi-automatic production of encounter decision aids from evidence summaries and guidelines. This approach proved effective in real-life guideline development, reducing the resources required for creating or updating encounter decision aids.

3. User-testing of encounter decision aids revealed that patients wanted to discuss practical issues and burden of treatment, while clinicians lacked insights and tools to support these conversations. Consequently, we developed a generic framework for practical issues and implemented it in encounter decision aids. This framework effectively facilitated the identification and discussion of practical issues and treatment burdens for both clinicians and patients.

4. We successfully integrated the practical issues framework into MAGICapp and demonstrated the feasibility of including these issues in the real-life production and publication of guidelines in BMJ RapidRecs. This integration proved valuable as it complemented the SoF tables and influenced the development of guidelines. In certain instances, it impacted the strength of recommendations.

5. To inform a guideline panel creating recommendations on TAVI and SAVR, we conducted a systematic review on patients' values and preferences. The current body

of evidence was of suboptimal rigor, only addressing a minority of practical issues and treatment burdens and lacking patient-important outcomes defined by patient partners in the guideline panel. This highlights the inadequacy of even high-quality guideline development methods in informing guideline panels about patients' values and preferences. It raises the question of whether searches should extend beyond traditional sources of evidence to comprehensively capture all relevant aspects of patients' values and preferences.

6. The concept of practical issues holds significance within the evidence ecosystem as it allows for the inclusion of a broader range of data and evidence in decision-making and as a method to make guidelines more useful in making care fit. However, the appropriate fit of practical issues within the evidence ecosystem is still uncertain.

7. Perspectives for the future

The work presented in this thesis began more than a decade ago with the concept of connecting encounter decision aids to guidelines. Fast forward a decade, amidst the pandemic, guidelines, and recommendations, often characterized by uncertain evidence of low quality, became subjects of public debate. Unlike many other recommendations, COVID-19 guidelines often included detailed descriptions of practical issues (245, 246), which were frequently requested by the public.

Despite this, there is still a lack of optimal methods for effectively searching, appraising, and integrating practical issues and treatment burden within guidelines. Patients' values and preferences also suffer from inadequate inclusion in evidence synthesis and dissemination. This ongoing gap between research and patient priorities contributes to research waste, treatment burden, and mismatched care (247). Addressing these issues requires further research and innovative approaches within the evidence ecosystem.

Within a healthcare system with limited resources, the use of MAGICapp and innovations such as living systematic reviews and living guidelines, exemplifies how technology can enhance the evidence ecosystem (149). Nevertheless, the current evidence ecosystem faces common challenges in producing evidence summaries that adequately address complexity, multimorbidity, and the potential limitations in patient applicability.

With the higher prevalence of chronic conditions, multimorbidity and multiple treatment interventions, the burden of treatment and the work of being a patient is expected to increase and often change over time (135, 248). Most guidelines and decision aids do not support decision-making for patients with multiple and often chronic conditions. Implementation of SDM in practice goes beyond what tools for SDM can achieve on their own, and no tool can guarantee that a decision is shared. In this way, the work on decision aids mimics the challenges of translating recommendations into clinical practice. This argues for an even greater emphasis on contextualized care, argued to be an essential clinical skill (142)

Future research can explore the inclusion of existing frameworks, such as the practical issues framework and burden of treatment taxonomy, to promote more contextualized care within the evidence ecosystem through the development of tools that meet the needs of patients and clinicians. The optimal methods and form of these tools require further investigation.

From this perspective, the work to date offers promise in a future in which a bidirectional pathway between guidelines and clinical decision making exist and evolve within an enhanced evidence ecosystem. A future in which the SHARE-IT encounter decision aids will increasingly meet the needs of making care fit at the point of care just as guidelines help fulfill the promise of evidence-based care. A future in which the overarching vision of MAGIC becomes a reality.

8. References

1. MAGIC Evidence Ecosystem Foundation www.magicevidence.org [Available from: www.magicevidence.org.
2. Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. Chest. 2013;144(2):381-9.
3. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med. 2012;366(9):780-1.
4. Elwyn G, Frosch D, Rollnick S. Dual equipoise shared decision making: definitions for decision and behaviour support interventions. Implement Sci. 2009;4:75.
5. In: Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Clinical Practice Guidelines We Can Trust. Washington (DC)2011.
6. van der Weijden T, Legare F, Boivin A, Burgers JS, van Veenendaal H, Stiggelbout AM, et al. How to integrate individual patient values and preferences in clinical practice guidelines? A research protocol. Implement Sci. 2010;5:10.
7. Hoffmann TC, Montori VM, Del Mar C. The connection between evidence-based medicine and shared decision making. JAMA. 2014;312(13):1295-6.
8. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med. 2012;27(10):1361-7.
9. Frosch DL, Kaplan RM. Shared decision making in clinical medicine: past research and future directions. Am J Prev Med. 1999;17(4):285-94.
10. Evidence-Based Medicine Working G. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA. 1992;268(17):2420-5.
11. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2.
12. Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ. 2012;344:e256.
13. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017;390(10092):415-23.
14. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2017;2017(4).
15. Wieringa TH, Rodriguez-Gutierrez R, Spencer-Bonilla G, de Wit M, Ponce OJ, Sanchez-Herrera MF, et al. Decision aids that facilitate elements of shared decision making in chronic illnesses: a systematic review. Syst Rev. 2019;8(1):121.
16. Breslin M, Mullan RJ, Montori VM. The design of a decision aid about diabetes medications for use during the consultation with patients with type 2 diabetes. Patient Educ Couns. 2008;73(3):465-72.
17. Montori VM, Breslin M, Maleska M, Weymiller AJ. Creating a conversation: insights from the development of a decision aid. PLoS Med. 2007;4(8):e233.
18. Montori VM, LeBlanc A, Buchholz A, Stilwell DL, Tsapas A. Basing information on comprehensive, critically appraised, and up-to-date syntheses of the scientific evidence: a quality dimension of the International Patient Decision Aid Standards. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S5.
19. Elwyn G, Lloyd A, Joseph-Williams N, Cording E, Thomson R, Durand MA, et al. Option Grids: shared decision making made easier. Patient Educ Couns. 2013;90(2):207-12.

20. Elwyn G, Kreuwel I, Durand MA, Sivell S, Joseph-Williams N, Evans R, et al. How to develop web-based decision support interventions for patients: a process map. Patient Educ Couns. 2011;82(2):260-5.
21. Hoffmann TC, Bakhit M, Durand MA, Perestelo-Perez L, Saunders C, Brito JP. Basing Information on Comprehensive, Critically Appraised, and Up-to-Date Syntheses of the Scientific Evidence: An Update from the International Patient Decision Aid Standards. Med Decis Making. 2021;41(7):755-67.
22. Elwyn G, Scholl I, Tietbohl C, Mann M, Edwards AG, Clay C, et al. "Many miles to go ...": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC Med Inform Decis Mak. 2013;13 Suppl 2(Suppl 2):S14.
23. Greenhalgh T, Howick J, Maskrey N, Evidence Based Medicine Renaissance G. Evidence based medicine: a movement in crisis? BMJ. 2014;348:g3725.
24. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol. 2011;64(4):395-400.
25. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490.
26. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.
27. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.
28. Alonso-Coello P, Schunemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.
29. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401-6.
30. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol. 2013;66(7):726-35.
31. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol. 2013;66(7):719-25.
32. Schunemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. J Clin Epidemiol. 2020;122:129-41.
33. Schunemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. J Clin Epidemiol. 2020;122:142-52.
34. Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. J Clin Epidemiol. 2020;121:62-70.
35. Zhang Y, Alonso-Coello P, Guyatt GH, Yepes-Nunez JJ, Akl EA, Hazlewood G, et al. GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. J Clin Epidemiol. 2019;111:94-104.

36. Zhang Y, Coello PA, Brożek J, Wiercioch W, Etxeandia-Ikobaltzeta I, Akl EA, et al. Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. Health and Quality of Life Outcomes. 2017;15(1):52.
37. Zhang Y, Coello PA, Guyatt GH, Yepes-Nunez JJ, Akl EA, Hazlewood G, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. J Clin Epidemiol. 2019;111:83-93.
38. GRADE working group [Available from: https://www.gradeworkinggroup.org.
39. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64(12):1311-6.
40. Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group clarifies the construct of certainty of evidence. J Clin Epidemiol. 2017;87:4-13.
41. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94.
42. Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. J Clin Epidemiol. 2010;63(6):620-6.
43. Rosenbaum S. Improving the user experience of evidence. https://aho.brage.unit.no/aho-xmlui/bitstream/handle/11250/93062/Rosenbaum_thesis.pdf?sequence=12010.
44. Vandvik PO, Otto CM, Siemieniuk RA, Bagur R, Guyatt GH, Lytvyn L, et al. Transcatheter or surgical aortic valve replacement for patients with severe, symptomatic, aortic stenosis at low to intermediate surgical risk: a clinical practice guideline. BMJ. 2016;354:i5085.
45. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci. 2018;13(Suppl 1):2.
46. Downe S, Finlayson KW, Lawrie TA, Lewin SA, Glenton C, Rosenbaum S, et al. Qualitative Evidence Synthesis (QES) for Guidelines: Paper 1 - Using qualitative evidence synthesis to inform guideline scope and develop qualitative findings statements. Health Res Policy Syst. 2019;17(1):76.
47. Kristiansen A, Brandt L, Alonso-Coello P, Agoritsas T, Akl EA, Conboy T, et al. Development of a novel, multilayered presentation format for clinical practice guidelines. Chest. 2015;147(3):754-63.
48. Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P, et al. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med. 2012;156(7):525-31.
49. Hill J, Bullock I, Alderson P. A summary of the methods that the National Clinical Guideline Centre uses to produce clinical guidelines for the National Institute for Health and Clinical Excellence. Ann Intern Med. 2011;154(11):752-7.
50. World Health O. WHO handbook for guideline development. 2nd ed ed. Geneva: World Health Organization; 2014 2014.
51. Murad MH. Clinical Practice Guidelines: A Primer on Development and Dissemination. Mayo Clin Proc. 2017;92(3):423-33.
52. Makarski J, Brouwers MC, Enterprise A. The AGREE Enterprise: a decade of advancing clinical practice guidelines. Implement Sci. 2014;9:103.

53. Jue JJ, Cunningham S, Lohr K, Shekelle P, Shiffman R, Robbins C, et al. Developing and Testing the Agency for Healthcare Research and Quality's National Guideline Clearinghouse Extent of Adherence to Trustworthy Standards (NEATS) Instrument. Ann Intern Med. 2019;170(7):480-7.
54. Chen Y, Yang K, Marusic A, Qaseem A, Meerpohl JJ, Flottorp S, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. Ann Intern Med. 2017;166(2):128-32.
55. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. BMJ. 2008;336(7652):1049-51.
56. Guyatt G. Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice. Second edition ed: McGraw-Hill Professional; 2008.
57. Dwamena F, Holmes-Rovner M, Gaulden CM, Jorgenson S, Sadigh G, Sikorskii A, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. Cochrane Database Syst Rev. 2012;12:CD003267.
58. Boivin A, Currie K, Fervers B, Gracia J, James M, Marshall C, et al. Patient and public involvement in clinical guidelines: international experiences and future perspectives. Qual Saf Health Care. 2010;19(5):e22.
59. van der Weijden T, Boivin A, Burgers J, Schunemann HJ, Elwyn G. Clinical practice guidelines and patient decision aids. An inevitable relationship. J Clin Epidemiol. 2012;65(6):584-9.
60. Krahn M, Naglie G. The next step in guideline development: incorporating patient preferences. JAMA. 2008;300(4):436-8.
61. Armstrong MJ, Bloom JA. Patient involvement in guidelines is poor five years after institute of medicine standards: review of guideline methodologies. Res Involv Engagem. 2017;3:19.
62. Kung J, Miller RR, Mackowiak PA. Failure of clinical practice guidelines to meet institute of medicine standards: Two more decades of little, if any, progress. Arch Intern Med. 2012;172(21):1628-33.
63. Siemieniuk RA, Agoritsas T, Macdonald H, Guyatt GH, Brandt L, Vandvik PO. Introduction to BMJ Rapid Recommendations. BMJ. 2016;354.
64. Petkovic J, Riddle A, Akl EA, Khabsa J, Lytvyn L, Atwere P, et al. Protocol for the development of guidance for stakeholder engagement in health and healthcare guideline development and implementation. Syst Rev. 2020;9(1):21.
65. Agoritsas T, Merglen A, Heen AF, Kristiansen A, Neumann I, Brito JP, et al. UpToDate adherence to GRADE criteria for strong recommendations: an analytical survey. BMJ Open. 2017;7(11):e018593.
66. Elwyn G, Durand MA, Song J, Aarts J, Barr PJ, Berger Z, et al. A three-talk model for shared decision making: multistage consultation process. BMJ. 2017;359:j4891.
67. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ. 2010;341:c5146.
68. Beauchamp T L CJF. Principles of biomedical ethics. 5th ed. ed: Oxford University Press; 2001.
69. Hoffmann TC, Legare F, Simmons MB, McNamara K, McCaffery K, Trevena LJ, et al. Shared decision making: what do clinicians need to know and why should they bother? Med J Aust. 2014;201(1):35-9.
70. Lov om pasient- og brukerrettigheter (pasient- og brukerrettighetsloven), (2019).
71. Legare F, Thompson-Leduc P. Twelve myths about shared decision making. Patient Educ Couns. 2014;96(3):281-6.
72. Veatch RM. Models for ethical medicine in a revolutionary age. What physician-patient roles foster the most ethical realtionship? Hastings Cent Rep. 1972;2(3):5-7.

73. Stiggelbout AM, Pieterse AH, De Haes JC. Shared decision making: Concepts, evidence, and practice. Patient Educ Couns. 2015;98(10):1172-9.
74. Elwyn G, Frosch D, Volandes AE, Edwards A, Montori VM. Investing in deliberation: a definition and classification of decision support interventions for people facing difficult health decisions. Med Decis Making. 2010;30(6):701-11.
75. Berger Z, Galasinski D, Scalia P, Dong K, Blunt HB, Elwyn G. The submissive silence of others: Examining definitions of shared decision making. Patient Educ Couns. 2021.
76. Bomhof-Roordink H, Gartner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open. 2019;9(12):e031763.
77. Agency for Healthcare Research and Quality. The SHARE Approach. 2020 [Available from: https://www.ahrq.gov/health-literacy/professional-training/shared-decision/index.html.
78. Hargraves IG, Montori VM, Brito JP, Kunneman M, Shaw K, LaVecchia C, et al. Purposeful SDM: A problem-based approach to caring for patients with shared decision making. Patient Educ Couns. 2019;102(10):1786-92.
79. Montori VM, Ruissen MM, Branda ME, Hargraves IG, Kunneman M. Problem-based shared decision making: The role of canonical SDM steps. Health Expect. 2023;26(1):282-9.
80. Kunneman M, Branda ME, Hargraves I, Pieterse AH, Montori VM. Fostering Choice Awareness for Shared Decision Making: A Secondary Analysis of Video-Recorded Clinical Encounters. Mayo Clin Proc Innov Qual Outcomes. 2018;2(1):60-8.
81. Legare F, Stacey D, Pouliot S, Gauvin FP, Desroches S, Kryworuchko J, et al. Interprofessionalism and shared decision-making in primary care: a stepwise approach towards a new model. J Interprof Care. 2011;25(1):18-25.
82. Legare F, Ratte S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient Educ Couns. 2008;73(3):526-35.
83. Pel-Littel RE, Snaterse M, Teppich NM, Buurman BM, van Etten-Jamaludin FS, van Weert JCM, et al. Barriers and facilitators for shared decision making in older patients with multiple chronic conditions: a systematic review. BMC Geriatr. 2021;21(1):112.
84. Waddell A, Lennox A, Spassova G, Bragge P. Barriers and facilitators to shared decision-making in hospitals from policy to practice: a systematic review. Implement Sci. 2021;16(1):74.
85. Scholl I, LaRussa A, Hahlweg P, Kobrin S, Elwyn G. Organizational- and system-level characteristics that influence implementation of shared decision-making and strategies to address them - a scoping review. Implement Sci. 2018;13(1):40.
86. Charles C, Gafni A, Whelan T, O'Brien MA. Cultural influences on the physician-patient encounter: The case of shared treatment decision-making. Patient Educ Couns. 2006;63(3):262-7.
87. Tan NQP, Maki KG, Lopez-Olivo MA, Geng Y, Volk RJ. Cultural influences on shared decision-making among Asian Americans: A systematic review and meta-synthesis of qualitative studies. Patient Educ Couns. 2023;106:17-30.
88. Joseph-Williams N, Abhyankar P, Boland L, Bravo P, Brenner AT, Brodney S, et al. What Works in Implementing Patient Decision Aids in Routine Clinical Settings? A Rapid Realist Review and Update from the International Patient Decision Aid Standards Collaboration. Med Decis Making. 2021;41(7):907-37.
89. Hoffmann TC, Bennett S, Tomsett C, Del Mar C. Brief training of student clinicians in shared decision making: a single-blind randomized controlled trial. J Gen Intern Med. 2014;29(6):844-9.

90. Dobler CC, Sanchez M, Gionfriddo MR, Alvarez-Villalobos NA, Singh Ospina N, Spencer-Bonilla G, et al. Impact of decision aids used during clinical encounters on clinician outcomes and consultation length: a systematic review. BMJ Qual Saf. 2019;28(6):499-510.
91. LeBlanc A, Kenny DA, O'Connor AM, Legare F. Decisional conflict in patients and their physicians: a dyadic approach to shared decision making. Med Decis Making. 2009;29(1):61-8.
92. O'Connor AM, Drake ER, Fiset V, Graham ID, Laupacis A, Tugwell P. The Ottawa patient decision aids. Eff Clin Pract. 1999;2(4):163-70.
93. O'Connor AM, Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, et al. A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. Patient Educ Couns. 1998;33(3):267-79.
94. Simons M, Rapport F, Zurynski Y, Stoodley M, Cullis J, Davidson AS. Links between evidence-based medicine and shared decision-making in courses for doctors in training: a scoping review. BMJ Open. 2022;12(4):e057335.
95. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ. 2006;333(7565):417.
96. Montori VM, Kunneman M, Brito JP. Shared Decision Making and Improving Health Care: The Answer Is Not In. JAMA. 2017;318(7):617-8.
97. Raats CJ, van Veenendaal H, Versluijs MM, Burgers JS. A generic tool for development of decision aids based on clinical practice guidelines. Patient Educ Couns. 2008;73(3):413-7.
98. International Patient Decision Aid Standards (IPDAS) Collaboration [Available from: http://ipdas.ohri.ca.
99. Volk RJ, Llewellyn-Thomas H, Stacey D, Elwyn G. Ten years of the International Patient Decision Aid Standards Collaboration: evolution of the core dimensions for assessing the quality of patient decision aids. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S1.
100. Stacey D, Volk RJ, Leads IEU. The International Patient Decision Aid Standards (IPDAS) Collaboration: Evidence Update 2.0. Med Decis Making. 2021;41(7):729-33.
101. Witteman HO, Maki KG, Vaisson G, Finderup J, Lewis KB, Dahl Steffensen K, et al. Systematic Development of Patient Decision Aids: An Update from the IPDAS Collaboration. Med Decis Making. 2021;41(7):736-54.
102. Feldman-Stewart D, Brundage MD, McConnell BA, MacKillop WJ. Practical issues in assisting shared decision-making. Health Expect. 2000;3(1):46-54.
103. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van der Weijden T. A systematic development process for patient decision aids. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S2.
104. Dannenberg MD, Durand MA, Montori VM, Reilly C, Elwyn G. Existing evidence summarization methods cannot guarantee trustworthy patient decision aids. J Clin Epidemiol. 2018;102:69-77.
105. Martin RW, Brogard Andersen S, O'Brien MA, Bravo P, Hoffmann T, Olling K, et al. Providing Balanced Information about Options in Patient Decision Aids: An Update from the International Patient Decision Aid Standards. Med Decis Making. 2021;41(7):780-800.
106. Edwards A, Elwyn G, Covey J, Matthews E, Pill R. Presenting risk information--a review of the effects of "framing" and other manipulations on patient outcomes. J Health Commun. 2001;6(1):61-82.
107. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PK, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S7.

108. Trevena LJ, Bonner C, Okan Y, Peters E, Gaissmaier W, Han PKJ, et al. Current Challenges When Using Numbers in Patient Decision Aids: Advanced Concepts. Med Decis Making. 2021;41(7):834-47.
109. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Making. 2007;27(5):696-713.
110. Fagerlin A, Pignone M, Abhyankar P, Col N, Feldman-Stewart D, Gavaruzzi T, et al. Clarifying values: an updated review. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S8.
111. Witteman HO, Ndjaboue R, Vaisson G, Dansokho SC, Arnold B, Bridges JFP, et al. Clarifying Values: An Updated and Expanded Systematic Review and Meta-Analysis. Med Decis Making. 2021;41(7):801-20.
112. Noseworthy PA, Branda ME, Kunneman M, Hargraves IG, Sivly AL, Brito JP, et al. Effect of Shared Decision-Making for Stroke Prevention on Treatment Adherence and Safety Outcomes in Patients With Atrial Fibrillation: A Randomized Clinical Trial. J Am Heart Assoc. 2022;11(2):e023048.
113. Kunneman M, Branda ME, Ridgeway JL, Tiedje K, May CR, Linzer M, et al. Making sense of diabetes medication decisions: a mixed methods cluster randomized trial using a conversation aid intervention. Endocrine. 2022;75(2):377-91.
114. Karagiannis T, Liakos A, Branda ME, Athanasiadou E, Mainou M, Boura P, et al. Use of the Diabetes Medication Choice Decision Aid in patients with type 2 diabetes in Greece: a cluster randomised trial. BMJ Open. 2016;6(11):e012185.
115. Perestelo-Perez L, Rivero-Santana A, Boronat M, Sanchez-Afonso JA, Perez-Ramos J, Montori VM, et al. Effect of the statin choice encounter decision aid in Spanish patients with type 2 diabetes: A randomized trial. Patient Educ Couns. 2016;99(2):295-9.
116. LeBlanc A, Herrin J, Williams MD, Inselman JW, Branda ME, Shah ND, et al. Shared Decision Making for Antidepressants in Primary Care: A Cluster Randomized Trial. JAMA Intern Med. 2015;175(11):1761-70.
117. Kew KM, Malik P, Aniruddhan K, Normansell R. Shared decision-making for people with asthma. Cochrane Database Syst Rev. 2017;10:CD012330.
118. Quigley M, Dillon MP, Fatone S. Development of shared decision-making resources to help inform difficult healthcare decisions: An example focused on dysvascular partial foot and transtibial amputations. Prosthet Orthot Int. 2018;42(4):378-86.
119. Knops AM, Legemate DA, Goossens A, Bossuyt PM, Ubbink DT. Decision aids for patients facing a surgical treatment decision: a systematic review and meta-analysis. Ann Surg. 2013;257(5):860-6.
120. Volk RJ, Lowenstein LM, Leal VB, Escoto KH, Cantor SB, Munden RF, et al. Effect of a Patient Decision Aid on Lung Cancer Screening Decision-Making by Persons Who Smoke: A Randomized Clinical Trial. JAMA Netw Open. 2020;3(1):e1920362.
121. Tikkinen KAO, Dahm P, Lytvyn L, Heen AF, Vernooij RWM, Siemieniuk RAC, et al. Prostate cancer screening with prostate-specific antigen (PSA) test: a clinical practice guideline. BMJ. 2018;362:k3581.
122. Scalia P, Durand MA, Berkowitz JL, Ramesh NP, Faber MJ, Kremer JAM, et al. The impact and utility of encounter patient decision aids: Systematic review, meta-analysis and narrative synthesis. Patient Educ Couns. 2019;102(5):817-41.
123. Hoffmann TC, Del Mar C. Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Intern Med. 2015;175(2):274-86.
124. Hoffmann TC, Del Mar C. Clinicians' Expectations of the Benefits and Harms of Treatments, Screening, and Tests: A Systematic Review. JAMA Intern Med. 2017;177(3):407-19.

125. Legare F, Stacey D, Turcotte S, Cossi MJ, Kryworuchko J, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev. 2014(9):CD006732.
126. Wyatt KD, Branda ME, Anderson RT, Pencille LJ, Montori VM, Hess EP, et al. Peering into the black box: a meta-analysis of how clinicians use decision aids during clinical encounters. Implement Sci. 2014;9:26.
127. May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ. 2009;339:b2803.
128. Eton DT, Ramalho de Oliveira D, Egginton JS, Ridgeway JL, Odell L, May CR, et al. Building a measurement framework of burden of treatment in complex patients with chronic conditions: a qualitative study. Patient Relat Outcome Meas. 2012;3:39-49.
129. Demain S, Goncalves AC, Areia C, Oliveira R, Marcos AJ, Marques A, et al. Living with, managing and minimising treatment burden in long term conditions: a systematic review of qualitative research. PLoS One. 2015;10(5):e0125457.
130. Tran VT, Montori VM, Eton DT, Baruch D, Falissard B, Ravaud P. Development and description of measurement properties of an instrument to assess treatment burden among patients with multiple chronic conditions. BMC Med. 2012;10:68.
131. Mair FS, May CR. Thinking about the burden of treatment. BMJ : British Medical Journal. 2014;349.
132. May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S. Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. BMC Health Serv Res. 2014;14.
133. Shippee ND, Shah ND, May CR, Mair FS, Montori VM. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. J Clin Epidemiol. 2012;65(10):1041-51.
134. Unruh KT, Pratt W. The Invisible Work of Being a Patient and Implications for Health Care: "[the doctor is] my business partner in the most important business in my life, staying alive". Conf Proc Ethnogr Prax Ind Conf. 2008;2008(1):40-50.
135. Sav A, King MA, Whitty JA, Kendall E, McMillan SS, Kelly F, et al. Burden of treatment for chronic illness: a concept analysis and review of the literature. Health Expect. 2015;18(3):312-24.
136. Tarlov AR, Ware JE, Jr., Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. JAMA. 1989;262(7):925-30.
137. Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, et al. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA. 1989;262(7):907-13.
138. Eton DT, Ridgeway JL, Egginton JS, Tiedje K, Linzer M, Boehm DH, et al. Finalizing a measurement framework for the burden of treatment in complex patients with chronic conditions. Patient Relat Outcome Meas. 2015;6:117-26.
139. Tran VT, Barnes C, Montori VM, Falissard B, Ravaud P. Taxonomy of the burden of treatment: a multi-country web-based qualitative study of patients with chronic conditions. BMC Med. 2015;13:115.
140. Tran VT, Montori VM, Ravaud P. Is My Patient Overwhelmed?: Determining Thresholds for Acceptable Burden of Treatment Using Data From the ComPaRe e-Cohort. Mayo Clin Proc. 2020;95(3):504-12.
141. Weiner SJ. Contextualizing medical decisions to individualize care: lessons from the qualitative sciences. J Gen Intern Med. 2004;19(3):281-5.
142. Weiner SJ. Contextualizing care: An essential and measurable clinical competency. Patient Educ Couns. 2022;105(3):594-8.

143. Spencer-Bonilla G, Quinones AR, Montori VM, International Minimally Disruptive Medicine W. Assessing the Burden of Treatment. J Gen Intern Med. 2017;32(10):1141-5.
144. Brandt L, Vandvik PO, Alonso-Coello P, Akl EA, Thornton J, Rigau D, et al. Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial. BMJ Open. 2017;7(2):e011569.
145. Vandvik PO, Brandt L. Future of Evidence Ecosystem Series: Evidence ecosystems and learning health systems: why bother? J Clin Epidemiol. 2020;123:166-70.
146. Ravaud P, Crequit P, Williams HC, Meerpohl J, Craig JC, Boutron I. Future of evidence ecosystem series: 3. From an evidence synthesis ecosystem to an evidence ecosystem. J Clin Epidemiol. 2020;123:153-61.
147. Crequit P, Boutron I, Meerpohl J, Williams HC, Craig J, Ravaud P. Future of evidence ecosystem series: 2. current opportunities and need for better tools and methods. J Clin Epidemiol. 2020;123:143-52.
148. Siemieniuk RA, Bartoszko JJ, Diaz Martinez JP, Kum E, Qasim A, Zeraatkar D, et al. Antibody and cellular therapies for treatment of covid-19: a living systematic review and network meta-analysis. BMJ. 2021;374:n2231.
149. Vandvik PO, Askie L, Glen F, Tendal B, Agoritsas T. COVID-19: living guidelines help fix cracks in evidence pipeline. Nature. 2021;595(7866):172.
150. Lamontagne F, Agarwal A, Rochwerg B, Siemieniuk RA, Agoritsas T, Askie L, et al. A living WHO guideline on drugs for covid-19. BMJ. 2020;370:m3379.
151. Turner T, Elliott J, Tendal B, Vogel JP, Norris S, Tate R, et al. The Australian living guidelines for the clinical care of people with COVID-19: What worked, what didn't and why, a mixed methods process evaluation. PLoS One. 2022;17(1):e0261479.
152. Hill K, English C, Campbell BCV, McDonald S, Pattuwage L, Bates P, et al. Feasibility of national living guideline methods: The Australian Stroke Guidelines. J Clin Epidemiol. 2022;142:184-93.
153. Witteman HO, Dansokho SC, Colquhoun H, Coulter A, Dugas M, Fagerlin A, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev. 2015;4:11.
154. Iivari J IN. Varieties of user-centredness: An analysis of four systems development methods. Information Systems Journal Wiley Online Library. 2011;2011;21(2):125–153.
155. P M. User experience design.[Available from: https://semanticstudios.com/user_experience_design/.
156. DECIDE. DECIDE https://www.decide-collaboration.eu2011-2015 [Available from: http://www.decide-collaboration.eu.
157. Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brozek J, et al. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. Implement Sci. 2013;8:6.
158. Siemieniuk RA, Agoritsas T, Macdonald H, Guyatt GH, Brandt L, Vandvik PO. Introduction to BMJ Rapid Recommendations. BMJ. 2016;354:i5191.
159. Agoritsas T, Heen AF, Brandt L, Alonso-Coello P, Kristiansen A, Akl EA, et al. Decision aids that really promote shared decision making: the pace quickens. BMJ. 2015;350:g7624.
160. Heen AF, Vandvik PO, Brandt L, Achille F, Guyatt G, Akl E, et al. Decision aids linked to evidence summaries and clinical practice guidelines: Results from user-testing in clinical encounters. Bmc Medical Informatics and Decision Making. 2021.
161. Elwyn G, Stiel M, Durand MA, Boivin J. The design of patient decision support interventions: addressing the theory-practice gap. Journal of evaluation in clinical practice. 2011;17(4):565-74.

162. Wroblewski L. Mobile first: A Book Apart; 2011.
163. Cates C. Visual Rx www.nntonline.net/visualrx2012 [Available from: www.nntonline.net/visualrx.
164. Blueprint. http://blueprintcss.org2012 [Available from: http://blueprintcss.org.
165. Guyatt GH, Akl EA, Crowther M, Schunemann HJ, Gutterman DD, Lewis SZ. Introduction to the ninth edition: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012;141(2 Suppl):48S-52S.
166. Holbrook A, Schulman S, Witt DM, Vandvik PO, Fish J, Kovacs MJ, et al. Evidence-based management of anticoagulant therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012;141(2 Suppl):e152S-e84S.
167. Retningslinjer for antitrombotisk behandling og profylakse. https://files.magicapp.org/guideline/b1136bce-6dd7-4a62-adf6-1df450bffa24/1_1/pdf/published_guideline_414-1_1.pdf: Norsk Selskap for Trombose og Hemostase; 2015.
168. Davies C, Pan H, Godwin J, Gray R, Arriagada R, Raina V, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet. 2013;381(9869):805-16.
169. Edwards A, Elwyn G, Hood K, Robling M, Atwell C, Holmes-Rovner M, et al. The development of COMRADE--a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. Patient Educ Couns. 2003;50(3):311-22.
170. Edwards A, Elwyn G Fau - Hood K, Hood K Fau - Robling M, Robling M Fau - Atwell C, Atwell C Fau - Holmes-Rovner M, Holmes-Rovner M Fau - Kinnersley P, et al. The development of COMRADE--a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. (0738-3991 (Print)).
171. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-88.
172. Heen AF, Vandvik PO, Brandt L, Montori VM, Lytvyn L, Guyatt G, et al. A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines. J Clin Epidemiol. 2020;129:104-13.
173. DIPEX. Health Experiences Research Group 2019 [Available from: https://www.healthtalk.org/HERG.
174. HealthTalkOnline. HealthTalk.org 2016 [
175. Marrin K BK, Marie-Anne Durand, Adrian Edwards, Amy Lloyd, Victoria Thomas, Glyn Elwyn. Fast and frugal tools for shared decision making: How to develop Option Grids. European Journal for person centered healthcare. 2013;1(1).
176. Strauss A, Corbin J. The basics of qualitative research: grounded theory procedures and techniques. . 4th edition ed. Newbury Park, CA: SAGE; 1990.
177. Kennedy TJ, Lingard LA. Making sense of grounded theory in medical education. Med Educ. 2006;40(2):101-8.
178. Lytvyn L, Guyatt GH, Manja V, Siemieniuk RA, Zhang Y, Agoritsas T, et al. Patient values and preferences on transcatheter or surgical aortic valve replacement therapy for aortic stenosis: a systematic review. BMJ Open. 2016;6(9):e014327.
179. Lytvyn L, Siemieniuk RA, Dilmitis S, Ion A, Chang Y, Bala MM, et al. Values and preferences of women living with HIV who are pregnant, postpartum or considering

pregnancy on choice of antiretroviral therapy during pregnancy. BMJ Open. 2017;7(9):e019023.
180. Vernooij RWM, Lytvyn L, Pardo-Hernandez H, Albarqouni L, Canelo-Aybar C, Campbell K, et al. Values and preferences of men for undergoing prostate-specific antigen screening for prostate cancer: a systematic review. BMJ Open. 2018;8(9):e025470.
181. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6. www.training.cochrane.org/handbook: Cochrane; 2019.
182. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol. 2008;8:45.
183. Garner P, Hopewell S, Chandler J, MacLehose H, Schunemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ. 2016;354:i3507.
184. Selva A, Sola I, Zhang Y, Pardo-Hernandez H, Haynes RB, Martinez Garcia L, et al. Development and use of a content search strategy for retrieving studies on patients' views and preferences. Health Qual Life Outcomes. 2017;15(1):126.
185. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000;283(15):2008-12.
186. Siemieniuk RA, Agoritsas T, Manja V, Devji T, Chang Y, Bala MM, et al. Transcatheter versus surgical aortic valve replacement in patients with severe aortic stenosis at low and intermediate risk: systematic review and meta-analysis. BMJ. 2016;354:i5130.
187. Sandelowski M, Voils CI, Barroso J. Defining and Designing Mixed Research Synthesis Studies. Res Sch. 2006;13(1):29.
188. CASP Qualitative Checklist. https://casp-uk.net/casp-tools-checklists/2019 [
189. Lewin S, Bohren M, Rashidian A, Munthe-Kaas H, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 2: how to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table. Implement Sci. 2018;13(Suppl 1):10.
190. Colvin CJ, Garside R, Wainwright M, Munthe-Kaas H, Glenton C, Bohren MA, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 4: how to assess coherence. Implement Sci. 2018;13(Suppl 1):13.
191. MAGIC. SHARE-IT educational module [Available from: https://magicevidence.org/share-it/tour/.
192. Kristiansen A, Brandt L, Agoritsas T, Akl EA, Berge E, Flem Jacobsen A, et al. 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. Chest. 2014;146(3):735-61.
193. Kristiansen A, Brandt L, Agoritsas T, Akl EA, Berge E, Bondi J, et al. Adaptation of trustworthy guidelines developed using the GRADE methodology: a novel five-step process. Chest. 2014;146(3):727-34.
194. Witteman HO, Vaisson G, Provencher T, Chipenda Dansokho S, Colquhoun H, Dugas M, et al. An 11-Item Measure of User- and Human-Centered Design for Personal Health Tools (UCD-11): Development and Validation. J Med Internet Res. 2021;23(3):e15032.
195. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs. 2009;27(3):175-83.
196. Vaisson G, Provencher T, Dugas M, Trottier ME, Chipenda Dansokho S, Colquhoun H, et al. User Involvement in the Design and Development of Patient Decision Aids and Other Personal Health Tools: A Systematic Review. Med Decis Making. 2021;41(3):261-74.

197. Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud. 2020;108:103632.
198. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet. 2001;358(9280):483-8.
199. Chun Tie Y, Birks M, Francis K. Grounded theory research: A design framework for novice researchers. SAGE Open Med. 2019;7:2050312118822927.
200. Dobler CC, Harb N, Maguire CA, Armour CL, Coleman C, Murad MH. Treatment burden should be included in clinical practice guidelines. BMJ. 2018;363:k4065.
201. Shaikh HJ, Brodell BJ, Soin SP. "Social media perceptions of hip fractures: A two-year cross-sectional analysis". Injury. 2023.
202. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open. 2013;3(1).
203. Roberti J, Cummings A, Myall M, Harvey J, Lippiett K, Hunt K, et al. Work of being an adult patient with chronic kidney disease: a systematic review of qualitative studies. BMJ Open. 2018;8(9):e023507.
204. Vaughan Sarrazin MS, Cram P, Mazur A, Ward M, Reisinger HS. Patient perspectives of dabigatran: analysis of online discussion forums. Patient. 2014;7(1):47-54.
205. IPDAS - The implementation of patient decision support interventions into routine clinical practice: a systematic review. http://ipdas.ohri.ca/IPDAS-Implementation.pdf
206. Barry MJ, Harter M, Househ M, Steffensen KD, Stacey D. What can we learn from rapidly developed patient decision aids produced during the covid-19 pandemic? BMJ. 2022;378:e071530.
207. Zeballos-Palacios CL, Hargraves IG, Noseworthy PA, Branda ME, Kunneman M, Burnett B, et al. Developing a Conversation Aid to Support Shared Decision Making: Reflections on Designing Anticoagulation Choice. Mayo Clin Proc. 2019;94(4):686-96.
208. Zorzela L, Golder S, Liu Y, Pilkington K, Hartling L, Joffe A, et al. Quality of reporting in systematic reviews of adverse events: systematic review. BMJ. 2014;348:f7668.
209. Mahady SE, Schlub T, Bero L, Moher D, Tovey D, George J, et al. Side effects are incompletely reported among systematic reviews in gastroenterology. J Clin Epidemiol. 2015;68(2):144-53.
210. Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida P, Altman DG, et al. PRISMA harms checklist: improving harms reporting in systematic reviews. BMJ. 2016;352:i157.
211. Altman DG. The scandal of poor medical research. BMJ. 1994;308(6924):283-4.
212. Yepes-Nunez JJ, Li SA, Guyatt G, Jack SM, Brozek JL, Beyene J, et al. Development of the summary of findings table for network meta-analysis. J Clin Epidemiol. 2019;115:1-13.
213. Puhan MA, Schunemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ. 2014;349:g5630.
214. Magic Evidence Ecosystem Foundation. https://www.magicevidence.org/match-it [
215. Helsingen LM, Vandvik PO, Jodal HC, Agoritsas T, Lytvyn L, Anderson JC, et al. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline. BMJ. 2019;367:l5515.
216. Hao Q, Aertgeerts B, Guyatt G, Bekkering GE, Vandvik PO, Khan SU, et al. PCSK9 inhibitors and ezetimibe for the reduction of cardiovascular events: a clinical practice guideline with risk-stratified recommendations. BMJ. 2022;377:e069066.
217. Li S, Vandvik PO, Lytvyn L, Guyatt GH, Palmer SC, Rodriguez-Gutierrez R, et al. SGLT-2 inhibitors or GLP-1 receptor agonists for adults with type 2 diabetes: a clinical practice guideline. BMJ. 2021;373:n1091.

218. Bekkering GE, Agoritsas T, Lytvyn L, Heen AF, Feller M, Moutzouri E, et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. BMJ. 2019;365:l2006.
219. Kunneman M, Griffioen IPM, Labrie NHM, Kristiansen M, Montori VM, van Beusekom MM, et al. Making care fit manifesto. BMJ Evid Based Med. 2023;28(1):5-6.
220. Witteman HO, Gavaruzzi T, Scherer LD, Pieterse AH, Fuhrel-Forbis A, Chipenda Dansokho S, et al. Effects of Design Features of Explicit Values Clarification Methods: A Systematic Review. Med Decis Making. 2016;36(6):760-76.
221. Dowie J, Kaltoft MK. Strong Recommendations Are Inappropriate in Person-Centred Care: The Case of Anti-Platelet Therapy. Stud Health Technol Inform. 2019;262:110-3.
222. LeBlanc A, Wang AT, Wyatt K, Branda ME, Shah ND, Van Houten H, et al. Encounter Decision Aid vs. Clinical Decision Support or Usual Care to Support Patient-Centered Treatment Decisions in Osteoporosis: The Osteoporosis Choice Randomized Trial II. PLoS One. 2015;10(5):e0128063.
223. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med. 1997;44(5):681-92.
224. Frazao TDC, Camilo DGG, Cabral ELS, Souza RP. Multicriteria decision analysis (MCDA) in health care: a systematic review of the main characteristics and methodological steps. BMC Med Inform Decis Mak. 2018;18(1):90.
225. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med. 2020;3:17.
226. Brito JP, Castaneda-Guarderas A, Gionfriddo MR, Ospina NS, Maraka S, Dean DS, et al. Development and Pilot Testing of an Encounter Tool for Shared Decision Making About the Treatment of Graves' Disease. Thyroid. 2015;25(11):1191-8.
227. Lyu Y, Xu Q, Yang Z, Liu J. Prediction of patient choice tendency in medical decision-making based on machine learning algorithm. Front Public Health. 2023;11:1087358.
228. Jayakumar P, Moore MG, Furlough KA, Uhler LM, Andrawis JP, Koenig KM, et al. Comparison of an Artificial Intelligence-Enabled Patient Decision Aid vs Educational Material on Decision Quality, Shared Decision-Making, Patient Experience, and Functional Outcomes in Adults With Knee Osteoarthritis: A Randomized Clinical Trial. JAMA Netw Open. 2021;4(2):e2037107.
229. Biller-Andorno N, Biller A. Algorithm-Aided Prediction of Patient Preferences - An Ethics Sneak Peek. N Engl J Med. 2019;381(15):1480-5.
230. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46(3):205-11.
231. Reuben DB, Tinetti ME. Goal-oriented patient care--an alternative health outcomes paradigm. (1533-4406 (Electronic)).
232. Alonso-Coello P, Montori VM, Diaz MG, Devereaux PJ, Mas G, Diez AI, et al. Values and preferences for oral antithrombotic therapy in patients with atrial fibrillation: physician and patient perspectives. Health Expect. 2015;18(6):2318-27.
233. Muhlbacher AC, Juhnke C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? Appl Health Econ Health Policy. 2013;11(3):163-80.
234. Harrison M, Milbers K, Hudson M, Bansback N. Do patients and health care providers have discordant preferences about which aspects of treatments matter most? Evidence from a systematic review of discrete choice experiments. BMJ Open. 2017;7(5):e014719.

235. Gartner FR, Portielje JE, Langendam M, Hairwassers D, Agoritsas T, Gijsen B, et al. Role of patient preferences in clinical practice guidelines: a multiple methods study using guidelines from oncology as a case. BMJ Open. 2019;9(12):e032483.
236. Armstrong MJ, Mullins CD, Gronseth GS, Gagliardi AR. Impact of patient involvement on clinical practice guideline development: a parallel group study. Implement Sci. 2018;13(1):55.
237. Blackwood J, Armstrong MJ, Schaefer C, Graham ID, Knaapen L, Straus SE, et al. How do guideline developers identify, incorporate and report patient preferences? An international cross-sectional survey. BMC Health Serv Res. 2020;20(1):458.
238. Col NF, Otero D, Lindman BR, Horne A, Levack MM, Ngo L, et al. What matters most to patients with severe aortic stenosis when choosing treatment? Framing the conversation for shared decision making. PLoS One. 2022;17(8):e0270209.
239. Coylewright M, Palmer R, O'Neill ES, Robb JF, Fried TR. Patient-defined goals for the treatment of severe aortic stenosis: a qualitative analysis. Health expectations : an international journal of public participation in health care and health policy. 2016;19(5):1036-43.
240. Clifford AM, Ryan J, Walsh C, McCurtin A. What information is used in treatment decision aids? A systematic review of the types of evidence populating health decision aids. BMC Med Inform Decis Mak. 2017;17(1):22.
241. Rosbach M, Andersen JS. Patient-experienced burden of treatment in patients with multimorbidity - A systematic review of qualitative data. PLoS One. 2017;12(6):e0179916.
242. Eton DT, Elraiyah TA, Yost KJ, Ridgeway JL, Johnson A, Egginton JS, et al. A systematic review of patient-reported measures of burden of treatment in three chronic diseases. Patient Relat Outcome Meas. 2013;4:7-20.
243. Poolman RW, Agoritsas T, Siemieniuk RA, Harris IA, Schipper IB, Mollon B, et al. Low intensity pulsed ultrasound (LIPUS) for bone healing: a clinical practice guideline. BMJ. 2017;356:j576.
244. Siemieniuk RAC, Harris IA, Agoritsas T, Poolman RW, Brignardello-Petersen R, Van de Velde S, et al. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: a clinical practice guideline. BMJ. 2017;357:j1982.
245. World Health O. Advice for the public: Coronavirus disease (COVID-19) 2023 [Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.
246. NICE. COVID-19 rapid guideline: managing COVID-19 2021 [Available from: https://www.nice.org.uk/guidance/ng191/chapter/Recommendations.
247. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet. 2009;374(9683):86-9.
248. Russell LB, Suh DC, Safford MA. Time requirements for diabetes self-management: too much for many? J Fam Pract. 2005;54(1):52-6.

9. Appendix

I have only included the think aloud guides, information, and consent form from one study location. The text from the other study locations is identical except information about local investigators. The forms from the Norwegian study were translated to Norwegian.

9.1 "Think aloud" interview guide

"THINK ALOUD" INTERVIEW GUIDE WITH CLINICIANS

Title of Study: SHARE IT (Sharing Evidence to Inform Treatment decisions): Development of Generic Decision Aids linked to Recommendations in GRADE Guidelines to promote Shared Decision Making in the Clinical Encounter

---

OUTLINE:

Outside of the clinical encounter, investigators will conduct semi-structured interviews of clinicians in the form of "think aloud" sessions, to elicit feedback on presentation formats and usability of the Decision Aids (DA).

These comments will be further analyzed according to eight different facets of "user experience" (findability, usefulness, usability, understandability, credibility, desirability, affiliation & accessibility).

Practically during these sessions, we will provide clinicians the same DA their used in their clinical encounter with the patient. We will ask them to go through its features, thinking out loud to provide us with feedback on issues or problem they faced during the consultation, or are facing presently when looking at the DA.

The investigator will repeat the following instructions at the beginning of the session:

A short bit of repetition before we begin.

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.

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 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 you gave me your opinion..

Then the investigator will invite the clinician to use each feature of the DA (even those that were not used in the clinical encounter) and ask the some or all of the following questions at each relevant step:

- *What is your first reaction*
- *Is the (describe the feature) easily understandable? Explain in your own words what it means.*
- *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?*
- *Would you organize the information you have seen differently?*
- *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?*
- *Any suggestions for improving the user-testing?*

Complementary questions for semi-structured interview of patients

COMPLEMENTARY QUESTIONS FOR SEMI-STRUCTURED INTERVIEW

The doctor discussed some information about your treatment options using a computer tablet.

- **What did you think of this information?**
- **Was it clear, easy to understand?**
- **Did you like the way it was presented?**
- **Tell me about any good/bad points?**
- **What, if anything, could have been shown differently?**
- **How do you feel about your doctor using this tool?**
- **Do you think it helped you and your doctor in your decision today?**

9.2 Clinicians Information and Consent Form

McMaster University Health Sciences | Hamilton Health Sciences | St. Joseph's Healthcare Hamilton

Hamilton Integrated Research Ethics Board

CLINICIANS INFORMATION & CONSENT FORM

Title of Study: SHARE IT (Sharing Evidence to Inform Treatment decisions): Development of Generic Decision Aids linked to Recommendations in GRADE Guidelines to promote Shared Decision Making in the Clinical Encounter

Locally Responsible Investigator: Dr. Gordon Guyatt, Department of Medicine, McMaster University

Principal Investigator: Dr. Thomas Agoritsas, Department of Clinical Epidemiology & Biostatistics, McMaster University.

Co-Investigators: Dr. Fred Spencer, Department of Medicine, McMaster University; Dr. Kari Tikkinen, Dr. Ignacio Neumann, Dr. Yuquing Zhang, Dr. Alonso Carrasco-Labra, Department of Clinical Epidemiology & Biostatistics, McMaster University.

BACKGROUND

Communicating evidence regarding treatment options for shared-decision making is challenging. This process can be facilitated by the use of Decision Aids (DA). However, current DA can be misleading, as they are often not based on the best evidence, or become rapidly outdated. Moreover, their production is time-consuming and they tend to be unwieldy in real clinical practice. Linking the production of DA to current trustworthy guidelines that follow GRADE methodology (Grading of Recommendations Assessment, Development and Evaluation) offers unique opportunities to overcome these limitations.

OBJECTIVE AND METHODS OF THE STUDY

In this SHARE-IT project, we aim to

- To develop a framework for the production of generic Decision Aids (DA) from recommendations in GRADE guidelines.
- To design a set of interactive and adaptable presentation formats to be used in the clinical encounter to display patient relevant outcomes, evidence about benefits and harms of treatment alternatives, and uncertainty of the evidence.
- To test the feasibility of an automated production of DA from electronically published guidelines, and display them on a wide range of devices: computer tablets, desktops, smartphones, as well integrated in electronic medical records.

Through a process of brainstorming and consultation (see attached protocol for detailed description), we will develop a number of preliminary DA for a variety of clinical conditions and treatment alternatives. When ready for pilot use, prototypes displayable on computer tablets

(iPads) will be introduced into the clinical settings for iterative "user-testing" in real patient-clinicians encounters. We will observe the encounter involving the use of the DA, looking for patterns of the conversations and documenting the issues, problems and challenges. These sessions will also be audio-recorded (occasionally video-recorded) and transcribed for further analysis. The encounter will be followed by a short interview of the patient, based on the COMRADE questionnaire. In addition, we will conduct semi-structure interviews with clinicians, in the form of a "think aloud sessions", to elicit feedback on presentation formats and usability of the DA. All this information will contribute to further refinement of the DA and of instructions for their use.

WHAT WILL MY RESPONSIBILITIES BE, IF I TAKE PART IN THE STUDY?

If you volunteer to be part in this study, we will ask you participate to real life user-testing in elective consultations, or pre-specified days of clinical activity. Specifically we will ask you to:

1. Undergo a brief training on how you can use the DA for shared decision making in your practice.
2. Find suitable patients with whom you wish to perform shared decision making and use one of our preliminary DA.
3. Obtain informed consent, using the attached patient information sheet.
4. Have one of the investigators observe and audio-record your encounter with the patient. In specific cases, we may ask you and your patient to video-record the encounter. The investigator will not be present during a physical examination, nor record it.
5. Undergo a semi-structure interview outside the clinical encounter (in a time to be defined). This will take the form of a "think aloud sessions" to get your feedback on the presentation formats and usability of the DA.

WHAT ARE THE RISK & BENEFITS?

There are no known risks to you or your patient from taking part in this research study. The DA is designed to be adaptable to you and your patients needs. You will be free to use it to the extent you see fit (including stopping using it). We will not interfere with the clinical interaction, but will only observe it. You will be able to ask the investigator to leave at anytime.

You and your patient may benefit from the use of the DA during the clinical encounter, if it succeeds in enhancing shared-decision making, by increasing patients understanding of the treatment options and their satisfaction with the decision process.

Your participation is voluntary and you have the right to withdraw your consent at any time, and the investigator will leave.

HOW WILL INFORMATION BE STORED AND PROTECTED

The researcher will not record your name of any other personal information that could identify you or your patient. Audio (or video) recordings will be listened (viewed) only by members of the research team. They will be transcribed in text form and all personal information will be permanently removed from the transcript. All audio (or video) recording will be stored on a secured server at McMaster University and destroyed after 1 year.

If you have any questions about the research now or later, please contact:
Dr. Agoritsas, principal investigator of the study, at 416-890-7778.

CONSENT STATEMENT

Clinician:

I have read the preceding information thoroughly. I have had an opportunity to ask questions and all of my questions have been answered to my satisfaction. I agree to participate in this study. I understand that I will receive a signed copy of this form.

Name	Signature	Date

Investigator obtaining consent:

I have discussed this study in detail with the clinician. I believe the participant understands what is involved in this study.

Name, Role in Study	Signature	Date

This study has been reviewed by the Hamilton Integrated Research Ethics Board (HIREB). The HIREB is responsible for ensuring that participants are informed of the risks associated with the research, and that participants are free to decide if participation is right for them. If you have any questions about your rights as a research participant, please call the Office of the Chair, Hamilton Integrated Research Ethics Board at 905.521.2100 x 42013.

9.3 Participant Information and Consent Form

PARTICIPANT INFORMATION SHEET

Title of Study: Development of decision aids to promote shared decision making in the clinical encounter: the SHARE IT project

Locally Responsible Investigator: Dr. Gordon Guyatt,
Department of Medicine, McMaster University

Principal Investigator: Dr. Thomas Agoritsas, Department of Clinical Epidemiology & Biostatistics, McMaster University.

Co-Investigators: Dr. Fred Spencer, Department of Medicine, McMaster University;
Dr. Kari Tikkinen, Dr. Ignacio Neumann, Dr. Yuquing Zhang, Dr. Alonso Carrasco-Labra, Department of Clinical Epidemiology & Biostatistics, McMaster University.

WHAT IS THE PURPOSE OF THIS STUDY?

Dr. Gordon Guyatt and his colleagues are asking you to participate in a research study that will develop educational materials to help patients and clinicians decide on best treatment. Often, in medicine, decisions about treatment are not straightforward. Patients need to discuss with their doctors the utmost up to date information available to choose the treatment that is best for them. However, understanding this information is sometimes difficult. The purpose of our research is to develop new ways to present available information, that we call a "decision aid", to assist patients and clinicians in making the right decision together.

WHAT WILL HAPPEN IF I AGREE TO TAKE PART IN THE STUDY?

The study team is inviting you to participate in our research study since you are visiting your clinician today, and you may have a discussion concerning the choice of your treatment.

If you agree to participate a researcher at McMaster University will observe your visit today. The researcher is interested in the portion of your appointment when you and your clinician are discussing treatment. Your clinician may use the decision aid, as appropriate to your situation, to discuss your treatment options. The researcher will take notes during your visit and make an audio-recording of your discussion. In specific cases, he or she may ask your permission to video-record the discussion. The researcher will not be present during a physical examination.

After your visit, you will fill in a short questionnaire regarding your discussion of treatment options. The researcher will also ask your feedback about the educational material. This discussion will take about 20 minutes in addition to your visit. You will not receive payment for your participation.

WHAT ARE THE POSSIBLE RISK & BENEFITS FOR ME AND/OR FOR SOCIETY?

There are no known risks to you from taking part in this research study.

We cannot promise any personal benefits to you from your participation. However, the use of the decision aids may help you better understand treatment options when discussing them with your clinicians and help you make the best choice for you. Your participation may help other patients in the future as a result of information gathered in the research study.

WHAT IF I DO NOT WANT TO TAKE PART IN THE STUDY?

Your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time.

Choosing not to participate in this study will in no way affect your care or treatment. You may also refuse to answer any questions you don't want to answer and still remain in the study.

HOW WILL MY INFORMATION BE KEPT PRIVATE?

The researcher will not record your name of any other personal information that could identify you. Audio (or video) recording will be listened (viewed) only by members of the research team. They will be transcribed in text form and all personal information such as your name, address, phone number, family physician's name will be permanently removed from the transcript.

All audio (or video) recording will be stored on a secured server at McMaster University and destroyed after 1 year. You can ask for the destruction of the recordings at any time.

If the results of the study are published, your name will not be used and no information that discloses your identity will be released or published without your specific consent to the disclosure.

QUESTIONS?

If you have any questions about the research now or later, please contact:
Dr. Agoritsas, principal investigator of the study, at 416-890-7778.

CONSENT STATEMENT

Participant:

I have read the preceding information thoroughly. I have had an opportunity to ask questions and all of my questions have been answered to my satisfaction. I agree to participate in this study. I understand that I will receive a signed copy of this form.

Name	Signature	Date

Person obtaining consent:

I have discussed this study in detail with the participant. I believe the participant understands what is involved in this study.

Name, Role in Study	Signature	Date

This study has been reviewed by the Hamilton Integrated Research Ethics Board (HIREB). The HIREB is responsible for ensuring that participants are informed of the risks associated with the research, and that participants are free to decide if participation is right for them. If you have any questions about your rights as a research participant, please call the Office of the Chair, Hamilton Integrated Research Ethics Board at 905.521.2100 x 42013.

10. Articles

Article 1

BMJ 2015;350:g7624 doi: 10.1136/bmj.g7624 (Published 10 February 2015)

ANALYSIS

SPOTLIGHT: PATIENT CENTRED CARE

Decision aids that really promote shared decision making: the pace quickens

Decision aids can help shared decision making, but most have been hard to produce, onerous to update, and are not being used widely. **Thomas Agoritsas and colleagues** explore why and describe a new electronic model that holds promise of being more useful for clinicians and patients to use together at the point of care

Thomas Agoritsas *research fellow*[1 2], Anja Fog Heen *doctoral candidate*[3 4], Linn Brandt *doctoral candidate*[3 4], Pablo Alonso-Coello *associate researcher*[1 5], Annette Kristiansen *doctoral candidate*[3 4], Elie A Akl *associate professor*[1 6], Ignacio Neumann *assistant professor*[1 7], Kari AO Tikkinen *adjunct professor*[1 8], Trudy van der Weijden *professor*[9], Glyn Elwyn *professor*[10], Victor M Montori *professor*[11], Gordon H Guyatt *distinguished professor*[1], Per Olav Vandvik *associate professor*[3 4]

[1]Department of Clinical Epidemiology and Biostatistics, McMaster University, Faculty of Health Sciences, Hamilton, Ontario, Canada; [2]Division of General Internal Medicine, Division of Clinical Epidemiology, University Hospitals of Geneva, Switzerland; [3]Department of Medicine, Innlandet Hospital Trust, Gjøvik, Norway; [4]Institute for Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway; [5]Iberoamerican Cochrane Centre, Biomedical Research Institute Sant Pau—CIBER, Epidemiología y Salud Pública, Barcelona, Spain; [6]Department of Internal Medicine, American University of Beirut, Lebanon; [7]Department of Internal Medicine, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile; [8]Departments of Urology and Public Health, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland; [9]Department Family Medicine, School for Public Health and Primary Care, Maastricht University, Maastricht, Netherlands; [10]Dartmouth Center for Health Care Delivery Science, Dartmouth Institute for Health Policy and Clinical Practice, Hanover, USA; [11]Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, USA

Many, perhaps most, important decisions in medicine are not clear cut.[1 2] Patients and clinicians need to discuss the options using the best available evidence and make informed joint decisions that take account of patients' context, values, and preferences.[3 4] But implementing shared decision making is not easy. Doctors need the skills and tools to do it and to build trust; patients need information and support. Patients also need to have a greater role in developing strategies to improve the process.[5 6]

Access to best evidence is another key ingredient. Until now the production and dissemination of clinical practice guidelines and summaries of evidence has largely been tailored to meet the educational needs of clinicians. They are seldom provided in a format that supports shared decision making.[7] Patients meanwhile, struggle to find reliable and accessible summaries of evidence, although plain language summaries and patient versions of guidelines are being developed.[8]

In this article we highlight the limitations of current decision aids and discuss how the generic production of electronic decision aids designed for use in the clinical encounter, linked directly to trustworthy summaries of evidence from systematic reviews and guidelines, may help in the long march to realising effective shared decision making.

Challenge of shared decision making

Shared decision making depends on a good conversation[9] in which clinicians share information about the benefits, harms, and burden of alternative diagnostic and therapeutic options and patients explain what matters to them and their views on the choices they face.[4 10] It should follow the principles of patient centred care, promote informed choice, and result in care that patients value.[1-11] Many clinicians think they practice shared decision making, but evidence suggest a perception-reality gap[3] because of misconceptions about the nature of shared decision making, the skills it requires, the time it takes, and the degree to which patients, families, and carers wish to share in decision making.[12-14]

Each clinical encounter is influenced by many factors. These include patients' circumstances and medical needs as well as

Correspondence to: T Agoritsas, thomas.agoritsas@gmail.com

their beliefs, stemming from what they have read, personal experience, advice from family and friends, and the media. It is therefore important to provide patients with accurate, up to date evidence on the benefits and harms of alternative management strategies and their likely effect on outcomes that matter to them, although evidence may not always reflect the complexity and multimorbidity of individual patients and patients may choose to ignore the evidence**.** Good shared decision making requires clinicians to have access to detailed knowledge and ideally summaries of the latest evidence and the means to share it in a way that supports thoughtful deliberation, something that cannot be done on the fly.

Limitations of traditional decision aids

For the past two decades enthusiasts have advocated decision aids to facilitate shared decision making, and over 500 have been developed.[15 16] A systematic review of 115 randomised trials showed that their use was associated with a 13% absolute increase in patients' knowledge scores and an 82% relative increase in accurate expectations of possible benefits and harms. Effects on clinical outcomes, adherence to treatment, and use of services have not, however, been consistent.[15 17]

Most decision aids have been designed for patients to use independently outside the consultation, either in the waiting room or at home.[10] Although these decision aids promote understanding of the issues, they cannot guarantee that decisions in the consultation are shared,[3 18] and there is insufficient evidence to determine how their use influences the consultation.[18] Another problem is that use of decision aids in routine care is low,[13] mainly because of poor design and lack of ready access to them. Furthermore clinicians may find the format impractical to use in consultations and may be as unfamiliar as their patients with risk estimates and the inherent uncertainty associated with probabilities.[19]

Traditional decision aids are often not based on current evidence or rapidly outdated, at least in part because of limitations in funding after tool development—and may thus do more harm than good.[20] A rigorous systematic review is needed for each important outcome, and such reviews are often unavailable. A recent assessment found that although around two thirds of decision aids are based on systematic reviews or guidelines, many of these sources are of questionable quality, and only 5% of aids included an "expiry date" or a stated policy about updating.[20]

Ensuring the quality and timeliness of decision aids is a daunting challenge. The work required to summarise evidence for a trustworthy decision aid is similar to that for producing a systematic review or a guideline, suggesting the potential for synergy between the worlds of evidence based practice and shared decision making.[20-22]

Harnessing the potential of recent developments

New decision aids

Some newer decision aids have been designed to facilitate collaborative deliberation in the course of the clinical encounter.[3 10] Montori and colleagues pioneered a user centred approach to producing decision aids through iterative observations of discussions between doctors and patients.[9 23] Their approach resulted in succinct, easy to use tools that provide graphic displays of the benefits and harms of different options organised around concerns that are important to patients (http://shareddecisions.mayoclinic.org). In contrast to traditional aids, which patients use independently, they are not designed to be comprehensive and do not include explicit exercises to help patients clarify their values (such as the relative values of avoiding a stroke versus a gastrointestinal bleed) [24] Instead they rely on the unique conversations that take place between patients and clinicians, with clinicians providing just in time, tailored explanations and information.[10] Direct observations in randomised trials have shown that these short tools (so far available for diabetes, statins, and antidepressants) promote dialogue and increase joint deliberation.[25] They also shift the "body language" as patients and clinicians sit together to review the data.[23 26]

Other short point of care decision aids include Option Grids (www.optiongrid.co.uk).[27 28] These are one page summaries that provide answers to patients' frequently asked questions, covering clinical outcomes and practical concerns faced in daily life. Their value in routine care is being evaluated.[27]

Developments in appraisal and presentation of best evidence

The GRADE approach (Grading of Recommendations Assessment, Development and Evaluation) provides systematic, transparent, and explicit guidance for processing evidence from the medical literature, and has been widely adopted.[7-30] Use of the GRADE approach results in standardised and succinct evidence profiles or summary of findings tables, which specify the absolute effects of an intervention on outcomes important to patients rather than surrogate outcomes and provide a rating of the certainty in these estimates (high, moderate, low, or very low).[30] The recent international patient decision aids standards have emphasised the potential of GRADE for the production of decision aids[20], and it has been adopted by over 80 organisations (www.gradeworkinggroup.org).

Furthermore, clinical practice guidelines using GRADE now issue weak recommendations (in contrast to strong) when there is a close balance between desirable and undesirable outcomes among alternatives, low certainty in estimates of effect, or when there is large variability in patients' values and preferences. Weak recommendations, which dominate in recent high quality guidelines,[2] thus identify decisions where shared decision making is particularly important.[20 22]

Use of new technologies

The not-for-profit MAGIC project (Making GRADE the Irresistible Choice www.magicproject.org) has developed an online "app" with potential to produce electronic decision aids for use in the clinical encounter.[7] This MAGICapp (www.magicapp.org) allows authors of guidelines or systematic reviewers to write evidence summaries into a structured database and appraise them using GRADE criteria. The content can then be published on a web platform and presented in interactive formats on tablets, web portals, or electronic medical record systems.[31]

In the SHARE-IT project, we use this authoring and publication platform for the generic and semi-automated production of a large number of decision aids.[7] The aids can be used with the corresponding systematic review or clinical practice guidelines and the format modified and tailored to specific contexts—for example, published in different languages or adapted to national guidelines.[32 33] The electronic format facilitates continuous updating because the data in the decision aids will change automatically each time the underlying review is modified.[7]

Figure 1⇓ summarises the methods of the SHARE-IT project. In collaboration with DECIDE (www.decide-collaboration.eu),[34]

we gathered an international team of experts in evidence based medicine and shared decision making, clinicians, guideline developers, and designers, and developed an initial framework and electronic prototype for the translation of GRADE summaries into decision aids. We then applied an iterative and user centred design, directly involving patients and clinicians facing real decisions. We built 10 decision aids on antithrombotic drugs and modified the generic prototype in light of observations of their use in practice and individual feedback from patients and clinicians.

The video illustrates how the prototype uses interactive formats to present evidence summaries at varying levels of detail. The prototype shows that the approach is feasible, and preliminary experience suggests it is appreciated by both patients and clinicians (box). Across 16 clinical encounters, patients consistently reported high levels of satisfaction with the prototype in understanding risks and benefits and in enhancing their confidence in decisions (mean scores of 88.7 and 90.9 respectively (maximum 100) as assessed by COMRADE.[35]

Conclusion

No decision aid is sufficient to guarantee that clinical decision making is shared. Undergraduate, postgraduate, and continuing education programmes must teach health professionals about the importance of creating and fostering a culture of shared decision making and the skills needed to communicate evidence, and its limitations, in a way people can understand. Furthermore, the challenge of producing evidence summaries that deal optimally with complexity, multimorbidity, and potentially limited applicability to the patient remains.[36]

We are, however, now in a position to construct, test, and refine electronic evidence summaries for use in the clinical encounter for a wide variety of patient groups and clinical settings. Our prototype, built in the MAGICapp, demonstrates the feasibility of semiautomated production of decision aids from a large number of electronically published evidence summaries. We also plan to implement these formats in another similar platform, the GRADEpro Guideline Development Tool (www.guidelinedevelopment.org). We invite patient organisations, research groups, guideline developers, patients, and clinicians to partner with us (www.magicproject.org) and help us advance the science and art of truly shared and well informed decision making.

We thank Frankie Achille (interaction designer), Rob Fracisco (designer), and Deno Vichas and Chris Degiere (programmers) for their contributions in development of the online authoring and publication platform prototype (www.magicproject.org). TA was financially supported by a fellowship for prospective researchers grant No P3SMP3-155290/1 from the Swiss National Science Foundation, as well as by a fellowship grant from the University Hospitals of Geneva and from Eugenio Litta—Fondation Genevoise de Bienfaisance Valeria Rossi di Montelera. PA-C is funded by a Miguel Servet research contract from the Instituto de Salud Carlos III (CP09/00137). KAOT is funded by the Academy of Finland (#276046), Jane and Aatos Erkko Foundation, and Sigrid Jusélius Foundation. The Innlandet Hospital Trust, South-Eastern Norway Regional Health Authority and Innovation Norway have provided research grants for the MAGIC program (www.magicproject.org). This project has received funding from the European Union's Seventh Framework Programme for research, technological development and dissemination under grant agreement No 258583. (www.decide-collaboration.eu)

Contributors and sources: The SHARE-IT project was conceived and is mainly funded by the MAGIC program, in close collaboration with the DECIDE project and GRADE working group, to which most contributors are affiliated. We also received numerous feedbacks from stakeholders at international meetings. TA led and coordinated the project, supervised by GHG and POV. TA, AFH, LB, and POV developed and implemented the prototype, and all contributors provided feedback at different stages. TA, AFH, and POV performed user-testing in clinical encounters. TA drafted the manuscript and all authors critically revised the manuscript. TA is guarantor.

Competing interests: All authors have read and understood BMJ policy on declaration of interests and declare the following interests: TA, AFH, LB, AK, PAC, EAA, IN, KAOT, VMM, GHG, POV are members of the GRADE working group (www.gradeworkinggroup.org), as well as coinvestigators in the DECIDE project (www.decide-collaboration.eu). TA, AFH, LB, AK, GHG, POV are members of the MAGIC research and innovation program. GE leads the Option Grid collaborative. VMM designs and tests shared decision making tools at the KER UNIT in Mayo Clinic. These tools are then made available for free with no income generated for him, his unit, or his institution.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. *N Engl J Med* 2012;366:780-1.
2 Djulbegovic B, Guyatt GH. Evidence-based practice is not synonymous with delivery of uniform health care. *JAMA* 2014;312:1293-4.
3 Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. *BMJ* 2012;344:e256.
4 Elwyn G, Lloyd A, May C, van der Weijden T, Stiggelbout A, Edwards A, et al. Collaborative deliberation: a model for patient care. *Patient Educ Counsel* 2014;97:158-64.
5 Legare F, Ratte S, Stacey D, Kryworuchko J, Gravel K, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2010;5:CD006732.
6 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ* 2014;348:g3725.
7 Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. *Chest* 2013;144:381-9.
8 Boivin A, Currie K, Fervers B, Gracia J, James M, Marshall C, et al. Patient and public involvement in clinical guidelines: international experiences and future perspectives. *Qual Safe Health Care* 2010;19:e22.
9 Montori VM, Breslin M, Maleska M, Weymiller AJ. Creating a conversation: insights from the development of a decision aid. *PLoS Med* 2007;4:e233.
10 Elwyn G, Frosch D, Volandes AE, Edwards A, Montori VM. Investing in deliberation: a definition and classification of decision support interventions for people facing difficult health decisions. *Med Decis Making* 2010;30:701-11.
11 Dwamena F, Holmes-Rovner M, Gaulden CM, Jorgenson S, Sadigh G, Sikorskii A, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev* 2012;12:CD003267.
12 Hoffmann TC, Legare F, Simmons MB, McNamara K, McCaffery K, Trevena LJ, et al. Shared decision making: what do clinicians need to know and why should they bother? *Med J Aust* 2014;201:35-9.
13 Elwyn G, Legare F, van der Weijden T, Edwards A, May C. Arduous implementation: does the Normalisation Process Model explain why it's so difficult to embed decision support technologies for patients in routine clinical practice. *Implement Sci* 2008;3:57.
14 Cox K, Britten N, Hooper R, White P. Patients' involvement in decisions about medicines: GPs' perceptions of their preferences. *Br J Gen Pract* 2007;57:777-84.
15 Stacey D, Legare F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2014;1:CD001431.
16 Patient Decision Aids Inventory. Ottawa Hospital Research Institute and University of Ottawa. http://decisionaid.ohri.ca/index.html.
17 Walsh T, Barr PJ, Thompson R, Ozanne E, O'Neill C, Elwyn G. Undetermined impact of patient decision support interventions on healthcare costs and savings: systematic review. *BMJ* 2014;348:g188.
18 Hargraves I, Montori VM. Decision aids, empowerment, and shared decision making. *BMJ* 2014;349:g5811.
19 Perneger TV, Agoritsas T. Doctors and patients' susceptibility to framing bias: a randomized trial. *J Gen Intern Med* 2011;26:1411-7.
20 Montori VM, LeBlanc A, Buchholz A, Stilwell DL, Tsapas A. Basing information on comprehensive, critically appraised, and up-to-date syntheses of the scientific evidence: a quality dimension of the international patient decision aid standards. *BMC Med Inform Decis Making* 2013;13 Suppl 2:S5.
21 Hoffmann TC, Montori VM, Del Mar C. The connection between evidence-based medicine and shared decision making. *JAMA* 2014;312:1295-6.
22 Van der Weijden T, Boivin A, Burgers J, Schunemann HJ, Elwyn G. Clinical practice guidelines and patient decision aids. An inevitable relationship. *J Clin Epidemiol* 2012;65:584-9.
23 Breslin M, Mullan RJ, Montori VM. The design of a decision aid about diabetes medications for use during the consultation with patients with type 2 diabetes. *Patient Educ Counsel* 2008;73:465-72.
24 Llewellyn-Thomas HA, Crump RT. Decision support for patients: values clarification and preference elicitation. *Med Care Res Rev* 2013;70(1 suppl):50-79S.
25 Coylewright M, Branda M, Inselman JW, Shah N, Hess E, LeBlanc A, et al. Impact of sociodemographic patient characteristics on the efficacy of decision aids: a patient-level meta-analysis of 7 randomized trials. *Circ Cardiovasc Qual Outcomes* 2014;7:360-7.

Reaction to the decision aid

A haematologist expressed surprise that one decision aid regarding long term anticoagulation treatment for patients with unprovoked venous thromboembolism begins by inviting patients to choose which outcome to discuss first. She usually started by discussing the risk of recurrence, then bleeding before inviting patients' questions, omitting mortality.

After we clarified she could use the tool as she wanted, she began with the six month follow-up of a 47 year old man taking rivaroxaban for an unprovoked pulmonary embolism. She explained that, although the treatment was indicated after the acute event, the decision to continue rivaroxaban depended on his preferences. She accessed the decision aid and moved to sit next to the patient. Revising her prior plan to use her accustomed order, she used the trigger sentence offered: "What aspect of your medication would you like to discuss first?" The patient chose "practical consequences." In the conversation that followed, they further discussed risk of bleeding, recurrence, and associated mortality. The patient decided to discontinue rivaroxaban.

After the encounter, the clinician pointed out that the patient focused on practical consequences first, and she reflected on how the tool resulted in positive changes to her usual communication strategy. The patient reported that the decision aid made it easier to "digest the information and get the bigger picture." He explained he was first interested by "day-to-day stuff" before exploring "more intimidating" but important issues.

26 Wyatt KD, Branda ME, Anderson RT, Pencille LJ, Montori VM, Hess EP, et al. Peering into the black box: a meta-analysis of how clinicians use decision aids during clinical encounters. *Implement Sci* 2014;9:26.
27 Lloyd A, Joseph-Williams N, Edwards A, Rix A, Elwyn G. Patchy 'coherence': using normalization process theory to evaluate a multi-faceted shared decision making implementation program (MAGIC). Implement Sci 2013;8:102.
28 Giguere A, Labrecque M, Haynes R, Grad R, Pluye P, Legare F, et al. Evidence summaries (decision boxes) to prepare clinicians for shared decision-making with patients: a mixed methods implementation study. *Implement Sci* 2014;9:144.
29 Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6.
30 Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383-94.
31 Kristiansen A, Brandt L, Alonso-Coello P, Agoritsas T, Akl EA, Conboy T, et al. Development of a novel multilayered presentation format for clinical practice guidelines. *Chest* 2014 Oct 9 [Epub ahead of print].
32 Kristiansen A, Brandt L, Agoritsas T, Akl EA, Berge E, Bondi J, et al. Adaptation of trustworthy guidelines developed using the GRADE methodology: a novel 5-step process. *Chest* 2014;146:727-34.
33 Kristiansen A, Brandt L, Agoritsas T, Akl EA, Berge E, Jacobsen AF, et al. Applying new strategies for the national adaptation, updating and dissemination of trustworthy guidelines: results from the Norwegian adaptation of the American College of Chest Physicians evidence-based guidelines on antithrombotic therapy and the prevention of thrombosis, 9th wdition. *Chest* 2014;146:735-61.
34 Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brozek J, et al. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. Implement Sci 2013;8:6.
35 Edwards A, Elwyn G, Hood K, Robling M, Atwell C, Holmes-Rovner M, et al. The development of COMRADE—a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. *Patient Educ Counsel* 2003;50:311-22.
36 Uhlig K, Leff B, Kent D, Dy S, Brunnhuber K, Burgers JS, et al. A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med* 2014;29:670-9.

Cite this as: *BMJ* 2015;350:g7624

Figure

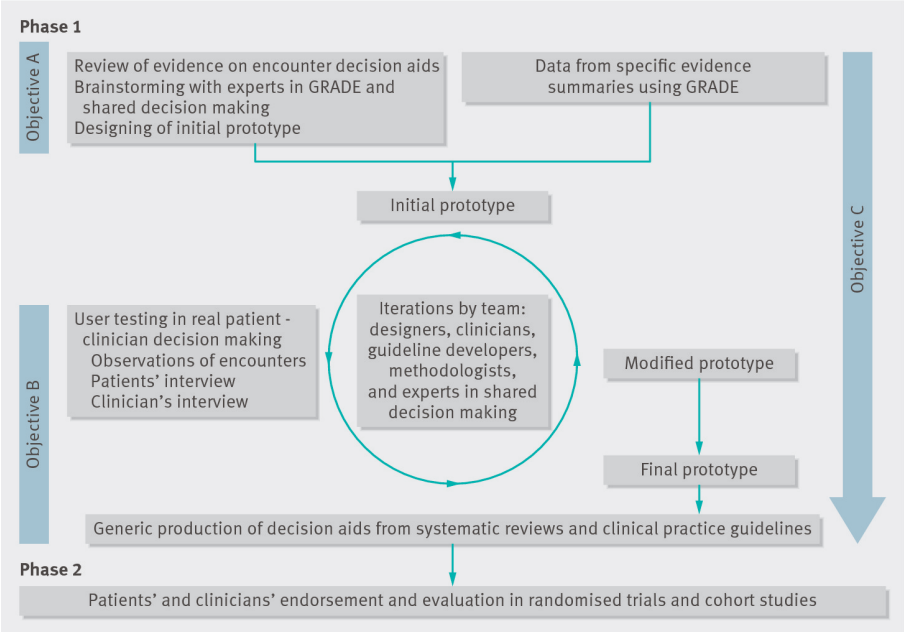


Fig 1 Outline of the methods and user-centred approach in the SHARE-IT project. Objective A=to develop a framework for the generic translation of GRADE evidence summaries into decision aids; Objective B=to design a set of interactive presentation formats for use in the clinical encounter; Objective C=to test the feasibility of an automated production of these decision aids from electronically published evidence summaries. Subsequent phases of the project involve the generic production of decision aids from real practice guidelines and their evaluation in randomised trials and cohort studies

Article 2

Heen *et al. BMC Med Inform Decis Mak* (2021) 21:202
https://doi.org/10.1186/s12911-021-01541-7

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

Decision aids linked to evidence summaries and clinical practice guidelines: results from user-testing in clinical encounters

Anja Fog Heen[1,2*], Per Olav Vandvik[2,3], Linn Brandt[1,2], Frankie Achille[2], Gordon Henry Guyatt[2,4], Elie A. Akl[5], Shaun Treewek[6] and Thomas Agoritsas[2,7]

Abstract

Background: Tools for shared decision-making (e.g. decision aids) are intended to support health care professionals and patients engaged in clinical encounters involving shared decision-making. However, decision aids are hard to produce, and onerous to update. Consequently, they often do not reflect best current evidence, and show limited uptake in practice. In response, we initiated the Sharing Evidence to Inform Treatment decisions (SHARE-IT) project. Our goal was to develop and refine a new generation of decision aids that are generically produced along digitally structured guidelines and evidence summaries.

Methods: Applying principles of human-centred design and following the International Patient Decision Aid Standards (IPDAS) and GRADE methods for trustworthy evidence summaries we developed a decision aid prototype in collaboration with the Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence project (DECIDE). We iteratively user-tested the prototype in clinical consultations between clinicians and patients. Semi-structured interviews of participating clinicians and patients were conducted. Qualitative content analysis of both user-testing sessions and interviews was performed and results categorized according to a revised Morville's framework of user-experience. We made it possible to produce, publish and use these decision aids in an electronic guideline authoring and publication platform (MAGICapp).

Results: Direct observations and analysis of user-testing of 28 clinical consultations between physicians and patients informed four major iterations that addressed readability, understandability, usability and ways to cope with information overload. Participants reported that the tool supported natural flow of the conversation and induced a positive shift in consultation habits towards shared decision-making. We integrated the functionality of SHARE-IT decision aids in MAGICapp, which has since generated numerous decision aids.

Conclusion: Our study provides a proof of concept that encounter decision aids can be generically produced from GRADE evidence summaries and clinical guidelines. Online authoring and publication platforms can help scale up production including continuous updating of electronic encounter decision aids, fully integrated with evidence summaries and clinical practice guidelines.

Keywords: Decision aids, Shared decision-making, Clinical practice guidelines

Background

Most medical decisions are highly context-dependant, and, when creating individual plans of care, current best evidence of potential benefits and harms requires

*Correspondence: anjaheen@gmail.com
[1] Department of Medicine, Innlandet Hospital Trust, Mailbox 104, 2381 Brumunddal, Norway
Full list of author information is available at the end of the article

interpretation in light of patients' values and preferences. Shared decision-making is the process in which patients and clinicians partner together and have a conversation to find the best option for that patient [1]. Communicating evidence for shared decision-making is challenging [2]. Trustworthy clinical practice guidelines (henceforth guidelines) are amongst the most reliable methods of translating evidence into statements to guide practice, but are typically not designed to support shared decision-making. Decision aids represent widely advocated tools for shared decision-making [3]. Decision aids improve patients' knowledge of options, their perception of feeling well-informed, and their clarity regarding what matters most to them [3].

Both guidelines and decision aids face similar challenges: their production and updating is highly resource-demanding, they are often not based on best available evidence, they may be hard to find and use, and their uptake is highly variable in practice [4]. We have previously reported how we have addressed these overarching challenges in the Sharing Evidence to Inform Treatment decisions (SHARE-IT) project [4]. SHARE-IT has resulted in a new generation of generic decision aids linked to trustworthy guidelines and evidence summaries in digitally structured formats [4, 5]. These encounter decision aids are designed to be used by clinicians and patients to explore together the management options and facilitate shared decision-making [4].

We report here our detailed approach to SHARE-IT encounter decision aids conceptual and technical development, and results from iterative user-testing to achieve user-friendly presentation formats. We also report how these encounter decision aids were integrated in MAGICapp, a digital authoring and publication platform for guidelines and evidence summaries. In MAGICapp, the evidence data is structured in a way that enables a semi-automated production of decision aids, and facilitate dissemination and dynamic updating of them, within the context of guidelines [4].

Methods

Overview and rationale

SHARE-IT was initiated in 2012 by the non-profit MAGIC Evidence Ecosystem Foundation [6]. Combining research with innovation and product development within a digital and trustworthy evidence ecosystem, MAGIC aims to provide clinicians and patients with user-friendly tools for decision support implemented at the point of care [4, 5]. Its online authoring and publication platform—the MAGICapp (Fig. 1)—was initially developed to apply GRADE methodology (Grading of Recommendations Assessment, Development and Evaluation) [7] to author, publish and dynamically update trustworthy guidelines in user-friendly formats [5]. We quickly identified the need to translate digitally structured data into tools that could support shared decision-making in the clinical encounter.

We conceived SHARE-IT in collaboration with the DECIDE project (Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence), a multi-national research project initiated by the GRADE working group and funded by the European Union [8–11]. After the DECIDE project ended in 2014, our team continued user-testing and developing the decision aids. A major consequent refinement was the addition of a display of practical issues to complement evidence on benefits and harms [12, 13].

Based on initial feedback from experts and stakeholders, principles of human-centred design were applied, and led to iterative revisions of the encounter decision aids through repeated observations with patients and clinicians engaged in real-life decision-making [14]. Figure 2 shows the three phases of our project, as defined in DECIDE: (1) brainstorming and stakeholder feedback with a multidisciplinary team to develop a conceptual framework and prototype decision aid; (2) iterative development and user testing of the decision aids; (3) their generic semi-automated production, from GRADE evidence summaries linked to guidelines, in MAGICapp.

Development of the decision aids

Sketching the initial template

We based our initial prototype on evidence regarding optimal formats for shared decision-making, with a particular focus on encounter decision aids. In particular, our template was inspired by decision aid cards centred on key outcomes and issues meaningful to patients pioneered by Dr. Montori and his team in the Mayo Clinic Knowledge and Evaluation Research Unit [15].

Our team combining expertise in GRADE methodology, shared decision-making and human-centered design, built several prototypes with the help of an interaction designer (FA). We followed a modified "mobile first" approach [16] in sketching and creating the initial template, using an online calculator [17] and Blueprint software [18] which allowed us to quickly customize and test our prototypes on tablet screens. We judged use of the tool on a desktop computer would not optimally facilitate face to face communication between patient and clinician in clinical encounters.

Stakeholder feedback and brainstorming on the next iterations

To move from the initial template to a conceptual framework and prototype decision aids linked to guidelines, we

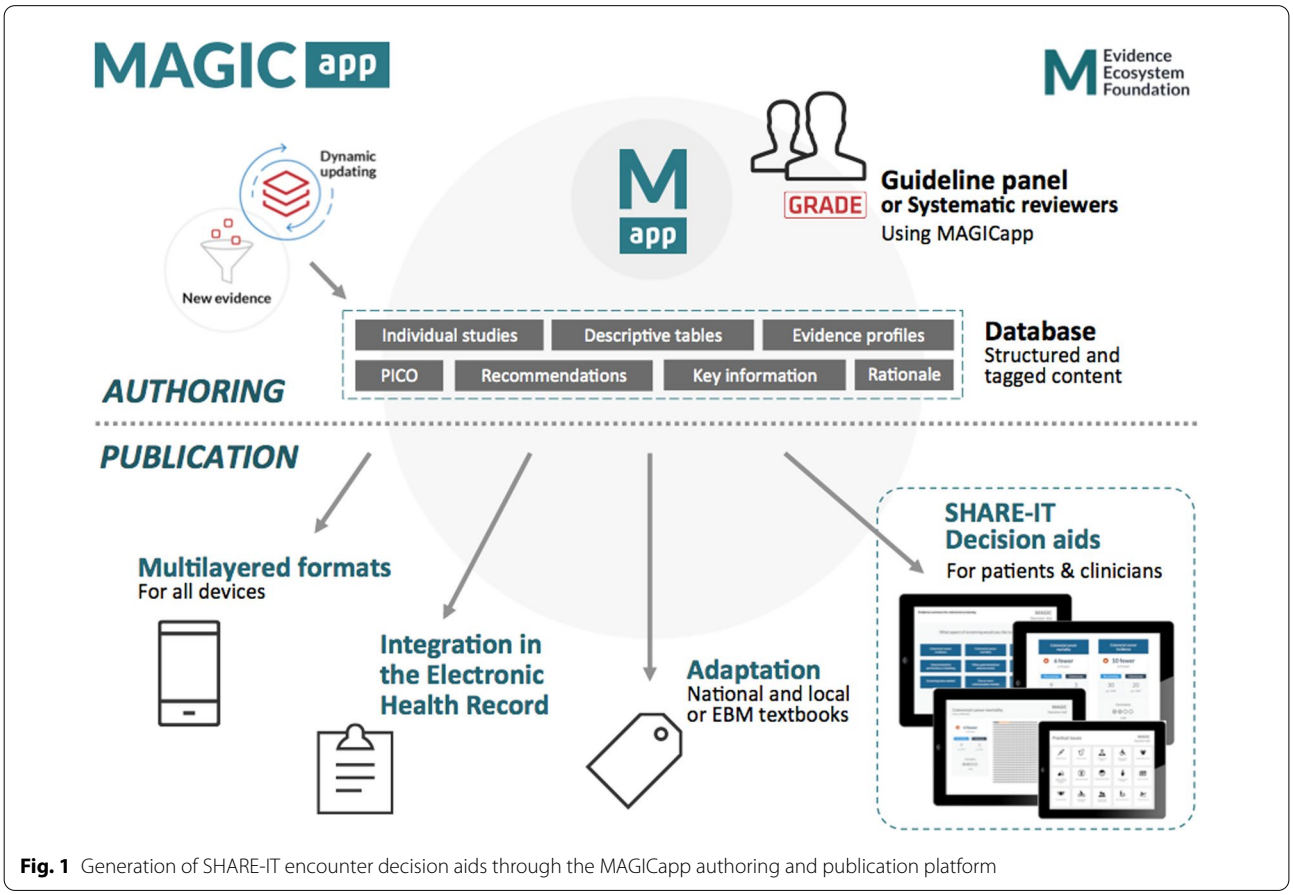


Fig. 1 Generation of SHARE-IT encounter decision aids through the MAGICapp authoring and publication platform

conducted three face-to-face meetings with stakeholders in DECIDE (Canada 2012, Italy 2013, and Peru 2013) [8]. The meetings involved clinicians and experts in shared decision-making, guideline development and designers. The experts evaluated the initial template and subsequent prototype decision aids and participated in brainstorming, discussion and feedback.

User testing

Following stakeholder feedback, the team prepared the prototype for formal user testing in clinical encounters to learn about the design from a user's perspective to improve its next iteration as opposed to developers or experts [19].

Materials and setting

Prototype encounter decision aids were built for a variety of clinical scenarios, including 21 decisions concerning antithrombotic therapy and one for cancer treatment [20, 21]. The choice of supporting evidence summaries was driven by the fact that several authors had conducted extensive GRADE evidence summaries related to an update of the American College of Chest Physicians Evidence-Based Clinical Practice Guidelines on the topics [20]. Antithrombotic therapy decisions addressed new oral anticoagulants (for pulmonary embolisms, deep vein thrombosis and atrial fibrillation) and thromboprophylaxis during pregnancy. We used GRADE evidence summaries published in digitally structured formats in MAGICapp [4, 5]. The cancer scenario addressed adjuvant tamoxifen treatment to prevent recurrence of breast cancer; we produced a GRADE evidence summary based on trial results [15]. All decision aids reflected decisions deemed particularly sensitive to patient values and preferences, typically accompanying weak recommendations according to the GRADE framework [22]. The decision aids were available in English and Norwegian.

The completed Standards for Reporting Qualitative Research checklist is included as Additional file 1.

Participants and recruitment

We performed user-testing of the decision aids in real-life consultations in secondary and tertiary health care facilities in Norway (Innlandet Hospital Trust, Gjøvik and Oslo University Hospital, Oslo), the United Kingdom (Ninewells Hospital, Dundee) and Canada

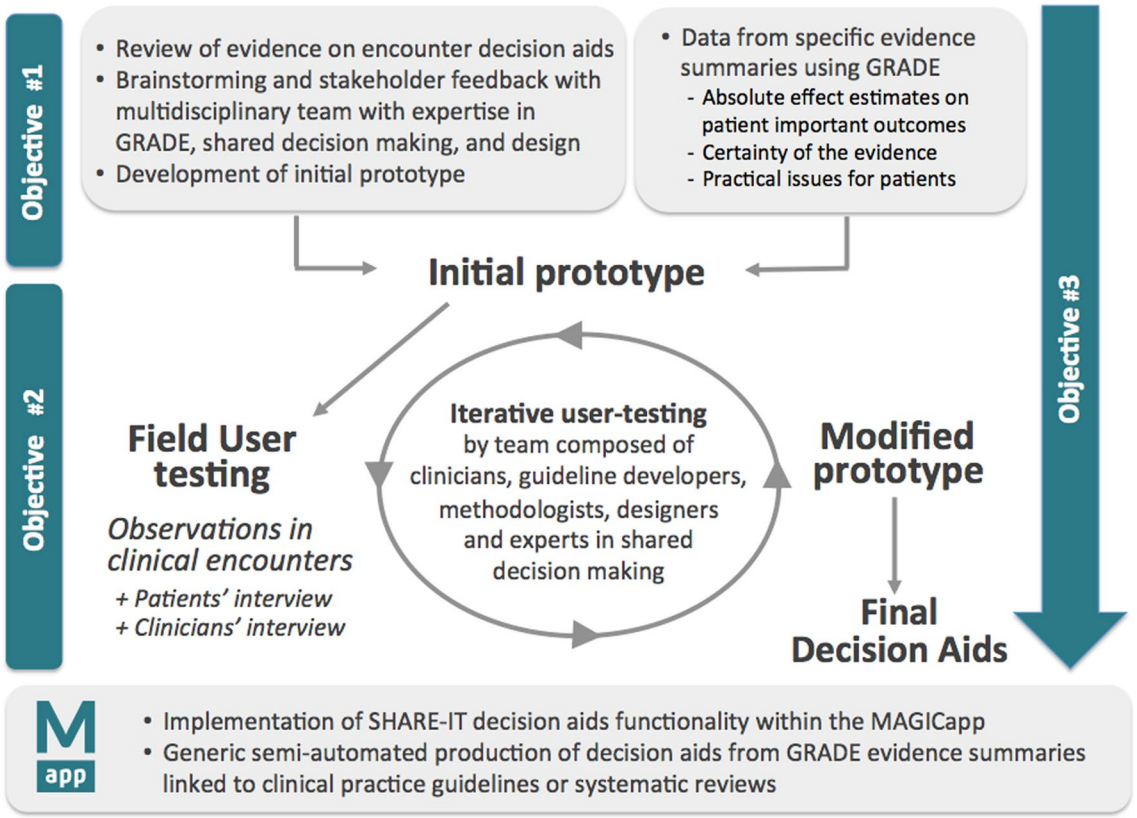


Fig. 2 Prototyping, User-testing and Implementation of SHARE-IT decision aids in the MAGICapp for their generic production

(McMaster University Hospital and Hamilton General hospital, Ontario). A convenience sample of physicians was recruited, with variable experience in risk communication and variable familiarity of the clinical topic. Patients were recruited through the participating physicians as part of either their outpatient clinic visits or acute hospital inpatient admissions.

Data collection

A team member provided a brief demonstration of the tool, typically less than 10 min, demonstrating to participating physicians the use of the encounter decision aid. A study member directly observed the clinical encounter, noting the use of the decision aids, and patients' questions regarding their management. We audio-recorded and transcribed the consultations, followed by professional translation to English for encounters in Norwegian.

Directly after the consultation, the team member who had observed the encounter conducted separate think-aloud sessions with patients and clinicians. We used a semi-structured interview guide with questions eliciting feedback on their experience and on the format and usability on the decision aid. The focus of our attention was their actual experience. Suggestions for improvement were also collected. At the end of the interview respondents completed the 20-item COMRADE instrument, which provides a quantitative assessment of *risk communication* and *confidence in the decision* [23]. COMRADE uses a 5-point scale from 1(strongly agree) to 5 (strongly disagree) [24].

Data analysis

We coded transcriptions of the audio-recordings of the clinical encounters and semi-structured interviews. Content analysis was performed through both deductive and inductive approaches, searching for units of meaning and condensing text [25]. We then compared and added codes to the results and searched for barriers, problems and facilitating elements or characteristics of the tool that influenced the user experience and the process of shared decision-making. Each element of meaning was coded using a revised version of Morville's framework (Fig. 3) categorizing eight different facets of "user experience" to sort results into categories: findability, usefulness, usability, understandability, credibility, desirability, identification and accessibility [26, 27]. Finally, each element was also coded with regards to the quality of the

Fig. 3 Modified Morville's model for testing the experience of users

experience—i.e., positive feedback, neutral experience, suggestions for improvement of the tool, minor frustration and major frustration ("show stoppers").

Results

Development of framework and prototype encounter decision aids

In the three DECIDE stakeholder meetings, 22 experts provided extensive feedback and suggestions to inform the conceptual framework and prototype decision aid formats. Core desirable features of the decision aids included: (1) communicating risk and uncertainty, (2) navigating the content, (3) facilitating use of the encounter decision aids both within and outside the clinical encounter, and (4) the inclusion of burden of treatment/practical issues. Following several iterations, the experts reached consensus on a prototype decision aid ready to undergo user-testing (Figs. 4, 5).

Iterative development through user testing

We performed four major iterations of the decision aid presentation formats, based on the observations and analyses of 28 real-life consultations with physicians and patients (median age 53, range 19–90, 64% women). Participants used tablet computers (e.g., iPads) in 47% of the consultations, desktop or laptop computers in the remainder. COMRADE response rate was 72.7% (n = 20). Patients rated both items of risk communication and items in their confidence in the decision with a median of 1 on the 5-point scale (i.e. "*strongly agree*").

Overview of user-experiences

Table 1 provides a quantitative summary of user-experiences with the encounter decision aids categorized according to the revised Morville's facets (Fig. 3) and the quality of the experience coded as: positive feedback, neutral experience, suggestions for improvement, minor frustration and major frustration. These were based on content analysis of transcripts of the consultations and semi-structured interviews. Elements of major or minor frustration, with or without suggestions for improvement, were the main drive for improvement of the tool across iterations, as they affected most the user experience. Neutral experience referred to statements voiced by users, which were neither positive nor negative, that provided insight on how they navigated across the different features or functionalities of the tool. Together with spontaneous positive feedback, they pointed at functionalities of the tools that worked smoothly in the course of the clinical encounter.

We coded 586 observed units of meaning across all interactions. Most reported issues involved understandability and usefulness, whereas findability and credibility aspects were least reported. Regarding the quality of the experience, there were no showstoppers. The majority of observations (43%) related to ways to use the tool in consultations, while 32% were expressions of positive feedback (e.g. praise, elements of delighted surprise), and 12% suggestions for improvement. We provide below a synthesis of the findings in each of the facets, with illustrative quotes directly from the consultations and interviews.

Accessibility

Across iterations, the majority of comments concerned the readability, font colours or size, or visual contrast (e.g. they needed to put on their glasses) with other expressions of aesthetic preferences. Patients perceived the tools worked well for themselves, but speculated on how it may not be as accessible for others, such as colour-blind people, or older patients who may be more averse to technology:

> *57-year-old man with venous pulmonary embolism: "If you were using this tool with other people, other than me, just people 65/70 years old and afraid of the new technology, the picture would be a little blurred."*
> *46-year-old woman with breast cancer: "Being faced with an iPad or a laptop may put off some older*

Fig. 4 Example of encounters decision aids; **a** first layer displaying outcomes and practical issues relevant to a given decision; **b** underlying layer for the exploration of practical issues

women"

Usability

The majority of users, both clinicians and patients, reported that the tool was easy and simple to use without need for explanation, with a design that supported usability.

> *Clinician: "Actually, it's quite self-explanatory, really, the whole app."*
> *Clinician: "Everything was presented in a very neutral way. That is, no scary fonts, no green or red colours that might imply certain values. I felt everything was easy to read and interpret"*

Physicians integrated the tool in their work-flow and conversations using expressions such as "let's go back and see", pointing at outcomes on the screen, asking what the patient wanted to look at first or leaving the direction of the conversation to the patient. Several did this together with the patient, describing the numbers, using the tablet together and the tool engaged both patients and clinicians:

> *Clinician: "So what do you think we should do with, what's most important for you do you think, when to choose a medicine?"*
> *Clinician: "Do you want to talk about the risk of bleeding first or the risk of clotting first or the practical considerations?"*

Two clinicians commented that it took some time to get used to the tool and get it fully integrated in their consultation, or struggling with finding the appropriate language:

> *Clinician: "Quite honestly, I felt a bit awkward using the tool, but it was my first time using it. Like any new tool, I am sure it takes practice to make it flow smoother."*
> *Clinician: "I thought the tool was a great idea. It was a little harder to come up with the language to use to discuss it with the patient than I had*

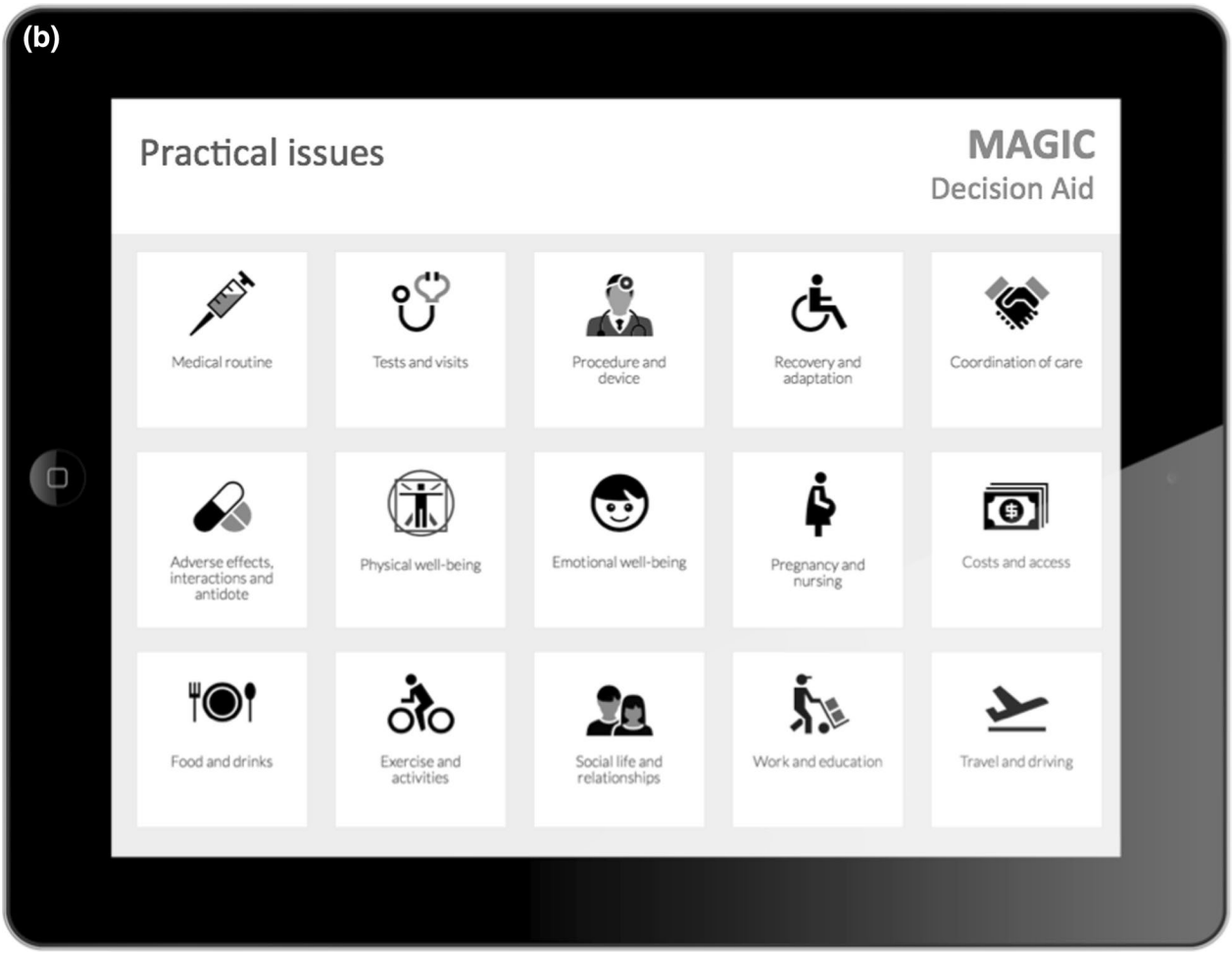


Fig. 4 continued

expected, but in general, I thought it worked well."

Understandability

Patients and clinicians used their own vocabulary to express how they understood how the tool could help them individualise the conversation related to risks, value elicitation and uncertainty in decision-making.

> *Clinician: This tool is supposed to help me explain, to compare the two [options] to help you decide what you would like to do at this point."*
> *Clinician: "We can reduce that number by 58 people [per 1000] if we give Rivaroxaban. So there is some value about taking it but there are some downsides. Now what's the downside you are worrying about most?"*

Visualisation of the evidence in the tool was informative, clear and easy to understand for most participants, while one patient reported that the pictographs were confusing:

> *76-year-old man with venous thromboembolism: "I liked the way it was presented [...] both numbers and figures were easy to understand"*
> *33-year-old pregnant patient with increased risk of venous thromboembolism: "Confusing looking at the board with figures [i.e. pictographs]"*

All consultations contained discussion about absolute risks of different outcomes. This part of the conversation using the tool was mostly led by the clinician. There was a broad variability in how clinicians and patients rephrased the risk estimates (e.g. "small" or "high" and also applying it to the specific patients' situation, particularly when less applicable:

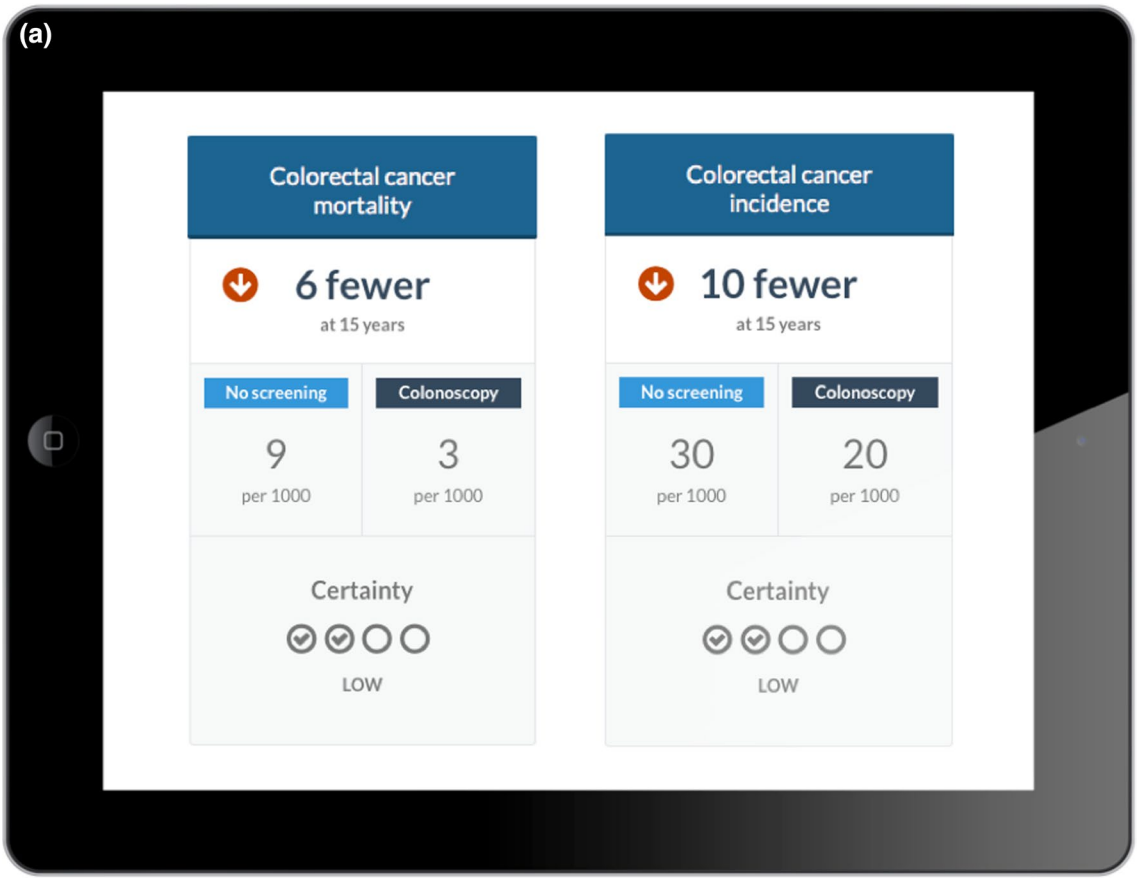


Fig. 5 Example of encounter decision aids; **a** Second layer comparing outcome cards; **b** third layer for further discussion of absolute risks and certainty

> *Clinician: "I wish I had data to say: okay if we had 1000 guys who rode a motorcycle all the time, what's the risk. I don't have that, and I will never be able to get that."*
> *Clinician: "I am not going to tell you can't play but I'm going to tell you, you have to be comfortable with carrying that risk. I don't want you to play [sport] scared right. You know, the other thing we talked about briefly is what happens if you bleed on Rivaroxaban, you know if somebody jabbed you or something you are going to bruise up."*

In the first iteration of the decision aid prototype, the certainty of the evidence was labelled as "confidence" without any further information.

> *33-year-old pregnant patient with increased risk of venous thromboembolism giving feedback on the decision aid: "The box where it said something about confidence in the results, it said low or high. It could explain if it was the medical confidence in the results or the users' confidence in the results."*

We then systematically incorporated in later iterations the main reason for the degree of certainty, taken from GRADE summary of findings. Most clinicians still often ignored certainty, except in specific conversations discussing mortality when faced with uncertainty, as they perceived patients would struggle to understand it:

> *Clinician: "I am not quite convinced that "uncertainty" is a concept that patients can grasp or that the way it is presented in the tool is all that helpful."*
> *Clinician: "I did appreciate having the quality of the evidence accessible as well. Though I don't recall using this feature more than one or two times in the encounter, it was nice to know that it was there."*

¾Users reported that medical abbreviations were not understandable, and the generic drug names needed explanation.

> *Clinician: "The only thing maybe ... one must always explain this with VTE, which is abbreviated."*

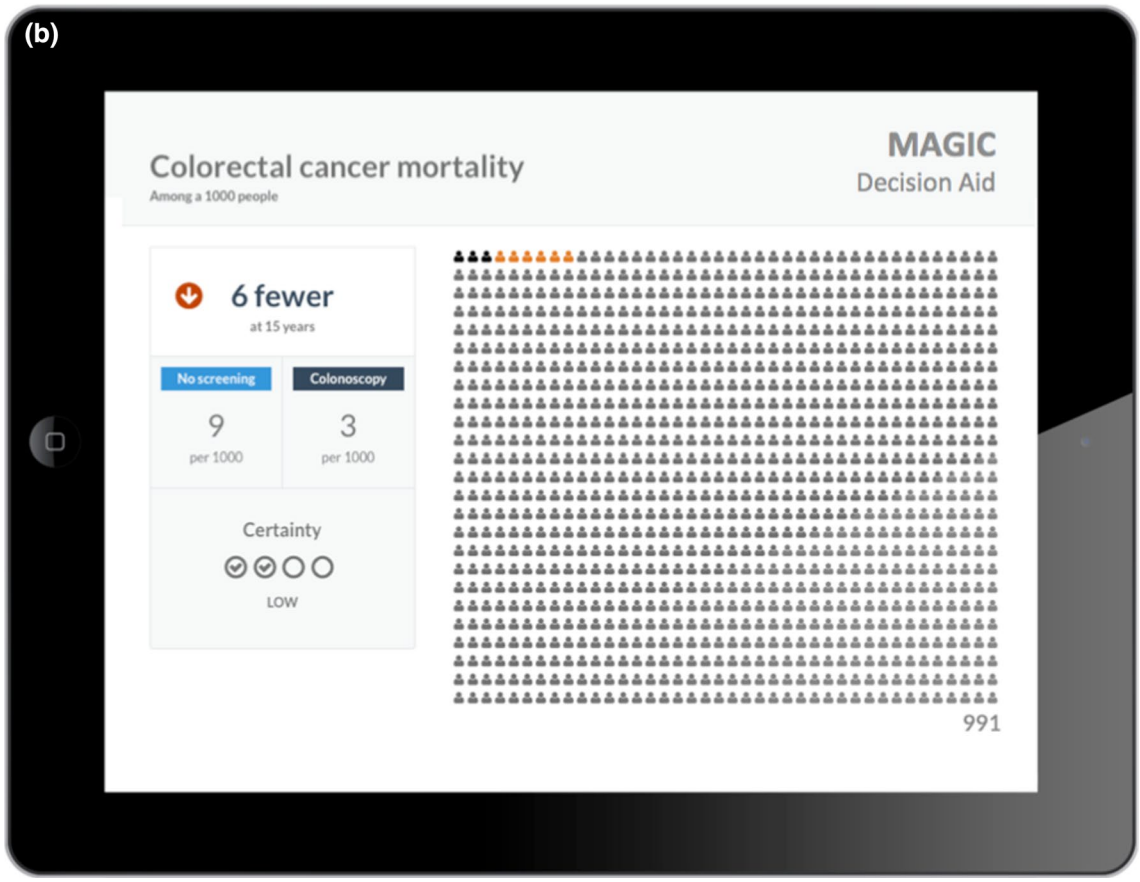


Fig. 5 continued

Table 1 Quantitative summary of facets of user-experiences and the quality of the experience using the decision aid

Facets of user-experience	Quality of the experience					
	Neutral experience	Positive feedback	Suggestions to improve	Minor frustrations	Major frustrations	Total
Accessibility	3	1	16	20	2	42
Credibility	1	6	1		–	8
Desirability	11	30	9	4	–	54
Findability	–	–	1	–	–	1
Identification	29	2	1	–	–	32
Understandability	126	34	11	11	6	188
Usability	43	33	14	26	2	118
Usefulness	40	81	17	5		143
Total	253	187	70	66	10	586

Usefulness

The majority of patients and physicians perceived that the tool was useful and supported better information, value clarification and shared decision-making. They felt the tool contributed to reaching a decision together, although some highlighted decision aids were not necessary to achieve a good consultation.

> *Clinician: "The patient thought that the tool made the benefits of her decision to continue taking tamoxifen the 10 years clearer; she had been told there was*

a benefit but did not know how much of a benefit before today."

After several consultations, both physicians and patients noted that they had been somehow surprised by the decision made. It allowed the presentation of useful information that would otherwise not have been brought up in the conversation:

73-year-oild man with venous thromboembolism: "Yes, if the tool wasn't used, I would probably not have gotten the information."
Clinician: "Surprisingly, the patient ended up choosing to stop medication after 3 months, congruent with his values. [It] probably wouldn't have been the mother's or father's decision. [They] would have preferred that he stopped basketball for this health".
Clinician: "At the time of consent, we were convinced using the tool wouldn't change anything. After the consultation, we thought it was really useful to look at the evidence, in particular graphically."

Another key observation related to the use of the tool, particularly with tablet computers, was that physicians and patients shifted posture from sitting across each other to side-by-side, looking at the tablet, and even holding it together when having a conversation.

64-year-old man with venous thromboembolism reflecting on the use of the decision aid: "It shown you graphs [...] rather than just sitting back verbally across the desk and saying... like Dr. X did."

The simplicity of the various presentation formats, allowing an overview, the comparison of benefits and harms, and the exploration of the same information in different formats, both visually and numerically, was highlighted as particularly useful:

52-year-old woman with breast cancer giving feedback on the decision aid: "The simplicity of it is actually one of its strengths I think"
64-year-old man with venous thromboembolism reflecting on the use of the decision aid: "I think a picture says 1000 words [...] it's giving you the stats, it's also showing you stats. [...] I think for many people it's easier to understand that when you see the graph than just to hear it and just see a number."

Users appreciated the possibility to easily compare and switch between different clinical outcomes, supporting the natural flow of the conversation rather than following a pre-defined pathway:

52-year-old woman with breast cancer giving feedback on the decision aid: "The most helpful feature is the flexibility - being able to switch between different clinical outcomes for any given clinical scenario. [...] I like the fact that you can bring them up side-by-side as well, I think that's really helpful rather than kind of exiting and entering, you know, and trying to remember the ones from before."
Clinician: "If the conversation shifts in a particular direction, e.g., the patient wants to talk more about bleeding, we can shift the tool in that direction. I really appreciated this flexibility, because it made my discussion more responsive and natural."

Views around the overall amount of information available or displayed were highly variable:

53-year-old woman with venous thromboembolism reflecting on the content of the decision aid: "A lot of information. Should not be less but is difficult to grasp all of it."
71-year-old man with venous thromboembolism reflecting on the content of the decision aid: Patient 15 (VTE): "No superfluous information. Very short and concise so rather have some more details."

Decision aids varied in the total number of outcomes that they included (i.e. up to 10 outcomes), which led to variable feedback on their optimal number or the order in which they may be presented at the top level of the tool, although users also recognized the value of choosing which one to focus on in the clinical conversation:

Clinician: "Actually, [the outcomes] are lost, the really important ones [...] perhaps a bit overwhelming."
Clinician: "I was going to say if there's a way of having the ones that are actually more relevant, but the point is that it's what the patient thinks is relevant isn't it?"

This issue led us to develop an authoring feature in MAGICapp that allows the selection of which outcomes of the GRADE summary of findings table to display in the decision aids, and the possibility of relabelling the outcomes (Table 2).

Finally, several patients highlighted that it would be useful to have written information to bring home to be able to remember the content of the conversation and to discuss with close ones:

47-year-old man with venous thromboembolism: "So, something that complements this that you can look on your own, at home, that's interesting. And spend a little extra time looking at it. Because you know, I'm going to go home, and my wife is going to ask me 100 questions."

Table 2 User-testing findings of barriers and issues and solutions to inform iterations of the encounter decision aids

	Barriers and issues discovered during the user testing	Changes in the subsequent iterations
Accessibility	Lack of contrast in text and pictographs Scrolling was needed to see all content when tablet was vertical Wi-Fi issues in hospitals	Enhanced contrasts, changed colours Scrolling removed Created off-line version and print version
Usability	Suggestion of a top layer to ease the introduction to the tool Difficulty coming up with language to use the tool Suggestion to combine the tool with information provided to patient before encounter Suggestion to have the possibility to change the denominator in the icons (and possibly in the numbers)	Supportive sentence "What aspect would you like to discuss next? Choose and compare" outcomes to raise choice awareness Possibility to change data entry and display directly in MAGICapp feeding in the interactive decision aid content
Understandability	Concept of certainty Medical abbreviations difficult to understand Generic drugs names confusing	Main reason for uncertainty made available one click away Names and descriptions of outcomes can be edited
Usefulness	Great variability in the perception of the appropriate amount of information, in particular the number and order of outcomes Useful to have something to bring home Suggestion of a feature that could compare several options	Number of outcomes and their order can be selected independently of underlying evidence profile Print version developed Multiple comparisons prototype in development
Identification	The patient's risk might be different from what is shown in the tool	Highlight during demonstration and in quick educational modules that this is encounter decision aid to be used together with a clinician, who can adapt content to each patient, highlighting potential similarities or differences
Credibility	Different colour of outcome card for practical issues could lead to selection bias	Specific design developed to display practical issues and navigate across them [12]
Findability	Clinician needed more information on evidence behind estimates in decision aids	Integration with MAGICapp with decision aids directly linked to GRADE evidence summaries

This issue led to the development of a printable version of the decision aid (Table 2).

Identification

The patients identified with the content and felt the tool was about their own choice. Physicians used the tool to enhance awareness of choices or to find out what mattered most, for example steering the conversation towards the daily life implications for patients:

> *90-year-old woman with venous thromboembolism: "So I just need to be careful not to prick myself with the needle when I'm sewing"*

Some patients felt that the physicians' knowledge was more relevant to their own decision than what was presented in the tool. Clinicians also spontaneously clarified when the patient's risk might differ from what is shown:

> *Clinician: Now this data, this stuff that we constructed from big studies, but this is a little different from you."*

Credibility

Both the physicians and patients perceived the tool as trustworthy, both in content and the way it was presented:

> *47-year-old man with venous thromboembolism reflecting on the decision aid: "I feel confident I saw all important information to take a decision"*
> *Clinician: "The order was correct: why you take the medication, what prevents it, the most important complications"*

Desirability

Many clinicians and patients expressed a preference in having the tool used in a consultation, rather than not, and one patient thought that the tool would empower patients.

> *74-year-old man with venous thromboembolism reflecting on the use of the decision aid: "I feel a bit privileged coming here, cause other patients that go to their GP might not get the same introduction"*
> *Clinician: "I think that a problem with many of these sorts of decision aids they just get too complicated, so I think this is quite nice" ... "I think it's great, I, I'd like to be able to use it in the clinic actually, because I think it's quite, quite a helpful way of practically explaining things to people with, with some detail, but not too much detail"*

Findability

Since the decision aid was directly provided for each encounter, we were unable to explore issues related to how challenging it would be to find it during an encounter. The only aspect that came up related to physicians' needs to have easy access to the supporting evidence for the estimates-effect provided in the decision aid. This issue was solved by the integration of decision aids in MAGICapp where all underlying evidence is directly linked to the decision aid (Table 3).

Changes made in presentation formats across iterations of the prototype

We performed four major iterations of the decision aid presentation formats based on user-experiences. Table 2 summarizes the identified issues and barriers followed by specific solutions that were implemented across iterations. Final versions of the generic decision aids were reached after the team reached consensus that the decision aid prototype successfully involved patients in shared decision-making and satisfied the needs of patients and physicians.

A final version of the generic decision aids was reached and read for integration in an authoring and publication platform for their generic and semi-automated creation. Table 3 summarized the main features in the decision aids.

Integration in MagicApp

We integrated the prototype in MAGICapp (Figs. 4, 5) The technical integration of the final version of the decision aid prototype specifically resulted in a: (1) automatically generated decision aids for all available GRADE evidence summaries linked to recommendations in the platform, (2) access to all underlying evidence, (3) automatic update of decision aids when the evidence summary is updated and (4) selecting the number of displayed outcomes and changing labels for more lay language wording whenever relevant.

MAGICapp has numerous (> 1000) available decision aids. Since users and customers of the platform are responsible for producing the evidence summary and own it, we have not performed a formal quality assurance of accuracy and clinical relevance of all available decision aids. The integration in MAGICapp also makes it possible to easily generate widgets so the decision aids can be integrated on other online platforms (e.g. button links to decision aids from the BMJ Rapid Recommendations.)

Discussion

We have developed encounter decision aids linked to evidence summaries that have informed trustworthy guidelines to facilitate shared decision-making with patients at the point of care. User-testing in real clinical encounters revealed opportunities for improvement in readability, understandability, usability and information overload that we addressed through four design iterations. After addressing these issues, user-testing demonstrated that the developed decision aids are understandable and intuitive; support conversation on issues that matter most to patients; and help clinicians share evidence regarding benefits, harms, their associated degree of certainty, along with practical issues relevant to each management option.

Table 3 Main concepts and features of the decision aids

- Electronic generic framework for decision aids integrated in an authoring and publication platform for guidelines and evidence summaries (MAGICapp)
- Decision aids are semi-automatically produced and updated based on content in MAGICapp with adaptation possibilities (e.g. wording and number of outcomes, language)
- Multi-layered presentation format:
  - First layer displays the list of patient-important outcomes and practical issues (Fig. 4a)
  - Second layer displays interactive outcome cards with evidence estimates, certainty, and patient-important practical issues across 15 generic categories. Possibility to interactively compare two or more outcomes in parallel (Fig. 5a)
  - Third layer displays a corresponding set of pictographs showing the absolute risk with each option (Fig. 5b) and practical issues related to the treatment option (Fig. 4b)
- Educational module developed http://magicproject.org/161128/ and integrated in MAGICapp. Content was generated to mimic the very short demonstration used during user-testing
- Print functionality of decision aids create pdf files that can be printed or used for notetaking and/or to bring home
- Prototype for comparisons between multiple options are developed and implemented in a BMJ Rapid Recommendation [28]
- Offline feature so decision aids can be used without use of Internet
- Widgets from MAGICapp to grab and show a given decision aid on any other online platform Example: Rapid Recommendation on Prostate cancer screening (https://www.bmj.com/content/362/bmj.k3581 to BMJ infographic) which links to MAGICapp content, including widgets to decision aids for various profile of patients

Strengths and limitations

Strengths of our project include the user-testing of the decision aids in real-life consultations and in a variety of clinical settings. Suggestions for improvements from users resulted in changes that produced a higher degree of usability and accessibility.

The brief introduction to the tool proved sufficient that clinicians and patients described it as easy to use and understand.

A key element is the perceived trustworthiness of the content, which was captured by the user experience dimensions of credibility and identification. Clinicians also outlined their need to link back to the detailed evidence summary and sources of uncertainty, which the tool provides.

In regard to limitations, our study may have selected clinicians who were more versed in, and more enthusiastic about innovative approaches for risk communication. Moreover, the current study focused only on situations in which patients face two management alternatives and did not explore decision aids for multiple comparisons. Development a tool dealing with multiple options is in progress, and is currently included in recent BMJ Rapid Recommendations, for example on screening for colorectal cancer [28].

Shared decision-making hinges on clinicians having access to up-to-date and quality appraised evidence [29]. This was achieved by integrating the framework in MAGICapp to semi-automatically produce decision aids based on content from guidelines and evidence summaries. This, however, requires someone to carry out the updating process, which remains a hit-or-miss phenomenon.

User-testing was performed before mandatory social-distancing required by COVID-19 restrictions. Generalizability to virtual consultations remains to be confirmed, although the online nature of the tool allows its use from afar.

Implications for encounter decision aid production

Information overload is a critical challenge in the development of evidence-based tools. This is particularly true for decision aids, which risk excessive information that may compromise useful conversations. To that end, the design of our generic decision aids was heavily inspired by the work of Montori and colleagues who identified the need for encounter decision aids to be as "quiet" as possible: i.e. that the tool does not impose a necessary sequence of predefined algorithms of questions and answers, that pushes the interaction into a pre-defined script, but instead organizes information so as to support the actual conversation that occurs between clinicians and their patients on what matter most to them [15, 30].

We implemented a similar approach through our interactive multi-layered formats. User testing allowed us to explore those elements that were better to highlight in top levels and those that could be presented in deeper layers. The final version of the decision aids has a top layer displaying only the list of outcomes and practical issues, without any numbers. Intermediate layers provide a synoptic view of each potential benefit or harm, followed by deeper layers providing detailed pictographs and underlying information, such as reasons for uncertainty in the estimates of benefits and harms.

Such information was sometimes useful and other times distracting. Iterative user-testing demonstrated that patients appreciated the flexibility of this approach, as well as the possibility to easily switch between different outcomes and issues. Moreover, as the number and labelling of outcomes in the decision aids sometimes needed to differ from the supporting GRADE evidence summary, we implemented the functionality to edit the decision aids automatically generated in MAGICapp.

The multi-layered approach allows the display of more outcomes than static GRADE summary of findings tables allow (usually not more than 7 most critical and important outcomes) [31]. In common with other encounter tools tested by Montori et al., patients reached most decisions after exploring only a selection of outcomes (usually 3 to 4) [15, 30].

SHARE-IT represents the first successful, user-tested effort to fully integrate production of decision aids with the production and dissemination of evidence summaries, recommendations and guidelines. This integration also makes it possible to adapt the content (e.g. to national guidelines or policies or certain populations). The content and quality of the decision aids are, however, dependent on the quality of the evidence summaries.

Education and training are also central in any implementation strategy. Use of SHARE-IT decision aids required minimal demonstration of the tool, as shown in our short online education module [32]. This was sufficient to explore it intuitively during a real clinical encounter. As piloted by several clinical educators in our team, example of such decision aids linked to guidelines can be used in rounds and bedside teaching, a strategy that warrants further evaluation. This may help to overcome an important barrier: the benefits of using decision aids (as well as engaging in shared decision-making altogether) are really known after one has experienced it.

Conclusion

Our study provides a proof of concept that encounter decision aids can be generically produced from GRADE evidence summaries or recommendations for clinical practice. Further evaluation is needed in more clinical

contexts and as part of educational and broader implementation strategies. This would require that decision aids are available for a large number of clinical decisions. The integration of SHARE-IT decision aids in MAGICapp offers great potential in scaling up their production and continuous update along with evidence summaries and clinical practice guidelines.

Abbreviations

DECIDE: Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence project; IPDAS: International Patient Decision Aid Standards; GRADE: Grading of Recommendations Assessment, Development and Evaluation; SHARE-IT: Sharing Evidence to Inform Treatment decisions project.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-021-01541-7.

Additional file 1. Standards for Reporting Qualitative Research checklist.

Acknowledgements

We thank Frankie Achille (interaction designer/developer), Rob Fracisco (designer/developer), and Deno Vichas and Chris Degiere (developers) for their contributions in development of the online authoring and publication platform (www.magicevidence.org).

Authors' contributions

The SHARE-IT project was conceived by the MAGIC Evidence Ecosystem Foundation, in close collaboration with the DECIDE project and GRADE working group, to which most contributors are affiliated (AFH, POV, EA, GHG, TA). We also received numerous comments from stakeholders at international meetings. AFH and TA led and coordinated the project, supervised by GHG and POV. TA, AFH, FA, LB, and POV developed and implemented the prototype, and all contributors provided comments and feedback at different stages. TA, AFH, ST and POV performed user-testing in clinical encounters. AFH drafted the manuscript and all authors critically revised the manuscript. AFH is guarantor. All authors read and approved the final manuscript.

Funding

AFH was financially supported by a PhD fellowship from Innlandet Hospital Trust and have received innovation grants from South-Eastern Norway Regional Health Authority. TA was financially supported by a fellowship for prospective researchers Grant No P3SMP3-155290/1 from the Swiss National Science Foundation. The funding body had no role in design of the study, collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the research ethics committees at the respective participating sites. Patients and physicians participating in the study provided written informed consent. The Scottish study was approved by National Research Ethics Service Committee South West—Frenchay (15/SW/0127). The Norwegian study was approved by Regional Committees for Medical and Health Research Ethics (Ref. nr.: 2013/1630). The Canadian study was approved by Hamilton Integrated Research Ethics board (Ref. 13-373).

Consent for publication

Not applicable.

Competing interests

POV, LB, GG, and TA are board members of MAGIC Evidence Ecosystem Foundation (www.magicevidence.org), a not-for-profit organization which provides authoring and publication software (MAGICapp) for evidence summaries, guidelines, and decision aids.

Author details

[1]Department of Medicine, Innlandet Hospital Trust, Mailbox 104, 2381 Brumunddal, Norway. [2]MAGIC Evidence Ecosystem Foundation, Oslo, Norway. [3]Institute for Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway. [4]Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada. [5]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon. [6]Health Services Research Unit, University of Aberdeen, Aberdeen, UK. [7]Division of Clinical Epidemiology and Division of General Internal Medicine, University Hospitals of Geneva, Geneva, Switzerland.

Received: 15 February 2021 Accepted: 25 May 2021
Published online: 29 June 2021

References

1. Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ. 2012;344:e256. https://doi.org/10.1136/bmj.e256.
2. Trevena LJ, Zikmund-Fisher BJ, Edwards A, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak. 2013;13(Suppl 2):S7. https://doi.org/10.1186/1472-6947-13-S2-S7.
3. Stacey D, Légaré F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2017. https://doi.org/10.1002/14651858.CD001431.pub5.
4. Agoritsas T, Heen AF, Brandt L, et al. Decision aids that really promote shared decision making: the pace quickens. BMJ. 2015;350:g7624. https://doi.org/10.1136/bmj.g7624.
5. Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. Chest. 2013;144(2):381–9. https://doi.org/10.1378/chest.13-0746.
6. MAGIC Evidence Ecosystem Foundation. www.magicevidence.org. Accessed 28 Apr 2021.
7. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6. https://doi.org/10.1136/bmj.39489.470347.AD.
8. Treweek S, Oxman AD, Alderson P, et al. Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. Implement Sci. 2013;8:6. https://doi.org/10.1186/1748-5908-8-6.
9. Kristiansen A, Brandt L, Alonso-Coello P, et al. Development of a novel, multilayered presentation format for clinical practice guidelines. Chest. 2015;147(3):754–63. https://doi.org/10.1378/chest.14-1366.
10. Brandt L, Vandvik PO, Alonso-Coello P, et al. Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial. BMJ Open. 2017;7(2):e011569. https://doi.org/10.1136/bmjopen-2016-011569.
11. DECIDE. https://www.decide-collaboration.eu. 2011–2015. Accessed 1 July 2020.
12. Heen AF, Vandvik PO, Brandt L, et al. A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines. J Clin Epidemiol. 2020;129:104–13. https://doi.org/10.1016/j.jclinepi.2020.10.002.
13. Heen AF, Vandvik PO, Agoritsas T, et al. Including practical issues and patient perspectives in rapid recommendations. BMJ. 2018;363:k4871. https://doi.org/10.1136/bmj.k4871.
14. Witteman HO, Dansokho SC, Colquhoun H, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev. 2015;4:11. https://doi.org/10.1186/2046-4053-4-11.
15. Montori VM, Breslin M, Maleska M, et al. Creating a conversation: insights from the development of a decision aid. PLoS Med. 2007;4(8):e233. https://doi.org/10.1371/journal.pmed.0040233.
16. Wroblewski L. Mobile first: a book apart 2011.

17. Cates C. Visual Rx. http://www.nntonline.net/visualrx/introduction/. Accessed 1 July 2020.
18. Blueprint. http://blueprintcss.org. Accessed 1 July 2020.
19. Yepes-Nunez JJ, Li SA, Guyatt G, et al. Development of the summary of findings table for network meta-analysis. J Clin Epidemiol. 2019;115:1–13. https://doi.org/10.1016/j.jclinepi.2019.04.018.
20. Guyatt GH, Akl EA, Crowther M, et al. Introduction to the ninth edition: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. 2012;141(2 Suppl):48S-52S. https://doi.org/10.1378/chest.11-2286.
21. Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet. 2013;381(9869):805–16. https://doi.org/10.1016/S0140-6736(12)61963-1.
22. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490. https://doi.org/10.1136/bmj.328.7454.1490.
23. Edwards A, Elwyn G, Hood K, et al. The development of COMRADE—a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. Patient Educ Couns. 2003;50(3):311–22. https://doi.org/10.1016/s0738-3991(03)00055-7.
24. Edwards A, Elwyn G, Hood K, Robling M, et al. The development of COMRADE—a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. Patient Educ Couns. 2003;50:311–22.
25. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277–88. https://doi.org/10.1177/1049732305276687.
26. P M. User experience design. https://semanticstudios.com/user_experience_design/. 2004. Accessed 15 May 2017.
27. Rosenbaum S. Improving the user experience of evidence. A design approach to evidence-informed health care. 2011.
28. Helsingen LM, Vandvik PO, Jodal HC, et al. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline. BMJ. 2019;367:l5515. https://doi.org/10.1136/bmj.l5515.
29. Hoffmann TC, Legare F, Simmons MB, et al. Shared decision making: what do clinicians need to know and why should they bother? Med J Aust. 2014;201(1):35–9. https://doi.org/10.5694/mja14.00002.
30. Wyatt KD, Branda ME, Anderson RT, et al. Peering into the black box: a meta-analysis of how clinicians use decision aids during clinical encounters. Implement Sci. 2014;9:26. https://doi.org/10.1186/1748-5908-9-26.
31. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–94. https://doi.org/10.1016/j.jclinepi.2010.04.026.
32. MAGIC. SHARE-IT Education Module. https://magicevidence.org/share-it/tour/.

Publisher's Note

Standards for Reporting Qualitative Research (SRQR)*

http://www.equator-network.org/reporting-guidelines/srqr/

Page/line no(s).

Title and abstract

Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	Line 1-2
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	Line 44-75

Introduction

Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	Line 91-111
Purpose or research question - Purpose of the study and specific objectives or questions	Line 112-119

Methods

Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	Line 235-250
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	Line 216-221
Context - Setting/site and salient contextual factors; rationale**	Line 206-213
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	Line 210-213
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	Line 657-664
Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	Line 215-230

Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	Line 216-230
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	Line 262-269
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	Line 216-230
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	Line 238-250
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	Line 239-244

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	Line 253-554
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Line 253-554

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	Line 556-643
Limitations - Trustworthiness and limitations of findings	Line 572-594

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	Line 673-677
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	Line 681-688

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

Reference:
O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. **Standards for reporting qualitative research: a synthesis of recommendations.** *Academic Medicine*, Vol. 89, No. 9 / Sept 2014
DOI: 10.1097/ACM.0000000000000388

Article 3

Journal of Clinical Epidemiology 129 (2021) 104−113

ORIGINAL ARTICLE

A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines

Anja Fog Heen[a,*], Per Olav Vandvik[b], Linn Brandt[a], Victor M. Montori[c], Lyubov Lytvyn[d], Gordon Guyatt[d], Casey Quinlan[e], Thomas Agoritsas[d,f]

[a]*Department of Medicine, Innlandet Hospital Trust, Gjøvik, Norway*
[b]*Institute for Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway*
[c]*Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA*
[d]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Ontario, Canada*
[e]*Mighty Casey Media LLC, Richmond, VA, USA*
[f]*Division of General Internal Medicine, University Hospitals of Geneva, Geneva, Switzerland*

Accepted 6 October 2020; Published online 10 October 2020

Abstract

Objectives: The objective of the study was to develop and test feasibility of a framework of patient-important practical issues.

Study Design and Setting: Guidelines and shared decision-making tools help facilitate discussions about patient-important outcomes of care alternatives, but typically ignore practical issues patients consider when implementing care into their daily routines. Using grounded theory, practical issues in the HealthTalk.org registry and in Option Grids were identified and categorized into a framework. We integrated the framework into the MAGIC authoring and publication platform and digitally structured authoring and publication platform and appraised its use in The BMJ Rapid Recommendations.

Results: The framework included the following 15 categories: medication routine, tests and visits, procedure and device, recovery and adaptation, coordination of care, adverse effects, interactions and antidote, physical well-being, emotional well-being, pregnancy and nursing, costs and access, food and drinks, exercise and activities, social life and relationships, work and education, travel and driving. Implementation in 15 BMJ Rapid Recommendations added 283 issues to 35 recommendations. The most frequently used category was procedure and device, and the least frequent was social life and relationship.

Conclusion: Adding practical issues systematically to evidence summaries is feasible and can inform guidelines and tools for shared decision-making. How this inclusion can improve patient-centered care remains to be determined.

Keywords: Shared decision-making tools; Decision aids; Patient experience; Clinical practice guidelines; Patient-important outcomes

Funding/Role of the sponsor: A.F.H. was financially supported by a PhD fellowship from Innlandet Hospital Trust and has received innovation grants from South-Eastern Norway Regional Health Authority. The funding sources had no involvement in study design, collection, analysis, interpretation of data, in writing of the report, or in the decision to submit the article for publication.

Declaration of interests: P.O.V., L.B., G.G., and T.A. are board members of MAGIC Evidence Ecosystem Foundation (www.magicevidence.org), a not-for-profit organization which provides authoring and publication software (MAGICapp) for evidence summaries, guidelines, and decision aids. V.M.M. is a co-developer of minimally disruptive medicine, shared decision-making tools, and has participated in the development of measures of burden of treatment. He derives no income from any of these activities. A.F.H., L.L., and C.Q. have no interests to declare.

Authors' contributions: A.F.H. conceived the idea and led the project, supervised by T.A. and P.O.V. V.M.M. made substantial contributions to development of the idea. T.A., A.F.H., L.B., and P.O.V. developed and implemented the framework in MAGICApp. A.F., L.L., T.A., P.O.V., C.Q., and G.G. contributed in the implementation in BMJ RapidRecs. T.A., A.F.H., and P.O.V. performed user testing in clinical encounters. A.F.H. drafted the manuscript, and all authors critically revised the manuscript. A.F.H. is the guarantor.

* Corresponding author. Department of Medicine, Innlandet Hospital Trust, Sykehuset Innlandet, Postboks 104, 2381 Brumundda, Norway. Tel.: +47 915 06 200.

E-mail address: anjaheen@gmail.com (A.F. Heen).

https://doi.org/10.1016/j.jclinepi.2020.10.002

What is new?

Key findings

- The practical issues related to how a treatment or test affects the patients' daily life are frequently missing in guidelines and in many tools for shared decision-making.
- We developed a generic framework that allows for inclusion of practical issues in guidelines, evidence summaries and tools for shared decision-making through the authoring and publication platform MAGICapp.
- Applying this framework to 15 guidelines in The BMJ Rapid Recommendations allowed the identification of 283 practical issues associated with 35 recommendations.

What this adds to what was known?

- With the help of a generic framework, key practical issues can be identified and integrated in evidence summaries informing recommendations.
- Using an authoring platform can digitally structure data on practical issues and be translated in generic tools for shared decision-making.

What is the implication and what should change now?

- Systematic inclusion of practical issues can support patients to make decisions that not only make intellectual sense, that is, they are evidence-based, but also make practical sense, that is, they are feasible given the patient's capacity to enact them in their daily lives without undue burden.
- Optimal production, evidence retrieval, and synthesis of relevant practical issues for decision-making will require further exploration.

1. Introduction

Implementing optimal approaches to resolving a patient's health situation is at the center of health care practice. Trustworthy, evidence-based decision support tools can aid this process. Shared decision-making (SDM) tools that support meaningful conversations between each patient and clinician [1−4] can complement clinical practice guidelines (hereafter guidelines) that provide recommendations for practice and group-level evidence summaries.

The goal of SDM tools is to provide individual care that makes intellectual, emotional, and practical sense to each patient [5,6]. However, many current decision support tools typically lack or have minimal information addressing the *practical issues* patients face when implementing treatment options or tests, and how implementation affects their daily life [7,8]. Such omission is problematic: Addressing practical issues often contributes to what has been described as ''the work of being a patient'' and, when excessive, constitutes an onerous burden of treatment [9−11].

Practical issues can include coordination of care, required tests or office visits, as well as effects on social activities, diet, work, or travel. Although some practical issues are sometimes reported as outcomes in published research—such as the average length of hospital stay after a procedure—other issues are much less frequently addressed. Such less frequently addressed issues include cost, preparation required, and typical patient experience of its implementation in daily life (e.g., impact on social life of antidiabetic injection). The wide range of practical issues may explain why in most evidence summaries from systematic reviews, they are typically overlooked. Another reason may be a physician-centered rather than a patient-centered perspective on the experience of using a treatment.

The lack of patient-important practical issues in guidelines and most SDM tools is problematic and calls for innovation in helping to remedy this serious omission. A new framework for patient-important practical issues may help to improve the situation and inform trustworthy guidelines and SDM tools.

In this study, we report on how we addressed three objectives:

1. The development of a generic framework of patient-important practical issues from patient experience databases.
2. The integration of this framework in the MAGIC authoring and publication platform (MAGICapp) for the digital structuring of practical issues' summaries and inclusion in guidelines and SDM tools.
3. The feasibility of using this framework and authoring tool in the production of international recommendations and linked SDM tools in The BMJ Rapid Recommendations (RapidRecs) [12].

This work was based on insights from the research and innovations from the MAGIC Evidence Ecosystem Foundation, a nonprofit initiative set up to facilitate the digital creation, publication, and updating of guidelines, evidence summaries, and SDM tools [2,12,13].

2. Methods

2.1. Development of a generic framework for practical issues

2.1.1. Data sources

To identify generic categories of patient-important practical issues, we chose a purposeful sample of two data sources covering a large and varied set of health conditions,

both of which applied a rigorous and trustworthy methodology in identifying patient experience and their most frequently asked questions, including practical issues.

The first data source was the Health Talk registry (www.healthtalk.org), from the Health Experiences Research Group and the University of Oxford, UK, a large sample of patient experience collected through thorough methodology using focus groups and standardized interviews on a large set of health conditions [14].

The second source was the Option Grids, which were produced, at the time of this analysis, at the Dartmouth Center for Health Care Delivery, USA, and Cardiff University, UK. Option Grids constitute at that time a specific example of decision aids that aimed to include patient experience in the form of frequently asked questions, which were elicited in a standardized methodology [15,16].

Our two large data sources thus allowed us broad and varied sample of practical issues so as to inform the development of a generic framework.

2.1.2. Data collection, coding, and categorization of practical issues

Using Strauss and Corbin's grounded theory approach [17] with an iterative study design, we collected all data, including videos, transcripts, and text available in our sources as of March 2014. Data were collected in single words, phrases, or paragraphs dependent on context and processed into a Microsoft Excel database. Two researchers (AFH and TA) conducted iterative coding and comparison in parallel with data collection. Data were analyzed by identifying specific codes to all practical issues in the data. We grouped codes with similar content together with each emerging theme (e.g., the following data were coded as practical issues related to "travel": "*Several people pointed out that traveling with a stoma required some preparation*" and "*… experienced problems when traveling, for example, long flights triggering seizures or adjusting their medicine-taking when traveling to different time zones*"). Using an inductive approach (constant comparative analysis), we compared codes, thus refining them and aggregating emerging themes into broader categories (e.g., in the previous examples ended up the broader category of "travel and holidays"). We performed axial coding to explore and define connections between categories [18]. This iterative inductive approach identified generic categories that informed our practical issue framework (Fig. 1).

Data collection included all issues relevant to management of options on either therapeutic or diagnostic alternatives, excluding issues that were only about experiencing a health condition. To validate the categories, we applied the categories to the original sample of data and, in April 2020, to a new data sample with a random extraction from the original sources.

2.2. Integration of the practical issues framework to facilitate inclusion in guidelines and SDM tools

The MAGICapp is an online authoring and publication platform for guidelines and evidence summaries. Here, evidence summaries are digitally structured, which allows for translation of data into various formats including multilayered guidelines and SDM tools. After defining the final set of generic categories, we integrated the framework in the data structure of the MAGICapp structuring of data so that practical issues are included in guidelines and supporting evidence summaries created through the online platform, in complement to traditional outcomes for benefits and harms [19].

We developed presentation formats of SDM tools as part of the SHARE-IT project, whose objective is to create decision aids supporting SDM in the clinical encounter. We have previously described the methods in SHARE-IT [19] and have conducted user testing in real clinical encounters (results are currently analyzed and will be reported in a manuscript in preparation [20]).

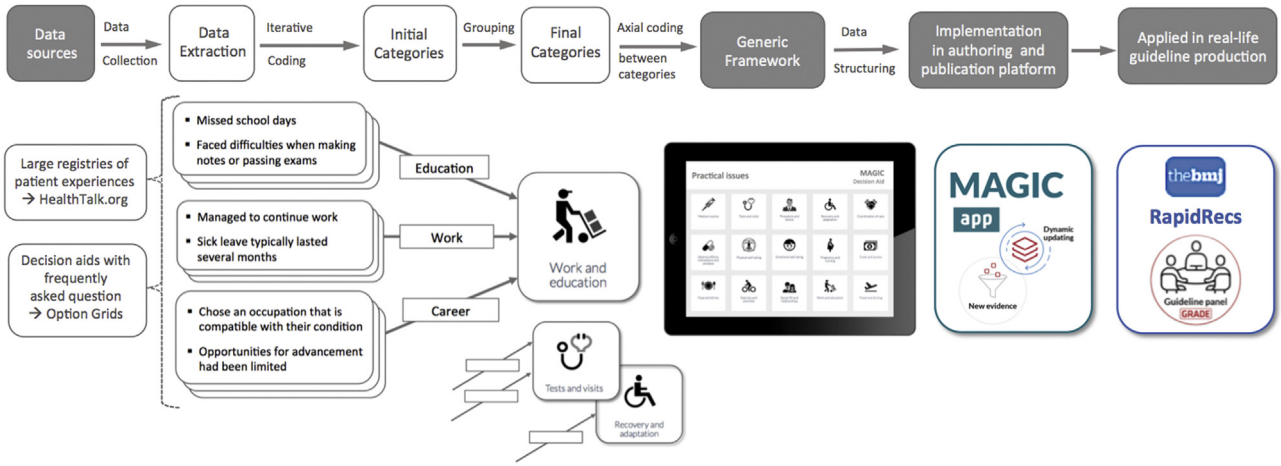


Fig. 1. Outline of the [1] development of the generic framework for practical issues from two large data sources, [2] integration in a digitally structured authoring and publication platform (MAGICapp), and [3] application in real-life guidelines (BMJ RapidRecs).

Implementing our framework in SHARE-IT decision allows for authoring and publication of narrative facts and evidence on practical issues in evidence summaries, guidelines, and SDM tools presented in multilayered formats with corresponding pictographs for each category (Fig. 1).

2.3. Feasibility of using the practical issues framework in the BMJ RapidRecs project

Guideline panels in the BMJ RapidRecs include patient partners, frontline clinicians, content experts, researchers, and guideline experts. Panelists respond to potential practice-changing evidence with trustworthy guideline recommendations, SDM tools, infographics displaying pertinent information, and key practical issues relevant for the recommendations developed and published in MAGICapp and in the BMJ [12].

To assess feasibility of using the framework in guideline development, it was introduced to the guideline panels of the BMJ RapidRecs. For each recommendation, we performed searches for practical issues and treatment burden in patient experience databases (e.g., Healthtalk.org), existing information leaflets and SDM tools, research of patients' values and preferences [21–23], and online evidence textbooks (e.g., medication side effects, from resources including UpToDate and DynaMed). We sought additional information from patient partners and other panel members who discussed the results and classified the emerging issues into the relevant categories of the practical issue framework.

Patient partners contribute to all aspects of the guideline and are fully included in the development process. In particular, they inform, review, and refine the practical issue section of all evidence summaries. This is carried out in dialog with frontline clinicians and other experts to ensure the integration of all elements relevant for decision-making in the body of evidence appraised in the process. We further asked the whole panel for feedback on the process and interviewed particular patient partners regarding the usefulness of the framework to identify practical issues.

3. Results

3.1. Development of the generic framework for practical issues

The team collected, coded, and analyzed data from 297 themes in the Health Talk registry and 29 Option Grids (Appendix 1) and then, to ensure that categories were generic and sufficiently inclusive across topics, performed data comparisons within each source and between sources. Researchers then combined codes into categories of

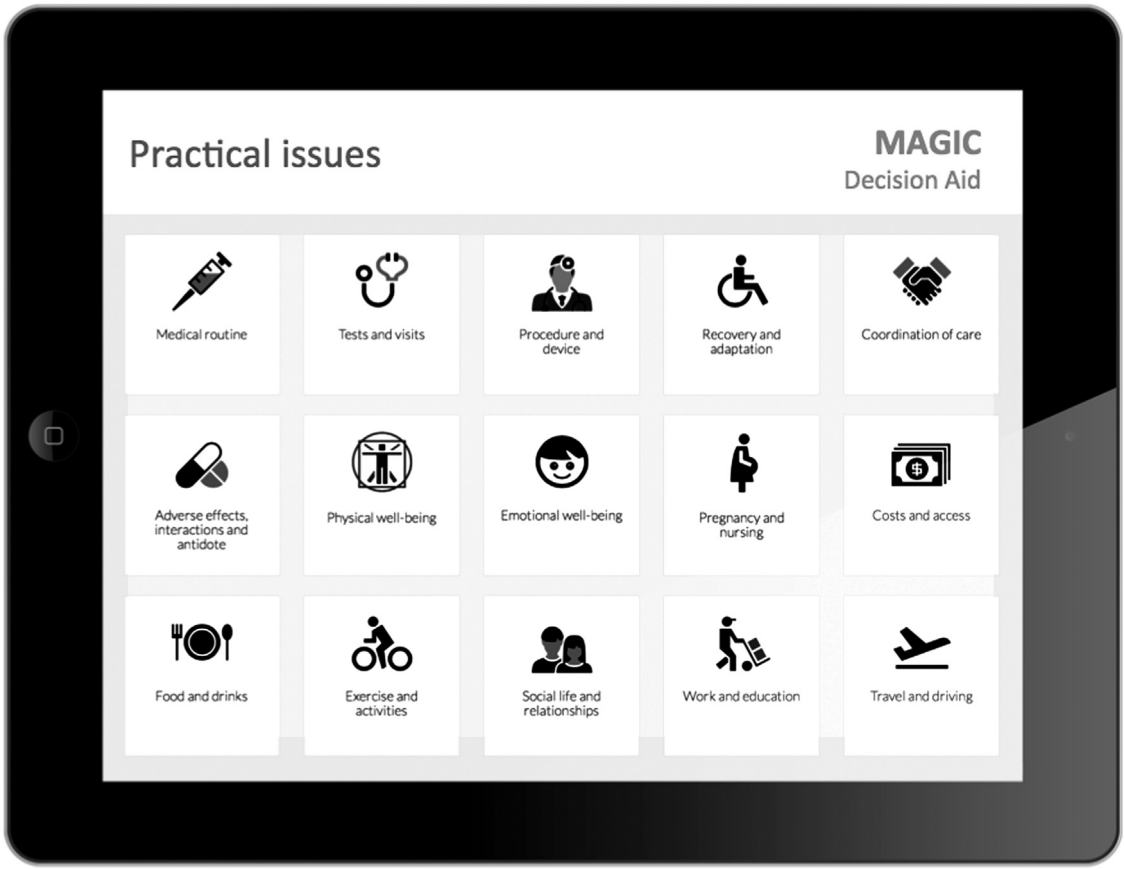


Fig. 2. The final practical issue framework including 15 categories and corresponding icons in SDM tools.

practical issues, initially resulting in 42 categories. Comparisons between categories and testing the categories against the original data resulted in categories being combined (Fig. 1).

Validation against a random extraction of the original sources did not result in new categories. Axial analysis resulted in a final set of 15 generic categories of practical issues (Fig. 2). Some are the direct implication of care and procedures (medication routine, tests and visits, procedure and device, recovery and adaptation, coordination of care). Other related to the impact on daily life (food and drinks, exercise and activities, social life and relationships, work and education, travel and driving). Additional categories included pregnancy and nursing, additional adverse effects, interactions, and antidote, physical well-being, emotional well-being and costs and access.

3.2. *Integration of practical issues in SDM tools—results from user testing*

We implemented the framework for practical issues in SDM tools in MAGICApp. To facilitate the inclusion in natural conversation, independent of health literacy [3], we developed pictographs and structured them in a top layer grid that presents each category of the framework in one screen (Fig. 3A). By clicking on each icon, users can access the content as a superimposed box displaying narrative information on the relevant practical issues, either as key words or short sentences (Fig. 3B). Categories that do not include any relevant content are still displayed on the top layer grid, as they may trigger meaningful questions from patients.

User testing of the SDM tools with practical issues, performed in 28 real-life consultations for a variety of clinical topics, demonstrated a high satisfaction with the presentation format displaying the practical issues. We report here some selected quotes from patients exposed to the SDM tools:

> Quote from a patient: ‘‘The information about practical issues was best, and what matters to me on day to day basis.’’

Most patients chose to explore the practical issues framework in the SDM tool being used in the conversation. Patients perceived the pictographs as intuitive and easily understood and were satisfied with the amount of information initially displayed.

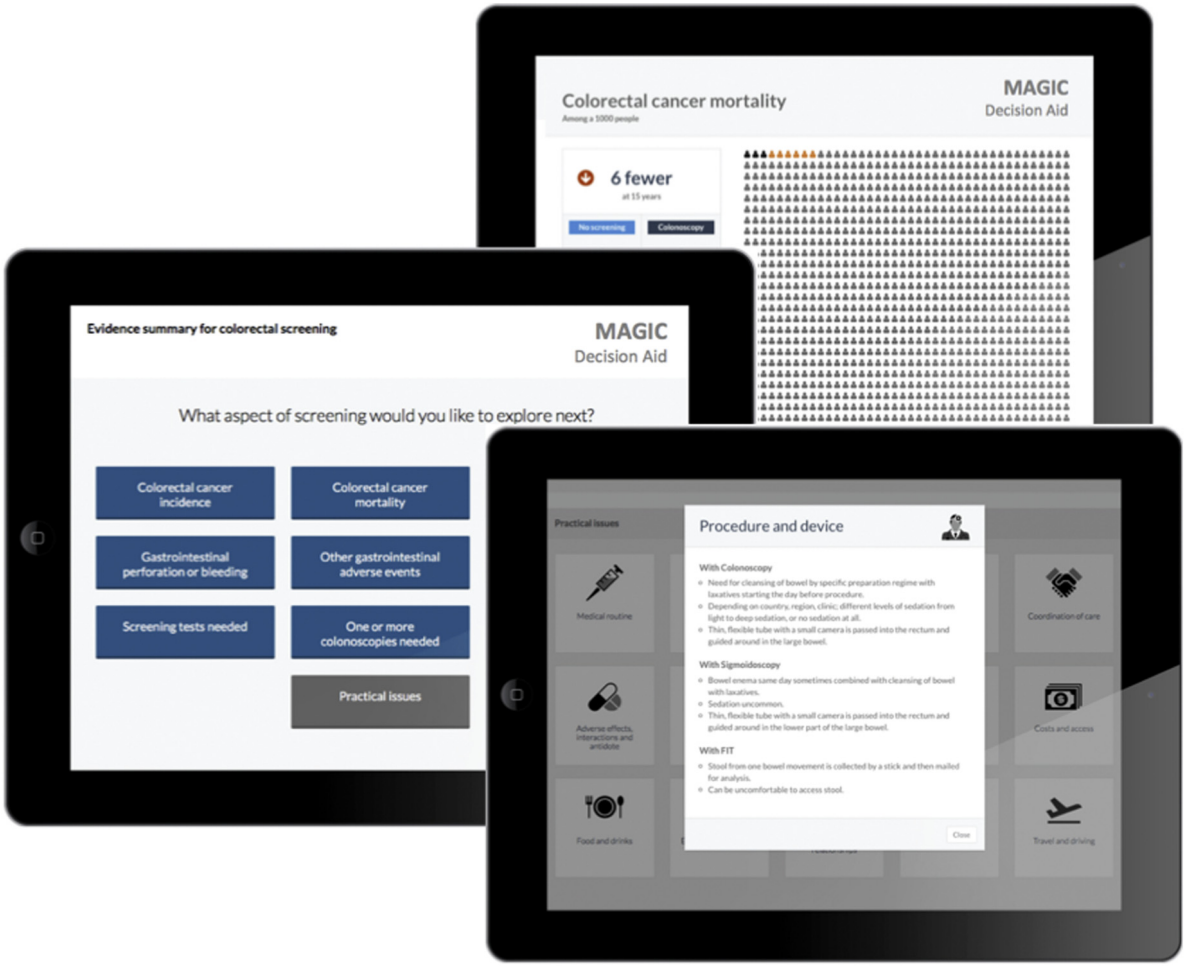


Fig. 3. The first layer of the decision aid (left) shows an overview of outcomes relevant for decision-making including practical issues. Deeper layers allow the display of absolute benefits and harms using pictographs as well as relevant practical issues.

> Quote from a patient: "I appreciated that text on practical consequences was short, or else I would start reading the small text."

The navigation in the tool and between practical issues and other outcomes was perceived as seamless and easy. Many patients asked for more information about the practical issues, and the layout helped them identify the issues that were the most important for them.

> Quote from a patient: "But actually, I'd like to have some more information about why we're not taking any blood tests? What's really the reason for that? Why can't you measure how is really… how [rivaroxaban] is really working on me."

Many patients and clinicians found that use of the tool resulted in them considering issues that they would otherwise likely not have discussed in the encounter. Several clinicians were surprised that the patient was interested in the practical issues and how this weighed in the decision-making process.

> Quote from a clinician: "At the same time it became clear that practical consequences were important for her. In particular, her travel to the doctor. This was positive, not sure if I had captured this without the decision aid. This was a good thing."

User testing also demonstrated that SDM tools displaying practical issues alongside evidence about health outcomes enhanced SDM [20]. Finally, displaying the 15 categories of our framework (Fig. 2) showed that even categories without a second layer of narrative information did trigger relevant conversations and collaborative deliberation about issues not listed in the tool, but that mattered to the patient.

3.3. Feasibility of using the practical issues framework and authoring tool in the BMJ RapidRecs

To prospectively assess their feasibility, we applied the practical issues framework in 15 BMJ RapidRecs from 2016 to December 2019 [12,24−38]. These recommendations included weak and strong recommendations according to GRADE [39] and covered both treatments, tests, and screening interventions. Key practical issues were also added to linked infographics that provide the gist of the evidence and links to detailed evidence summaries published online.

Feedback, suggestions for improvement, and challenges with using the framework were collected and solutions iteratively developed and implemented. Feedback included the need to identify evidence sources for practical issues, methods to synthesize and appraise them as many practical issues were not found by searching databases and repositories (e.g., PubMed), or when such studies were found, assessing their quality and validity remained often challenging. Guideline panel members felt that a combination of search in patient experience databases and patient information leaflets, in complement to medical literature searches, was the most feasible approach to identify key relevant practical issues. Direct deliberation between frontline clinicians and patients' partners about practical issues provided rich data useful for searching, appraising, and deciding on the importance of the different practical issues. Co-writing of statements with patient partners resulted in less detailed statements than when physicians wrote statements without the support of patient partners.

Procedure and device and adverse effects, interactions, and antidote constituted the categories to which practical issues were most frequently added and social life and relationship and coordination of care and travel and driving the least frequent. In total, 283 different practical issues were added to 35 recommendations across 15 different BMJ RapidRecs guidelines (Table 1).

4. Discussion

In the context of guideline development, we found that practical issues could be generated and applied across a wide range of clinical topics, in which practical issues were discussed along evidence summaries and contributed to recommendation making.

Digitally structured data proved valuable in reducing waste and resources used by developers of guidelines and SDM tools. Practical issues discussed in guideline panels will be instantly available in the corresponding SDM tools, with adaptation possibilities. Updating of the practical issues and other evidence will also only need to be performed at one place, reducing resources needed to keep SDM tools updated.

4.1. Strengths and weaknesses in relation to other studies

Our study includes user testing of SDM tools with practical issues in clinical consultations and testing feasibility in development of guidelines and SDM tools across a wide range of clinical topics. The framework proved sufficiently generic to work across a wide array of conditions, treatments, and tests.

In the context of guideline development, the framework supported the panel in structured discussions, parallel to discussions about traditional health outcomes, addressing the impact of the recommended treatments or tests on patients' lives.

Our study has some important limitations. The framework is based on two sources from the UK and USA, potentially resulting in selection bias in the mapping and categorization as well as missed categories of practical

Table 1. Summary of the practical issues produced in the BMJ RapidRecs

	BMJ RapidRec	Medication routine	Tests & visits	Procedure & device	Recovery & adaptation	Coordination of care	Adverse #, interactions, antidote
1	Transcatheter vs. surgical aortic valve replacement	3	2	11	6	1	6
2	Low-intensity pulsed ultrasound (LIPUS) for bone healing	–	3	4	–	–	2
3	Arthroscopic surgery for degenerative knee arthritis and meniscal tears	–	2	8	3	-	-–
4	Antiretroviral therapy in pregnant women living with HIV	5	4	–	–	-	12
5	Corticosteroids for treatment of sore throat	2	1	–	–	–	1
6	Antibiotics for uncomplicated skin abscesses	4	1	–	–	–	3
7	Atraumatic (pencil-point) vs. conventional needles for lumbar puncture	–	–	4	–	1	1
8	Corticosteroid therapy for sepsis	3	3	–	1	–	–
9	Patent foramen ovale closure or drug therapy for management of cryptogenic stroke?	2	2	2	1	–	1
10	Prostate cancer screening	1	3	4	1	1	2
11	Oxygen therapy for acutely ill medical patients	2	–	–	2	2	2
12	Dual vs. single antiplatelet therapy	3	1	-	1	–	2
13	Subacromial decompression surgery for adults with shoulder pain	3	2	2	1	1	-
14	Thyroid hormones for subclinical hypothyroidism	1	2	–	–	–	2
15	Colorectal cancer screening	3	3	7	2	2	3
Total		32	29	42	18	8	37

issues. We aimed for a generic approach, so this framework is flexible for use across clinical decisions and settings. However, some aspects would have better been captured by an approach focused on specific conditions. We did not do a formal evaluation of feasibility but received positive feedback on the usefulness of the framework. More formal evaluation is needed to assess this more rigorously. Finally, the 15 selected categories are not entirely mutually exclusive, which leaves open the possibility that guideline panels will choose different categories for similar content across guidelines (e.g., adverse effects can also fit in the category of physical well-being).

4.2. Implications for practice

Most guideline panels overlook practical issues, or tend to limit this discussion only to very specific aspects related to burden of treatment [8]. They rarely appraise and incorporate descriptions of the patient experience across management alternatives. This situation is problematic.

Current trustworthiness standards for guideline development such as those from Institute of Medicine [40] or GRADE do not specifically include guidance on practical issues or burden of treatment. GRADE has recently introduced its evidence to the decision framework, in which practical issues may be conceptualized within domains of applicability and feasibility [41].

It remains unknown if guideline panels using this framework are explicit in considering practical issues. Similarly, international patient decision aid standards remain vague on their guidance on practical issues. They do require provision of information in sufficient detail for decision-making in the context of benefits and harms, but are not explicit on how to best meet that objective [42]. Using our framework of 15 prespecified categories of practical issues, driven by actual patient experience, may help improve the consideration of practical issues both in guidance for clinicians and in clinical encounters using SDM tools.

In addition to absence of specific guidance, there may be several other reasons why practical issues are largely omitted in guidelines. First, the production of rigorous evidence on burden of treatment remains challenging [9], although investigators are developing feasible assessment approaches [43]. Second, guideline developers may put more value on prognostic quantitative outcomes (e.g., risk of dying) than on "softer" (and often narrative) outcomes such as the extent to which interventions impact a person's quality of life, or otherwise add burden to their daily lives [44]. Third, guideline developers may not be familiar with the relevant evidence, how to find it, or believe its appraisal and incorporation is beyond the scope of their work. Finally, although this may be changing rapidly, until recently only a minority of guidelines—around 16% in a large sample assessed in 2012—include patients in their

Physical well-being	Emotional well-being	Pregnancy & nursing	Costs access	Food & drinks	Exercise & activities	Social life & relationships	Work & education	Travel & driving	Total
2	2	–	2	1	3	–	3	2	44
–	–	–	2	–	–	–	-	1	12
–	–	–	–	–	2	–	1	1	17
–	–	-	10	2	–	–	–	–	33
	1	2	1	1	–	–	–	–	9
–	–	5	2	4	–	–	–	–	19
2	–	–	2	–	–	–	–	–	10
–	1	1	1	1	–	–	–	–	11
–	–	1	1	1	2	–	1	1	15
1	1	-	1	1	1	2	1	1	21
–	2	-	2	–	–	–	–	–	12
1	1	1	1	2	1	–	1	–	15
–	–	–	2	-	4	–	2	1	18
–	3	–	1	2	–	–	–	–	11
3	1	–	1	3	2	2	3	1	36
9	12	10	29	18	15	4	12	8	283

panels [45]. Patient co-authors may highlight practical issues related to the recommended course of action.

Our experience in exploring how practical issues can be considered when issuing trustworthy and actionable guidelines—namely 15 BMJ RapidRecs at the time of assessment—had demonstrated that our generic framework contributes to a more systematic and comprehensive consideration of all relevant practical issues. Including patient partners throughout this process helps identify, highlight, and refine the most relevant practical issues [12,46]. Formally assessing the exact contribution of this novel approach on other elements of the guideline making process will, however, require further research.

Further research is needed to fully address the potential benefits and additional burden of implementing our framework in systematic reviews addressing treatment options. One main advantage of including practical issues, in complement to traditional evidence synthesis of benefits and harms, is to take advantage of the systematic and methodologically rigorous approach to search, screen, extract, and analyze best current evidence. Another advantage is to directly feed trustworthy evidence on practical issues in subsequent guidelines and SDM tools. The approach comes, however, with additional burden and added workload to systematic reviews already struggling to be timely and informative for end users. How to best incorporate evidence syntheses on practical issues warrants further exploration, for example, within the Cochrane collaboration.

5. Conclusion

Practical issues are central in decision-making and need to be considered as a part of evidence summaries informing guidelines and SDM tools. We developed a generic framework that allows for such inclusion that proved feasible and, when implemented in guidelines and SDM, enormously useful.

How such an approach can impact on the process and recommendations issued by guideline panels requires further research. Optimal production, evidence retrieval, and synthesis of relevant practical issues for decision-making will require further exploration. Finally, our work suggests a new avenue for research on how to support patients to make decisions that not only make intellectual sense, that is, they are evidence-based, but also make practical sense, that is, they are feasible given the patient's capacity to enact them in their daily lives without undue burden [47].

Acknowledgments

We thank Shaun Treweek for his help with user testing and valuable feedback on the manuscript. We thank Frankie Achille (interaction designer), Rob Fracisco (designer), Deno Vichas and Chris Degiere (programmers) for their contributions in development of the online authoring and

How to use the framework in development of guidelines, systematic reviews, and decision aids:

- For each clinical question, walk through each of the 15 categories to identify practical issues that are most meaningful to inform the decision.
- Identify potential data sources to feed these categories. These can include qualitative evidence, systematic surveys, patient experience databases, patient information leaflets, lay summaries of procedures, drug information, and so forth.
- Include patient partners throughout the process as well as any other relevant stakeholders.
- Test the final wording for readability and understanding among target audience, especially patients and caregivers.

publication platform (www.magicproject.org). We also thank Will Stahl-Timmins for his contributions in the BMJ RapidRecs project.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.10.002.

References

[1] Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017;(4):CD001431.
[2] Agoritsas T, Heen AF, Brandt L, Alonso-Coello P, Kristiansen A, Akl EA, et al. Decision aids that really promote shared decision making: the pace quickens. BMJ 2015;350:g7624.
[3] Stiggelbout AM, Weijden TVd, Wit MPTD, Frosch D, Légaré F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012;344:e256.
[4] Elwyn G, Quinlan C, Mulley A, Agoritsas T, Vandvik PO, Guyatt G. Trustworthy guidelines − excellent; customized care tools − even better. BMC Med 2015;13(1):199.
[5] Fog Heen A, Montori VM. Achieving care: promoting alternatives to industrial healthcare. Swiss Med Wkly 2019;149:w20115.
[6] Zeballos-Palacios CL, Hargraves IG, Noseworthy PA, Branda ME, Kunneman M, Burnett B, et al. Developing a conversation aid to support shared decision making: reflections on designing anticoagulation choice. Mayo Clin Proc 2019;94(4):686−96.
[7] Clifford AM, Ryan J, Walsh C, McCurtin A. What information is used in treatment decision aids? A systematic review of the types of evidence populating health decision aids. BMC Med Inform Decis Mak 2017;17(1):22.
[8] Dobler CC, Harb N, Maguire CA, Armour CL, Coleman C, Murad MH. Treatment burden should be included in clinical practice guidelines. BMJ 2018;363:k4065.
[9] Mair FS, May CR. Thinking about the burden of treatment. BMJ 2014;349:g6680.
[10] May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S. Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. BMC Health Serv Res 2014;14:281.
[11] Shippee ND, Shah ND, May CR, Mair FS, Montori VM. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. J Clin Epidemiol 2012;65:1041−51.
[12] Siemieniuk RA, Agoritsas T, Macdonald H, Guyatt GH, Brandt L, Vandvik PO. Introduction to BMJ Rapid recommendations. BMJ 2016;354:i5191.
[13] Vandvik PO, Brandt L, Alonso-Coello P, Treweek S, Akl EA, Kristiansen A, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. Chest 2013;144:381−9.
[14] DIPEX. Health Experiences research group. 2019. Available at https://www.healthtalk.org/HERG. Accessed March 3, 2020.
[15] Elwyn G, Lloyd A, Joseph-Williams N, Cording E, Thomson R, Durand MA, et al. Option Grids: shared decision making made easier. Patient Educ Couns 2013;90(2):207−12.
[16] Marrin KBK, Durand MA, Edwards A, Lloyd A, Thomas V, Elwyn G. Fast and frugal tools for shared decision making: how to develop Option Grids. Eur J Pers Cent Healthc 2013;1(1):240−5.
[17] Strauss A, Corbin J. The basics of qualitative research: grounded theory procedures and techniques. 4th ed. Newbury Park, CA: SAGE; 1990.
[18] Kennedy TJ, Lingard LA. Making sense of grounded theory in medical education. Med Educ 2006;40(2):101−8.
[19] Agoritsas T, Heen AF, Brandt L, Alonso-Coello P, Kristiansen A, Akl EA, et al. Decision aids that really promote shared decision making: the pace quickens. BMJ 2015;350:g7624.
[20] Heen AF, Vandvik PO, Brandt L, Guyatt G, Treweek S, Agoritsas T. A generic framework for shared-decision-making tools linked to evidence summaries, recommendations and clinical practice guideline: results from development and user-testing 2020.
[21] Lytvyn L, Guyatt GH, Manja V, Siemieniuk RA, Zhang Y, Agoritsas T, et al. Patient values and preferences on transcatheter or surgical aortic valve replacement therapy for aortic stenosis: a systematic review. BMJ Open 2016;6(9):e014327.
[22] Lytvyn L, Siemieniuk RA, Dilmitis S, Ion A, Chang Y, Bala MM, et al. Values and preferences of women living with HIV who are pregnant, postpartum or considering pregnancy on choice of antiretroviral therapy during pregnancy. BMJ Open 2017;7(9):e019023.
[23] Vernooij RWM, Lytvyn L, Pardo-Hernandez H, Albarqouni L, Canelo-Aybar C, Campbell K, et al. Values and preferences of men for undergoing prostate-specific antigen screening for prostate cancer: a systematic review. BMJ Open 2018;8(9):e025470.
[24] Vandvik PO, Otto CM, Siemieniuk RA, Bagur R, Guyatt GH, Lytvyn L, et al. Transcatheter or surgical aortic valve replacement for patients with severe, symptomatic, aortic stenosis at low to intermediate surgical risk: a clinical practice guideline. BMJ 2016;354:i5085.
[25] Poolman RW, Agoritsas T, Siemieniuk RA, Harris IA, Schipper IB, Mollon B, et al. Low intensity pulsed ultrasound (LIPUS) for bone healing: a clinical practice guideline. BMJ 2017;356:j576.
[26] Siemieniuk RAC, Harris IA, Agoritsas T, Poolman RW, Brignardello-Petersen R, Van de Velde S, et al. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: a clinical practice guideline. BMJ 2017;357:j1982.
[27] Siemieniuk RAC, Lytvyn L, Mah Ming J, Mullen RM, Anam F, Otieno T, et al. Antiretroviral therapy in pregnant women living with HIV: a clinical practice guideline. BMJ 2017;358:j3961.
[28] Aertgeerts B, Agoritsas T, Siemieniuk RAC, Burgers J, Bekkering GE, Merglen A, et al. Corticosteroids for sore throat: a clinical practice guideline. BMJ 2017;358:j4090.
[29] Vermandere M, Aertgeerts B, Agoritsas T, Liu C, Burgers J, Merglen A, et al. Antibiotics after incision and drainage for

uncomplicated skin abscesses: a clinical practice guideline. BMJ 2018;360:k243.
[30] Rochwerg B, Almenawer SA, Siemieniuk RAC, Vandvik PO, Agoritsas T, Lytvyn L, et al. Atraumatic (pencil-point) versus conventional needles for lumbar puncture: a clinical practice guideline. BMJ 2018;361:k1920.
[31] Kuijpers T, Spencer FA, Siemieniuk RAC, Vandvik PO, Otto CM, Lytvyn L, et al. Patent foramen ovale closure, antiplatelet therapy or anticoagulation therapy alone for management of cryptogenic stroke? A clinical practice guideline. BMJ 2018;362:k2515.
[32] Lamontagne F, Rochwerg B, Lytvyn L, Guyatt GH, Moller MH, Annane D, et al. Corticosteroid therapy for sepsis: a clinical practice guideline. BMJ 2018;362:k3284.
[33] Tikkinen KAO, Dahm P, Lytvyn L, Heen AF, Vernooij RWM, Siemieniuk RAC, et al. Prostate cancer screening with prostate-specific antigen (PSA) test: a clinical practice guideline. BMJ 2018;362:k3581.
[34] Siemieniuk RAC, Chu DK, Kim LH, Guell-Rous MR, Alhazzani W, Soccal PM, et al. Oxygen therapy for acutely ill medical patients: a clinical practice guideline. BMJ 2018;363:k4169.
[35] Prasad K, Siemieniuk R, Hao Q, Guyatt G, O'Donnell M, Lytvyn L, et al. Dual antiplatelet therapy with aspirin and clopidogrel for acute high risk transient ischaemic attack and minor ischaemic stroke: a clinical practice guideline. BMJ 2018;363:k5130.
[36] Vandvik PO, Lahdeoja T, Ardern C, Buchbinder R, Moro J, Brox JI, et al. Subacromial decompression surgery for adults with shoulder pain: a clinical practice guideline. BMJ 2019;364:l294.
[37] Bekkering GE, Agoritsas T, Lytvyn L, Heen AF, Feller M, Moutzouri E, et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. BMJ 2019;365:l2006.
[38] Helsingen LM, Vandvik PO, Jodal HC, Agoritsas T, Lytvyn L, Anderson JC, et al. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline. BMJ 2019;367:l5515.
[39] Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches the GRADE Working Group. BMC Health Serv Res 2004;4:38.
[40] Medicine Io. In: Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, editors. Clinical practice guidelines we can trust. Washington, DC: The National Academies Press; 2011:290.
[41] Alonso-Coello P, Schunemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 2016;353:i2016.
[42] Volk RJ, Llewellyn-Thomas H, Stacey D, Elwyn G. Ten years of the International Patient Decision Aid Standards Collaboration: evolution of the core dimensions for assessing the quality of patient decision aids. BMC Med Inform Decis Mak 2013;13(Suppl 2):S1.
[43] Spencer-Bonilla G, Quinones AR, Montori VM, International Minimally Disruptive Medicine W. Assessing the burden of treatment. J Gen Intern Med 2017;32:1141−5.
[44] Tinetti ME, Naik AD, Dodson JA. Moving from disease-centered to patient goals−directed care for patients with multiple chronic conditions: patient value-based care. JAMA Cardiol 2016;1(1):9−10.
[45] Kung J, Miller RR, Mackowiak PA. Failure of clinical practice guidelines to meet institute of medicine standards: two more decades of little, if any, progress. Arch Intern Med 2012;172:1628−33.
[46] Guyatt GAT, Lytvyn L, Siemieniuk R, Vandvik P. BMJ Rapid recommendations: a possible revolution in clinical practice guidelines. Can J Gen Intern Med 2019;14(1):6−12.
[47] Shippee ND, Shah ND, May CR, Mair FS, Montori VM. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. J Clin Epidemiol 2012;65:1041−51.

Appendix 1 Overview of Option Grids and HealthTalk Online themes

NO	Source	Topic
1.1	Option Grid	Angina
1.2	Option Grid	Atrial fibrillation
1.3	Option Grid	Breast cancer
1.4	Option Grid	Breast reconstruction
1.5	Option Grid	Chronic kidney disease
1.6	Option Grid	Crohn's disease
1.7	Option Grid	Employment and mental health
1.8	Option Grid	Epilepsy surgery
1.9	Option Grid	Epilepsy treatments when considering pregnancy
1.10	Option Grid	Glue ear
1.11	Option Grid	Heavy menstrual bleeding
1.12	Option Grid	HIV test
1.13	Option Grid	Implantable cardioverter defibrillator
1.14	Option Grid	Insulin treatment
1.15	Option Grid	Language options for deaf newborns
1.16	Option Grid	Localised prostate cancer - low risk
1.17	Option Grid	Newborn circumcision
1.18	Option Grid	Osteoarthritis of the Hip
1.19	Option Grid	Osteoarthritis of the knee: Managing knee pain and activity level
1.20	Option Grid	Osteoarthritis of the knee: Self Management
1.21	Option Grid	Ovarian Cancer Risk: Oophorectomy after the menopause
1.22	Option Grid	Ovarian Cancer Risk: Oophorectomy before the menopause
1.23	Option Grid	Prostate specific antigen (PSA) test
1.24	Option Grid	Sciatica - slipped (herniated) disc
1.25	Option Grid	Sore throat
1.26	Option Grid	Spinal stenosis
1.27	Option Grid	Testing for Down's syndrome in pregnancy:Amniocentesis
1.28	Option Grid	Testing for Down's syndrome in pregnancy: Down's syndrome screening

1.29	Option Grid	Tonsillitis
2.1	HealthTalk	ALS
2.2	HealthTalk	AIDS (acquired immune deficiency syndrome)
2.3	HealthTalk	Abortion for fetal abnormality
2.4	HealthTalk	Abortion: young people's views (in 'Sexual Health') (Young People)
2.5	HealthTalk	Ageing (link to later life category)
2.6	HealthTalk	Alcohol: young people's experiences (Young People)
2.7	HealthTalk	Alzheimer's disease: carers of people with
2.8	HealthTalk	Amyotrophic lateral sclerosis
2.9	HealthTalk	Anemia, sickle cell in young people (in 'Long term health conditions') (Young People)
2.10	HealthTalk	Angina (in 'Heart failure')
2.11	HealthTalk	Antenatal Screening
2.12	HealthTalk	Antenatal screening: for sickle cell and beta thalassaemia
2.13	HealthTalk	Antenatal screening: general
2.14	HealthTalk	Antidepressants
2.15	HealthTalk	Arthritis in young people (Young People)
2.16	HealthTalk	Arthritis in young people (in 'Long term health conditions') (Young People)
2.17	HealthTalk	Arthritis: rheumatoid
2.18	HealthTalk	Asperger's syndrome: parents' experiences
2.19	HealthTalk	Asthma in young people (in 'Long term health conditions') (Young People)
2.20	HealthTalk	Autism: Grandparents experiences
2.21	HealthTalk	Autism: Siblings experiences
2.22	HealthTalk	Autism: adults
2.23	HealthTalk	Autism: parents' experiences
2.24	HealthTalk	BME mental health
2.25	HealthTalk	BME mental health carers
2.26	HealthTalk	BRCA mutations (in 'Jewish Health')
2.27	HealthTalk	Baby: breastfeeding
2.28	HealthTalk	Baby: weaning
2.29	HealthTalk	Bereavement and grieving

2.30	HealthTalk	Bereavement due to suicide
2.31	HealthTalk	Bereavement due to traumatic death
2.32	HealthTalk	Beta thalassaemia: screening
2.33	HealthTalk	Biobanking
2.34	HealthTalk	Biobanking
2.35	HealthTalk	Bipolar affective disorder
2.36	HealthTalk	Birth after a caesarean
2.37	HealthTalk	Birth: life threatening conditions
2.38	HealthTalk	Blood Pressure
2.39	HealthTalk	Bomb blast: bereavement due to (in 'Bereavement due to traumatic death')
2.40	HealthTalk	Bones and Joints (category)
2.41	HealthTalk	Bones: osteoporosis
2.42	HealthTalk	Bowel Cancer
2.43	HealthTalk	Bowel Screening
2.44	HealthTalk	Breast Cancer in men
2.45	HealthTalk	Breast Cancer in women
2.46	HealthTalk	Breast Screening
2.47	HealthTalk	Breastfeeding
2.48	HealthTalk	CF (in 'Jewish Health')
2.49	HealthTalk	CF in young people (in 'Long term health conditions') (Young People)
2.50	HealthTalk	CFS in young people (in 'Long term health conditions') (Young People)
2.51	HealthTalk	CGIN
2.52	HealthTalk	CIN3
2.53	HealthTalk	Caesarean: making decisions about birth after
2.54	HealthTalk	Cancer
2.55	HealthTalk	Cancer (in Jewish health)
2.56	HealthTalk	Cancer screening for breast cancer
2.57	HealthTalk	Cancer: Bowel
2.58	HealthTalk	Cancer: Breast (men)
2.59	HealthTalk	Cancer: Breast (women)

2.60	HealthTalk	Cancer: Cervical (Cervix)
2.61	HealthTalk	Cancer: Colon
2.62	HealthTalk	Cancer: Colorectal
2.63	HealthTalk	Cancer: Leukaemia
2.64	HealthTalk	Cancer: Living with and beyond
2.65	HealthTalk	Cancer: Lymphoma
2.66	HealthTalk	Cancer: Pancreatic (Pancreas)
2.67	HealthTalk	Cancer: Penile (Penis)
2.68	HealthTalk	Cancer: Prostate
2.69	HealthTalk	Cancer: Rectal (Rectum)
2.70	HealthTalk	Cancer: Testicular
2.71	HealthTalk	Cancer: screening for bowel cancer
2.72	HealthTalk	Cancer: screening for cervical cancer
2.73	HealthTalk	Cancer: screening for prostate cancer
2.74	HealthTalk	Cancer: terminal (in ‘Living with Dying’)
2.75	HealthTalk	Cancer: terminal (in ‘carers of people with a terminal illness’)
2.76	HealthTalk	Cardiomyopathy (in ‘Heart failure’)
2.77	HealthTalk	Carers of people with dementia
2.78	HealthTalk	Carers: Alzheimer’s
2.79	HealthTalk	Carers: autism (children)
2.80	HealthTalk	Carers: congenital heart disease (children)
2.81	HealthTalk	Carers: dementia
2.82	HealthTalk	Carers: mental health (ethnic minority)
2.83	HealthTalk	Carers: terminal illness
2.84	HealthTalk	Caring for someone with a terminal illness
2.85	HealthTalk	Catheters
2.86	HealthTalk	Cerebrovascular accident (stroke)
2.87	HealthTalk	Cervical Cancer
2.88	HealthTalk	Cervical Glandular Intra-epithelial Neoplasia
2.89	HealthTalk	Cervical Intra-epithelial Neoplasia

2.90	HealthTalk	Cervical Screening
2.91	HealthTalk	Cervical abnormalities: CIN
2.92	HealthTalk	Cervical abnormalities: CIN3 and CGIN
2.93	HealthTalk	Chronic Fatigue Syndrome in young people (in ‘Long term health conditions’) (Young People)
2.94	HealthTalk	Chronic Pain
2.95	HealthTalk	Chronic health issues in young people (Young People)
2.96	HealthTalk	Chronic pain in young people (in ‘Long term health conditions’) (Young People)
2.97	HealthTalk	Clinical Trials
2.98	HealthTalk	Clinical trials: Parents’ experiences
2.99	HealthTalk	Coeliac disease in young people (in ‘Long term health conditions’) (Young People)
2.100	HealthTalk	Colon cancer (in ‘Colorectal cancer’)
2.101	HealthTalk	Colorectal Cancer
2.102	HealthTalk	Conditions that threaten women’s lives in childbirth & pregnancy
2.103	HealthTalk	Congenital heart disease in children: parents' experience
2.104	HealthTalk	Crohn’s Disease (in ‘Jewish Health’)
2.105	HealthTalk	Cystic fibrosis (in ‘Jewish Health’)
2.106	HealthTalk	Cystic fibrosis in young people (in ‘Long term health conditions’) (Young People)
2.107	HealthTalk	DCIS
2.108	HealthTalk	Death (in ‘Organ Donation’)
2.109	HealthTalk	Death: from traumatic circumstances
2.110	HealthTalk	Death: terminal illness
2.111	HealthTalk	Death: terminal illness (carers)
2.112	HealthTalk	Decisions (‘Shared decision-making’)
2.113	HealthTalk	Dementia: carers of people with
2.114	HealthTalk	Depression
2.115	HealthTalk	Depression and low mood in young people (Young People)
2.116	HealthTalk	Depression and recovery in Australia
2.117	HealthTalk	Depression: antidepressants
2.118	HealthTalk	Diabetes Type 2
2.119	HealthTalk	Diabetes type 1 in young people (Young People)

2.120	HealthTalk	Diabetes type 1 in young people (in ‘Long term health conditions’) (Young People)
2.121	HealthTalk	Donating organs
2.122	HealthTalk	Drugs and alcohol: young people’s experiences (Young People)
2.123	HealthTalk	Drugs: young people's experiences (Young People)
2.124	HealthTalk	Duchenne Muscular Dystrophy in young people (in ‘Long term health conditions’) (Young People)
2.125	HealthTalk	Ductal Carcinoma In Situ
2.126	HealthTalk	Ductal Carcinoma in Situ (DCIS)
2.127	HealthTalk	Dying
2.128	HealthTalk	Eating disorders in young people (Young People)
2.129	HealthTalk	Eczema in young people (in ‘Long term health conditions’) (Young People)
2.130	HealthTalk	Ending a pregnancy for fetal abnormality
2.131	HealthTalk	Epilepsy
2.132	HealthTalk	Epilepsy in young people (Young People)
2.133	HealthTalk	Epilepsy in young people (in ‘Long term health conditions’) (Young People)
2.134	HealthTalk	Ethnic minority mental health
2.135	HealthTalk	Ethnic minority mental health carers
2.136	HealthTalk	Experiences of antidepressants
2.137	HealthTalk	Experiences of depression and recovery in Australia
2.138	HealthTalk	Experiences of psychosis
2.139	HealthTalk	Factor 11 deficiency (Jewish health)
2.140	HealthTalk	Familial Dysautonomia (Jewish health)
2.141	HealthTalk	Fertility (in ‘Infertility’)
2.142	HealthTalk	Fertility (in ‘Menopause’)
2.143	HealthTalk	Fetal abnormality (ending a pregnancy)
2.144	HealthTalk	Fetal abnormality (screening for)
2.145	HealthTalk	Fire: bereavement due to (in ‘Bereavement due to traumatic death’)
2.146	HealthTalk	Gaucher disease (in ‘Jewish health’)
2.147	HealthTalk	Gay, lesbian, bisexual young people (in ‘Sexual health’) (Young People)
2.148	HealthTalk	Glaucoma (in ‘Jewish health’)
2.149	HealthTalk	HIV

2.150	HealthTalk	HIV (human immunodeficiency virus)
2.151	HealthTalk	HIV in young people (in ‘Long term health conditions’) (Young People)
2.152	HealthTalk	Having a grandchild on the autism spectrum
2.153	HealthTalk	Having a sibling on the autism spectrum
2.154	HealthTalk	Health and weight in young people (Young People)
2.155	HealthTalk	Heart attack
2.156	HealthTalk	Heart disease in children: parents' experience
2.157	HealthTalk	Heart failure
2.158	HealthTalk	Heart valve disease
2.159	HealthTalk	High blood pressure
2.160	HealthTalk	High dependency unit
2.161	HealthTalk	High dependency unit (family and friends experiences)
2.162	HealthTalk	Hypertension
2.163	HealthTalk	Hypoplastic left heart syndrome in young people (in ‘Long term health conditions’)
2.164	HealthTalk	ICU
2.165	HealthTalk	Immunisation
2.166	HealthTalk	Immunisation: parents' decisions
2.167	HealthTalk	Industrial accident: bereavement due to (in ‘Bereavement due to traumatic death’)
2.168	HealthTalk	Infertility
2.169	HealthTalk	Infertility
2.170	HealthTalk	Insomnia: later life
2.171	HealthTalk	Intensive care: Patients' experiences
2.172	HealthTalk	Intensive care: conditions that threaten women’s lives childbirth and pregnancy
2.173	HealthTalk	Intensive care: experiences of family & friends
2.174	HealthTalk	Intensive care: experiences of family & friends
2.175	HealthTalk	Intensive care: patients' experiences
2.176	HealthTalk	Jewish Health
2.177	HealthTalk	Juvenile Idiopathic Arthritis (JIA) (Young People)
2.178	HealthTalk	Juvenile Idiopathic Arthritis in young people (in ‘Long term health conditions’) (Young People)
2.179	HealthTalk	LGBT young people (in ‘Sexual health’) (Young People)

2.180	HealthTalk	Learning disability and health
2.181	HealthTalk	Leukaemia
2.182	HealthTalk	Life on the Autism spectrum
2.183	HealthTalk	Living with Dying
2.184	HealthTalk	Living with a urinary catheter
2.185	HealthTalk	Living with and beyond cancer
2.186	HealthTalk	Londoners' experiences of life-changing injuries
2.187	HealthTalk	Long-term health conditions (young people) (Young People)
2.188	HealthTalk	Lou Gehrig's Disease
2.189	HealthTalk	Lung Cancer
2.190	HealthTalk	Lymphoma
2.191	HealthTalk	M.E. in young people (in 'Long term health conditions') (Young People)
2.192	HealthTalk	MMR: parents' decisions
2.193	HealthTalk	MND
2.194	HealthTalk	Making decisions about birth after caesarean
2.195	HealthTalk	Manslaughter: bereavement due to (in 'Bereavement due to traumatic death')
2.196	HealthTalk	Medical research
2.197	HealthTalk	Medicine: antidepressants
2.198	HealthTalk	Menopause
2.199	HealthTalk	Menstruation (in 'Sexual Health') (Young People)
2.200	HealthTalk	Mental health
2.201	HealthTalk	Mental health: antidepressants
2.202	HealthTalk	Mental health: depression
2.203	HealthTalk	Mental health: ethnic minority carers' experiences
2.204	HealthTalk	Mental health: ethnic minority experiences
2.205	HealthTalk	Mental illness
2.206	HealthTalk	Mesangiocapillary glomerulonephritis type 2 in young people (in 'Long term health conditions') (Young People)
2.207	HealthTalk	Minor Stroke
2.208	HealthTalk	Morphea in young people (in 'Long term health conditions') (Young People)

2.209	HealthTalk	Motor Neurone Disease (MND)
2.210	HealthTalk	Murder: bereavement due to (in ‘Bereavement due to traumatic death’)
2.211	HealthTalk	Muscular dystrophy in young people (in ‘Long term health conditions’) (Young People)
2.212	HealthTalk	Myocardial infarction (heart attack)
2.213	HealthTalk	Nerves and brain
2.214	HealthTalk	Obesity: young people (in ‘Health & Weight’) (Young People)
2.215	HealthTalk	Organ donation
2.216	HealthTalk	Osteoporosis
2.217	HealthTalk	Ovarian Cancer
2.218	HealthTalk	Overweight: young people (in ‘Health & Weight’) (Young People)
2.219	HealthTalk	PSA test for prostate cancer
2.220	HealthTalk	Pain: chronic
2.221	HealthTalk	Pancreatic Cancer
2.222	HealthTalk	Parents of children on the Autism Spectrum
2.223	HealthTalk	Parents of children with congenital heart disease
2.224	HealthTalk	Parents: clinical trials in children and young people
2.225	HealthTalk	Parkinson's disease
2.226	HealthTalk	Patient/doctor relationship
2.227	HealthTalk	Penile Cancer
2.228	HealthTalk	Periods (in ‘Sexual health’) (Young People)
2.229	HealthTalk	Placental disorders (in ‘Conditions that threaten women’s lives in childbirth & pregnancy)
2.230	HealthTalk	Pre-eclampsia (in ‘Conditions that threaten women’s lives in childbirth & pregnancy)
2.231	HealthTalk	Pregnancy
2.232	HealthTalk	Pregnancy and children
2.233	HealthTalk	Pregnancy: antenatal screening
2.234	HealthTalk	Pregnancy: ending a pregnancy for fetal abnormality
2.235	HealthTalk	Pregnancy: life-threatening conditions
2.236	HealthTalk	Prostate Cancer
2.237	HealthTalk	Prostate cancer screening
2.238	HealthTalk	Prostate-specific antigen test

2.239	HealthTalk	Psychotic Depression
2.240	HealthTalk	Railway accident: bereavement due to (in ‘Bereavement due to traumatic death’)
2.241	HealthTalk	Rectal cancer
2.242	HealthTalk	Research participation
2.243	HealthTalk	Research participation: children and young people (Young People)
2.244	HealthTalk	Research participation: parents of children and young people
2.245	HealthTalk	Rheumatoid Arthritis
2.246	HealthTalk	Road accident: bereavement due to (in ‘Bereavement due to traumatic death’)
2.247	HealthTalk	Schizophrenia (in 'Experiences of Psychosis')
2.248	HealthTalk	Scoliosis in young people (in ‘Long term health conditions’) (Young People)
2.249	HealthTalk	Screening for prostate cancer
2.250	HealthTalk	Screening for sickle cell and beta thalassaemia
2.251	HealthTalk	Screening for unrecognised heart valve disease
2.252	HealthTalk	Screening: antenatal
2.253	HealthTalk	Screening: for cervical cancer
2.254	HealthTalk	Senior Loken Syndrome in young people (in ‘Long term health conditions’) (Young People)
2.255	HealthTalk	Sepsis (in ‘Conditions that threaten women’s lives in childbirth & pregnancy)
2.256	HealthTalk	Shared decision making
2.257	HealthTalk	Sickle cell anemia in young people (in ‘Long term health conditions’) (Young People)
2.258	HealthTalk	Sickle cell: screening
2.259	HealthTalk	Sleep problems in later life
2.260	HealthTalk	Smear test
2.261	HealthTalk	Stroke
2.262	HealthTalk	Suicide: bereavement due to
2.263	HealthTalk	TIA and Minor Stroke
2.264	HealthTalk	Tay Sachs (in ‘Jewish health’)
2.265	HealthTalk	Teenagers: arthritis (Young People)
2.266	HealthTalk	Teenagers: clinical trials (Young People)
2.267	HealthTalk	Teenagers: depression and low mood (Young People)
2.268	HealthTalk	Teenagers: diabetes type 1 (Young People)

2.269	HealthTalk	Teenagers: drugs and alcohol (Young People)
2.270	HealthTalk	Teenagers: epilepsy (Young People)
2.271	HealthTalk	Teenagers: long term health conditions (Young People)
2.272	HealthTalk	Teenagers: sexual health (Young People)
2.273	HealthTalk	Terminal illness
2.274	HealthTalk	Terminal illness: carers
2.275	HealthTalk	Termination for fetal abnormality
2.276	HealthTalk	Testicular Cancer
2.277	HealthTalk	Torsion dystonia (in ‘Jewish health’)
2.278	HealthTalk	Transient Ischaemic Attack
2.279	HealthTalk	Transplant
2.280	HealthTalk	Traumatic death
2.281	HealthTalk	UK Biobank
2.282	HealthTalk	Urinary catheter
2.283	HealthTalk	Vaccination: parents' decisions
2.284	HealthTalk	Valvular disease (in ‘Heart failure’)
2.285	HealthTalk	Valvular heart disease: screening
2.286	HealthTalk	Weaning
2.287	HealthTalk	Weight: young people (Young People)
2.288	HealthTalk	Young people: cancer (Young People)
2.289	HealthTalk	Young people: clinical trials (Young People)
2.290	HealthTalk	Young people: depression and low mood (Young People)
2.291	HealthTalk	Young people: diabetes type 1 (Young People)
2.292	HealthTalk	Young people: drugs and alcohol (Young People)
2.293	HealthTalk	Young people: epilepsy (Young People)
2.294	HealthTalk	Young people: health and weight (Young People)
2.295	HealthTalk	Young people: long-term health conditions (Young People)
2.296	HealthTalk	Young people: sexual health (Young People)
2.297	HealthTalk	Young people:arthritis (Young People)

Appendix 2

1. round of iteration	2. round of iteration	3. round of iteration	Final categories
Activities	Sport and activities		
Daily life	Daily life		
Sport and activities	activities	activities	EXERCICE & ACTIVITIES
Social life	Social life	relationships	SOCIAL LIFE & RELATIONSHIPS
Personal relationships	Personal relationships		
Travel	Travel and holidays	Travel and holidays	TRAVEL & DRIVING
Driving	Driving and travel		
Driving and travel	transport		
Daily medication routine	Daily medication routine	routine	MEDICATION ROUTINE
Doctor/hospital visits			
Screening			
Treatment location			
Associated testing	Associated testing and doctors/hospital visits	Assosiated testing and doctor visits	TESTS & VISITS
Associated treatment	Associated treatment		
Additional treatment	Additional treatment		
Medical devices	Medical devices, implants and stents	Aids, equipment and adaptations	PROCEDURE & DEVICE
Blood transfusion	Transfusions		
Anaestetic	device		
Procedures			
Surgery			
Support groups and finding information	Support groups and finding information		
Help from others	Coordination of care	Coordination of care	COORDINATION OF CARE
Adaptation	Adaptation		
Interactions	Minor adverse effects, interactions and antidot	Adverse effects and antidote	ADVERSE EFFECTS & INTERACTIONS, ANTIDOTE

Antidot			
Side effects	Side effects		
Pregnancy	Pregnancy and nursing	nursing	PREGNANCY & NURSING
Nursing			
Personal care	Personal care		
Physical health	Physical health	Physical health	PHYSICAL WELL-BEING
Recovery	Recovery	adaptation	RECOVERY & ADAPTATION
Research and trials	Research and trials		
Spirituality and religion	Spirituality and religion		
Emotional health	Emotional health	Emotional health	EMOTIONAL WELL-BEING
Smoking and recreational drugs	Smoking and recreational drugs		
Alcohol	Alcohol		
Food and digestion	Food, drink and digestion	Food, drink and digestion	FOOD & DRINKS
Costs	Finances and benefits		
Finances and benefits	Finances, costs and benefits	Financial costs and support	COSTS & ACCESS
Income			
Work	Work, career and education		
Education		Work and career	WORK & EDUCATION

Article 4

OPEN ACCESS

Patient values and preferences on valve replacement for aortic stenosis: a systematic review

Anja Fog Heen ,[1] Lyubov Lytvyn,[2] Michael Shapiro,[3] Gordon Henry Guyatt,[2] Reed Alexander Cunningham Siemieniuk,[2] Yuan Zhang,[2] Veena Manja,[4,5] Per Olav Vandvik,[6] Thomas Agoritsas[2,7]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/heartjnl-2020-318334).

[1]Department of Medicine, Innlandet Hospital Trust Gjøvik Hospital, Brumunddal, Norway
[2]Department of Health Research Methods, Evidence, and Impact, McMaster University Faculty of Health Sciences, Hamilton, Ontario, Canada
[3]Chicago, Illinois, USA
[4]Department of Surgery, University of California Davis, Sacramento, California, USA
[5]Department of Medicine, Veterans Affairs Northern California Health Care System, Mather, California, USA
[6]Institute of Health and Society, University of Oslo, Oslo, Norway
[7]Division of General Internal Medicine, Department of Medicine, University Hospitals of Geneva, Geneva, Switzerland

Correspondence to
Dr Anja Fog Heen, Dept. of Medicine, Innlandet Hospital Trust Gjøvik Hospital, 2819 Gjøvik, Norway;
anjaheen@gmail.com

AFH and LL contributed equally.

Received 24 September 2020
Revised 15 January 2021
Accepted 18 January 2021
Published Online First 9 February 2021

Check for updates

To cite: Heen AF, Lytvyn L, Shapiro M, *et al*. *Heart* 2021;**107**:1289–1295.

ABSTRACT

The review aims to summarise evidence addressing patients' values, preferences and practical issues on deciding between transcatheter aortic valve insertion (TAVI) and surgical aortic valve replacement (SAVR) for aortic stenosis. We searched databases and grey literature until June 2020. We included studies of adults with aortic stenosis eliciting values and preferences about treatment, excluding medical management or palliative care. Qualitative findings were synthesised using thematic analysis, and quantitative findings were narratively described. Evidence certainty was assessed using CERQual (Confidence in the Evidence from Reviews of Qualitative Research) and GRADE (Grading of Recommendations Assessment, Development and Evaluation). We included eight studies. Findings ranged from low to very low certainty. Most studies only addressed TAVI. Studies addressing both TAVI and SAVR reported on factors affecting patients' decision-making along with treatment effectiveness, instead of trade-offs between procedures. Willingness to accept risk varied considerably. To improve their health status, participants were willing to accept higher mortality risk than current evidence suggests for either procedure. No study explicitly addressed valve reintervention, and one study reported variability in willingness to accept shorter duration of known effectiveness of TAVI compared with SAVR. The most common themes were desire for symptom relief and improved function. Participants preferred minimally invasive procedures with shorter hospital stay and recovery. The current body of evidence on patients' values, preferences and practical issues related to aortic stenosis management is of suboptimal rigour and reports widely disparate results regarding patients' perceptions. These findings emphasise the need for higher quality studies to inform clinical practice guidelines and the central importance of shared decision-making to individualise care fitted to each patient.

INTRODUCTION

Severe aortic stenosis is a common valvular disease occurring among approximately 3% of people over 75 years old that results in significant morbidity and mortality.[1] With increasing severity of stenosis, patients often experience chest pain, syncope and heart failure.[2] Treatment options include surgical aortic valve replacement (SAVR) or a minimally invasive approach, transcatheter aortic valve insertion (TAVI). Benefits of TAVI include shorter hospital stay and quicker recovery; however, long-term outcome data are scarce but emerging.[3]

In 2016 a BMJ Rapid Recommendations guideline (BMJ RapidRecs) was published regarding the choice of TAVI versus SAVR for patients with aortic stenosis at low to intermediate surgical risk.[4] To inform the guideline, a systematic review addressing patient values and preferences was conducted.[5] Since 2016, new trials with longer follow-up have been published,[6 7] requiring updated evidence synthesis and guidance. This article is an update of the previous review of patient values and preferences about TAVI versus SAVR.[5]

METHODS

We followed the MOOSE (Meta-analyses Of Observational Studies in Epidemiology) checklist (online supplemental appendix 1). The protocol was registered at PROSPERO (International Prospective Register of Systematic Reviews) (CRD42016041907).

Search strategy

We searched MEDLINE, EMBASE and PsycINFO via OVID, using a combination of keywords and subject headings for 'aortic stenosis' and 'valve replacement', as well as a validated methodological search filter for values and preferences studies.[8] We updated the previous search until 16 June 2020 (online supplemental appendix 2), without language or publication status restrictions. We searched for grey literature via relevant conference abstracts, theses and dissertations (using the keywords 'aortic stenosis' and 'preference' or 'experience'), and the reference lists of eligible studies.

Study selection

We included studies with participants ≥18 years with aortic stenosis whose values and preferences related to the decision to undergo TAVI or SAVR were elicited. We considered values and preferences as 'the relative importance patients placed on the outcomes' for treatment decisions.[9] We excluded studies not reporting original data, case reports, studies reporting health-related quality of life before and after treatment, and studies that transformed quality of life measures into utility values, because quality of life was assessed in the associated systematic review of treatment effectiveness informing the BMJ RapidRecs.[3] Our initial review[5] did not include studies reporting values and preferences focused solely on medical management or

palliative care of aortic stenosis. We therefore did not include them in this update and focused solely on TAVI and SAVR.

Data collection and synthesis

Two authors (AFH, LL) independently screened titles and abstracts using prespecified criteria after conducting calibration exercises. The authors reviewed full-text articles independently and in duplicate and resolved disagreements by discussion or consultation with a third reviewer (TA). We contacted the authors of two abstracts that were ultimately excluded and corresponded with two authors of included studies for further information.

Two reviewers (AFH, LL) independently abstracted participant demographics, clinical characteristics, methods and findings. We conducted thematic analysis on qualitative results,[10] coding and synthesising primary quotations from study participants and author-reported summaries and themes. Across eligible studies, we also abstracted patient-important practical issues (ie, how a treatment can affect patients' daily life) related to decisions to undergo treatment and categorised findings using a developed generic framework, described elsewhere.[11] The review authors resolved disagreements through discussion or by consulting a third party (TA).

Quality assessment

For studies reporting qualitative outcomes, we assessed study quality using the qualitative research checklist of the Critical Appraisal Skills Programme.[12] For studies reporting quantitative outcomes, we assessed risk of bias using the instrument developed by Zhang *et al*,[13] appraising the following domains: study population, measurement and data analysis.

Certainty of evidence

Beyond quality assessments of each study, we assessed the overall certainty of evidence using Grading of Recommendations Assessment, Development and Evaluation (GRADE) for quantitative findings[13 14] and Confidence in the Evidence from Reviews of Qualitative Research (CERQual) for qualitative findings.[15] We rated certainty of evidence as high, moderate, low or very low for each finding. Findings started at high certainty and rated them down if there were concerns in one or more domains.[16] For CERQual, certainty could be rated down for methodological limitations, coherence, adequacy and relevance.[15] For GRADE, certainty could be rated down for risk of bias, inconsistency, indirectness, imprecision and publication bias.[13 14]

Incorporation into BMJ RapidRecs

The BMJ RapidRecs are developed in a collaboration between the not-for-profit MAGIC Evidence Ecosystem Foundation[17] and The BMJ.[18] Recommendations and associated reviews are updated given potentially practice-changing new evidence,[4] and this update is part of this process. Findings will be appraised by an independent guideline panel, without conflict of interests, including patient partners, front-line clinicians and methodologists working together to translate emerging research to user-friendly and trustworthy recommendations, evidence summaries and tools for shared decision-making.[4 19]

RESULTS

We identified 1230 unique titles and abstracts and reviewed 51 in full text (figure 1). Eight studies, reported in ten articles, were deemed eligible, with new six studies since the original review.[20–25] Study findings are described narratively and include exemplar quotes from patients in the primary studies when available. Quantitative results are presented in table 1. Further details of the qualitative results are reported in online supplemental appendix table 5.

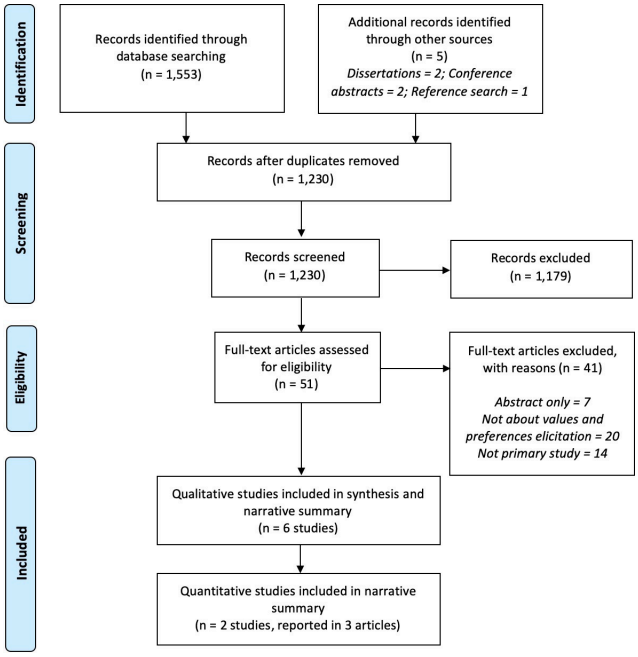


Figure 1 PRISMA study flow diagram. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Study characteristics

Studies were conducted in Canada, Norway, Sweden and USA (table 2). Of the quantitative studies, the sample sizes were 219[25] and 439.[26] Of the qualitative studies, one study included 333 participants,[20] while the others ranged from 10 to 46 participants.[21–24 27] Authors' conflicts of interest and study funding were variably reported. Two studies reported funding from a TAVI valve manufacturer (online supplemental appendix table 2).[20 25] All but one study included participants with confirmed severe aortic stenosis,[20–24 26 27] and the remaining included participants with self-reported diagnosis without specifying severity.[25] Participants were balanced in sex and were on average between 75 and 86 years old, except one study with 26% of participants aged 19–59 years old.[25] Surgical risk was variable across studies,[21 23 24 26 27] unknown or unspecified.[20 22 25]

Study quality and certainty of evidence

Most of the qualitative studies had methodological limitations,[20–22 24 27] the most common issues being inappropriate or unclear sampling and recruitment strategy, limited description of data analysis and strategies to enhance study rigour (online supplemental appendix table 3). For the quantitative studies, there were limitations in almost all domains, with the most concern being about participant selection, outcome presentation and data analysis (online supplemental appendix table 4). The certainty of findings ranged from low to very low (table 1, online supplemental appendix 1). The majority of studies assessed values and preferences on one intervention alone.

Table 1 GRADE summary of findings

Health state/outcome **(timeframe)**	Study design **(n=participants)**	Estimate of effect, **mean (SD) unless otherwise stated**	Certainty of evidence	Interpretation of finding
Mortality (30 days)	Adaptive swing weighting (109*)	Maximum acceptable increase in risk in exchange from SAVR to TAVI =3.7% (3.0)†.	Very low§¶**	The risk willingness of trading a reduction in mortality risk (30 days) for a less invasive procedure was uncertain and highly variable.
Mortality and aortic stenosis-related symptoms and concerns (lifetime)	Standard gamble (429)	Median risk willingness=25% (IQR 25%–50%). No risk (0%)=104 (23%). Low risk (0%–8%)=26 (6%). High risk (>8%–50%)=224 (51%). Prohibitive risk (>50%–95%)=68 (15%). 95%–100%=17 (4%).	Low§¶	The risk willingness of trading a reduction in mortality risk for full health with the procedure is highly variable among participants and across risk groups.
Disabling non-fatal stroke (30 days)	Adaptive swing weighting (110*)	Maximum acceptable increase in risk in exchange from SAVR to TAVI=6.7% (5.7)†.	Very low§¶**	The risk willingness of trading a reduction in risk of disabling stroke for a less invasive procedure was uncertain and highly variable.
Independence (30 days)	Adaptive swing weighting (131*)	Maximum acceptable reduction in benefit in exchange from SAVR to TAVI=13.9% (11.8)†.	Very low§¶**	The risk willingness of trading an increase of independence for a less invasive procedure was uncertain and highly variable.
Requirement for dialysis (1 year)	Adaptive swing weighting (132*)	Maximum acceptable increase in risk in exchange from SAVR to TAVI=6.2% (5.6)†.	Very low§¶**	The risk willingness of trading a reduction in the requirement for dialysis at 1 year for a less invasive procedure was uncertain and highly variable.
New permanent pacemaker (1 year)	Adaptive swing weighting (131*)	Maximum acceptable increase in risk in exchange from SAVR to TAVI=7.0% (5.7)‡.	Very low§¶**	The risk willingness of trading a reduction in permanent pacemaker insertion for a less invasive procedure was uncertain and highly variable.
Time over which the procedure has been proven to work	Adaptive swing weighting (131*)	Maximum acceptable decrease in duration that the procedure is known to work in exchange from SAVR to TAVI=17.4 years (16.9)‡.	Very low§¶**	The risk willingness of trading the expected duration or a new valve for a less invasive procedure was uncertain and highly variable.

*The total sample size was 219 participants, but they were not presented with all outcomes.
†Minimum acceptable reduction in benefit in exchange for reducing procedure invasiveness from 'invasive' to 'minimally invasive'.
‡Maximum acceptable increase in risk in exchange for reducing procedure invasiveness from 'invasive' to 'minimally invasive'.
§Serious risk of bias.
¶Serious imprecision.
**Serious indirectness.
GRADE, Grading of Recommendations Assessment, Development and Evaluation; SAVR, surgical aortic valve replacement; TAVI, transcatheter aortic valve insertion.

Values and preferences regarding outcomes of treatment

None of the studies presented participants' values and preferences based on a comprehensive assessment of the beneficial and adverse outcomes related to SAVR versus TAVI, nor did any studies report patient preferences about choosing between TAVI versus SAVR. Instead, studies focused on preferences about a selection of attributes in isolation. None of the studies addressed the lifelong management of treatment of valve failure.

Durability and valve reintervention

No study directly addressed how participants valued valve failure nor the risk and timing of reintervention. One study provided very low certainty of evidence regarding preferences about durability, illustrating considerable variability in patients' willingness to accept a shorter duration of the effectiveness of TAVI compared with SAVR.[25] A subgroup analysis suggested this variability may be partly explained by the fact that participants under 60 years old were more concerned with valve duration than those over 60.[25]

Mortality and risk willingness related to the decision to undergo treatment

All studies addressed mortality.[20–23 26 27] Studies did not explicitly distinguish between perioperative mortality, mortality from natural progression of disease or all-cause mortality. Participants viewed declining treatment to be worse than accepting the risk related to the procedure,[23] and thus were commonly willing to accept a high perioperative mortality risk. The importance of mortality can be illustrated by the following participant quote:

> And if I would have turned it [TAVI assessment] down, I mean, who knows how long I would last? Not much longer, probably, you know.[27]

Risk willingness varied considerably.[26] Overall, participants were willing to accept a higher mortality risk than current evidence suggests for TAVI, regardless of the fact that actual mortality risk is lower with TAVI than SAVR.[6 7 25]

For some participants, increasing life expectancy was more commonly a preference expressed by their families than by themselves,[23 24] as exemplified by the following quote:

> We did not discuss it too much the physician and I either. (…) He just asked if I wanted (the treatment) and I accepted. (…) I did it for the others' sake as well.[23]

Quality of life as reasons to undergo treatment

All but one study[26] reported improvements in health-related quality of life domains (eg, physical function, emotional well-being) as reasons to undergo treatment.[20 21 23–25 27] Common themes were desire for symptom relief and improved function. Respondents often described improved quality of life as the ability to do a specific activity, to regain or maintain independence,[21–24 27] to return to activities they had given up and to reconnect with their social network.[27] A participant's perspective was:

> We belong to a walking club […], but I've quit that in the last probably 3 or 4 months because I just couldn't keep up with them. They'd go and I said, "Well, I'll go half way" and they still got back before I did, so I said, "I guess I'll quit because it just hinders you guys."[27]

The desire to achieve the best possible health was closely intertwined with participants' ability to fulfil obligations towards family and friends and day-to-day activities when deciding on treatment.[20 21 23 24 27] Participants expressed not wanting to be a burden to relatives.[20 23 27] A participant noted the effect of their declining health on their partner, expressing:

> And this is passed on to my wife, of course. If I can't take [wife] to dance, she doesn't get to go either, you know what I mean?[27]

Concerns of pain

Pain was a concern with SAVR. One participant stated:

> Quite a bit of pain in the chest area, having your chest cracked open.[22]

Table 2 Study and participant demographics

Study	Country	Study design	Sample size	Patient population	Previous TAVI/SAVR	Age (years), mean (SD)	Sex (male), n (%)	Surgical risk STS score, median (IQR)	Heart failure symptoms NYHA class*, n (%)	Quality of life, symptoms, function, n
Quantitative studies										
Marsh *et al*[34]	USA	Adapted swing weighting	219†	Self-reported aortic stenosis; received treatment within 10 years or experiencing limitations in their physical activity due to aortic stenosis.	Undergone aortic stenosis treatment (unspecified)=80.4%	19–39=26.5%; 40–59=33.8%; 60–74=25.1%; 75–89=13.2%; 90+=1.4%	91 (41.6)	NR	Class I=78 (35.6%); class II=101 (46.1%); class III=40 (18.3%)	General health (past week): very good=55; good=85; fair=65; poor=13; very poor=1
Hussain *et al*[26]	Norway	Standard gamble	439	Severe aortic stenosis; referred for aortic valve treatment.	NR	75 (11)	264 (60)	11.9% (7.50%–17.10%)	Class I=11 (13%); class II=43 (50%); class III/IV=46 (53%)‡	SF-36§¶ physical component score=38 (10); mental component score=49 (10)
Qualitative studies										
Coylewright *et al*[24]	USA	Interview	46	Severe aortic stenosis; assessed for aortic valve treatment.	NR	68–74=5; 75–89=29; 90+=12	25 (54.3)	9% (4.9%)§	NR	KCCQ-12¶** 36 (4–76)
Olsson *et al*[21]	Sweden	Interview	24	Severe aortic stenosis.	NR	80.7 (7.4)	15 (62.5)	NR	Class III=11 (46%); class IV=13 (54%)	NR
Skaar *et al*[23]	Norway	Interview	10	Severe aortic stenosis.	NR	70–79=3; 80–89=7	4 (40)	Logistic EuroSCORE <10=2; 10–20=7; >20=1	Class I=1 (10%); class II=7 (70%); class III=2 (20%)	SPPB fit=3; intermediate=6; frail=1
Lauck *et al*[27]	Canada	Interview	15	Severe symptomatic aortic stenosis.	Undergone cardiac surgery (unspecified)=6	86 (75–92)††	9 (60)	6.4% (2.6%–16.3%)	Class II=11 (73%), *others not specified*	All but one participant were able to complete all activities of daily living.
Ontario Health Technology Assessment Series[22]	Canada	Interview	10	Aortic stenosis.	Undergone TAVI=9, undergone SAVR=1	NR	NR	NR	NR	NR
Frank *et al*/Styra *et al*[20 35]	Canada	Interview	333	Patients with aortic stenosis considering treatment options.	None	80.5 (52–97)††	181 (54.5)	NR	NR	NR

*NYHA class I=no symptoms and no limitation in ordinary physical activity; class II=mild symptoms and slight limitation during ordinary activity; class III=marked limitation in activity due to symptoms, even during less than ordinary activity, comfortable only at rest; class IV=severe limitations, experiences symptoms even while at rest, mostly bedbound patients.
†Baseline variables reported for 219 participants, but outcome data are for 109–132 participants (ie, not all participants were asked about all outcomes).
‡NYHA classification only reported for 86 of 439 participants. Class III and IV grouped together but only 2% were class IV.
§Mean, SD.
¶Maximum score=100.
**Median, range.
††Mean, range.
KCCQ-12, Kansas City Cardiomyopathy Questionnaire; NR, not reported; NYHA, New York Heart Association; SAVR, surgical aortic valve replacement; SF-36, Short-Form 36; SPPB, Short Physical Performance Battery; STS, Society of Thoracic Surgeons; TAVI, transcatheter aortic valve insertion.

Those who had TAVI described minimal pain, with a participant saying:

> And I didn't have any pain afterwards at all. I didn't even know that I'd had incisions in my groin. I just didn't know it was there. It was amazing.[22]

Acute kidney injury and stroke

Two studies addressed acute kidney injury.[24 25] One addressed the possibility of dialysis as a patient concern related to potential TAVI complications.[24] The other study provided evidence regarding patients' willingness to accept the risk of needing dialysis within 1 year after the procedure.[25] Patients in one study frequently expressed that they were afraid of the possibility of a stroke.[21]

Practical issues related to valve replacement

Several studies addressed participants' concerns regarding practical issues, such as invasiveness, length of hospital stay and recovery time.[21 22 25] Regarding TAVI, one participant stated:

> It's easy by comparison to an open-heart surgery. That is just a big plus. Can you imagine having your chest cut right open and taking months to recover?[27]

Overall, patients reported the longer hospital stay and recovery time with SAVR, compared with TAVI, as a major concern.[22] None of the studies mentioned the need for—and accessibility to—cardiac rehabilitation after SAVR or TAVI.

Decision-making process and support

Respondents perceived physicians as essential sources of information and decision-making guidance and as facilitators of referral for TAVI, and participants stressed the importance of a trusting relationship with their physician(s).[21 23] The experience of receiving rigorous advice from their physician was important in decision-making, illustrated by the following participant quote:

> When I'm with my doctor, I believe he is competent enough just to see what my problem is and how it can be treated.[27]

A number of studies, however, reported the possibility that physicians might not act in a trustworthy way, which motivated participants to seek a second opinion.[21 23 27] Overall, participants took into account a variety of medical, functional and social factors in their decision-making.[20–24 27]

Accessibility and cost of the procedure

Participants who lived away from hospitals that offered the procedure reported greater difficulty accessing TAVI.[22] Several studies reported participants' concern about burden of personal cost due to travel, meals and accommodation,[22 27] exemplified by the following participant quote:

> My family wanted to be there when I had the surgery, so there was … overnight accommodation … and meals, and so on. And someone to help with the driving … It was basically … personal expenses.[22]

Given the expected shorter length of hospital stay with TAVI, some patients perceived these costs to be much lower than with SAVR.[22]

DISCUSSION

Our search identified eight studies that examined patients' values, preferences and practical issues related to aortic stenosis treatment.[20–25] They provided limited evidence regarding how participants explicitly value and balance benefits and harms associated with TAVI and SAVR.[20–27] Most studies addressed only TAVI, and those that addressed both TAVI and SAVR did not specify the information they had provided to participants about the relative merits and burdens of the two procedures. Study participants were concerned about treatment complications, and willingness to accept procedural risk varied considerably. Participants of the qualitative studies rarely reported perspectives regarding specific outcomes (eg, stroke), but rather highlighted and valued fast return to function, independence, and social and daily activities. In terms of decision-making in general, trust in physicians and medical teams was very important in the decision. For practical issues, accessibility of the procedure and associated costs (eg, travel for themselves and their caregivers) were commonly reported.

Recent randomised trials,[28 29] as well as previously published trials with longer follow-up,[6 7] have added up to the current body of evidence comparing TAVI and SAVR.[3] Taken together, this evidence tends to show substantial short-term benefits of TAVI on outcomes important to patients with severe aortic stenosis at low and intermediate preoperative surgical risk, along with a substantially reduced burden of treatment.

However, valve durability with TAVI remains uncertain over the longer term due to limited follow-up compared with SAVR. An important concern is that TAVI might require valve reintervention much earlier than SAVR. This issue is particularly crucial for younger populations, as their life expectancy puts them at higher risk of needing one—or more—reinterventions. Unfortunately, our systematic review provides limited evidence on how patients may value differing valve durability and the risk of reinterventions. Indeed, only one study reported on patients' perceptions about willingness to accept a shorter duration of effectiveness of TAVI compared with SAVR, showing important variability. This study had methodological limitations and was funded by a valve device company.[25]

Another issue that varies with age is how the relative effects of TAVI translate in terms of absolute differences: because patients present a higher baseline mortality, TAVI is likely going to result in larger absolute reductions in deaths among older rather than younger patients. The balance of benefits and harms of TAVI versus SAVR will thus highly depend on age—as a proxy of life expectancy—as well as comorbidities.[4] The age or baseline risk threshold at which patients would consider the balance between benefits, harms and burden (including the risk of reintervention) in favour of either TAVI or SAVR remains thus far insufficiently explored. Current inference on these issues is further limited by the fact that several studies asked patients to trade off outcomes without basing the options on current best evidence. For example, they present unrealistic outcome risk options that were beyond the range of actual risks reported in trials.[4 6 7] The trade-off of outcomes may thus be misinformed or even misguided in such studies. Even less explored are patients' values and preference regarding the possible sequence of valve interventions.

Strengths and limitations

Our review has several strengths. First, we prospectively registered the protocol and followed study reporting criteria. Second, we conducted a comprehensive search, including grey literature, up to June 2020. Third, we assessed study quality using recommended instruments,[9 12 15] as well as using standardised methods to address the overall certainty of evidence for both quantitative and qualitative findings.[13 14 30] Fourth, the inclusion of a

patient partner as a coauthor enriched the framework used for thematic analysis. Finally, we abstracted data regarding patient-important practical issues, which shed light on areas important for decision-making that are rarely included when developing guidelines.

Our review has also limitations. First, we excluded studies looking at health-related quality of life for patients with aortic stenosis before and after therapy because these studies do not directly report on patient preferences. Second, due to the considerable heterogeneity of the types of studies included, we were not able to explore potential differences in values and preferences for subgroups of participants. Finally, our review highlights limitations of current evidence in the field, and particularly the lack of data on key outcomes and practical issues which guideline panels and patients need to inform decision-making.

CONCLUSION AND AVENUES FOR FUTURE RESEARCH

In parallel to new evidence on the effectiveness and durability of interventions, we need higher quality evidence on patients' values and preferences on all key outcomes, as well as better insight on what practical issues matter most to them. Future studies should be conducted in a broad and representative array of patients with severe, symptomatic aortic stenosis with variable risk profiles and comorbidities. They should also be informed by current best evidence on benefits and harms, rather than hypothetical (or even implausible) effects. Evidence from real-life decision-making, for example by using encounter decision aids, may better capture actual values and preferences to inform stakeholders such as guideline developers.[19 31]

Another priority should be to identify key practical issues for decision-making. New frameworks have been proposed to better structure searching, evidence synthesis and inclusion in the guideline-making process of patient-important practical issues.[19 32 33] Indeed, in highly preference-sensitive decisions such as whether to undergo TAVI or SAVR, practical issues related to each intervention and how they may affect patients' daily life may dominate shared decision-making conversations.[9]

Key messages

What is already known on this subject?

- Transcatheter aortic valve insertion (TAVI) is increasingly offered as an alternative treatment option to surgical aortic valve replacement (SAVR) for severe, symptomatic aortic stenosis, but its long-term durability remains uncertain.
- There is limited evidence on values, preferences and practical issues that are important to patients with aortic stenosis regarding the trade-offs of benefits and harms of TAVI compared with SAVR.

What might this study add?

- We provide a critical appraisal of empirical evidence on values and preferences related to aortic stenosis treatment.
- Current evidence suggests there is considerable variability among patients' values and preferences regarding the outcomes associated with TAVI or SAVR, as well as regarding the duration that the procedure has been proven to be effective.
- To improve their health status, participants were willing to accept higher mortality risk than current evidence suggests for either procedure, although this evidence was of low to very low certainty.
- Overall, participants preferred minimally invasive procedures with a shorter hospital stay and recovery time and also reported concerns regarding postsurgical pain and costs.
- An important limitation of this evidence is that no study presented participants current best evidence on all benefits and risks for both procedures, including valve durability, when enquiring for their preferred option.

How might this impact on clinical practice?

- Discussions regarding individual patients' values and preferences, focusing on the key outcomes and practical issues identified in this paper, can support shared decision-making about the best aortic stenosis treatment option for patients.
- This evidence can also inform the updates of health technology assessment and clinical practice guidelines on TAVI and SAVR.

Acknowledgements We thank the members of the BMJ Rapid Recommendations panel for critical feedback on outcomes and selection that informed this systematic review.

Contributors AFH led and coordinated the project. TA and POV provided supervision. AFH, LL and TA screened the studies for eligibility. AFH and LL extracted the data, assessed study risk of bias and synthesised the data. AFH, LL and TA assessed the quality of the body of evidence. All study authors were involved in the interpretation and discussion of the results. AFH and LL drafted the manuscript, and all authors critically revised the manuscript. All authors approved the final version of the article. AFH is the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests AFH, LL, GG, RACS, TA, POV and YZ are members of the GRADE working group. YZ designed the risk of bias tool and the GRADE evaluation for values and preferences studies. There are no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement statement Outcomes of interest included for this review were established by a multidisciplinary guideline panel that included three patient partners. One patient partner (MMS) from the guideline panel was included as a coauthor of this study. MMS was involved in the interpretation of study results and provided feedback on the manuscript.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Anja Fog Heen http://orcid.org/0000-0002-8873-7028

REFERENCES

1 Osnabrugge RLJ, Mylotte D, Head SJ, *et al*. Aortic stenosis in the elderly: disease prevalence and number of candidates for transcatheter aortic valve replacement: a meta-analysis and modeling study. *J Am Coll Cardiol* 2013;62:1002–12.
2 Fau TJ, Fau HO, Fau SF, *et al*. Spontaneous course of aortic valve disease. (0195-668X (print)).

3 Siemieniuk RA, Agoritsas T, Manja V, *et al*. Transcatheter versus surgical aortic valve replacement in patients with severe aortic stenosis at low and intermediate risk: systematic review and meta-analysis. *BMJ* 2016;354:i5130.
4 Vandvik PO, Otto CM, Siemieniuk RA, *et al*. Transcatheter or surgical aortic valve replacement for patients with severe, symptomatic, aortic stenosis at low to intermediate surgical risk: a clinical practice guideline. *BMJ* 2016:i5085.
5 Lytvyn L, Guyatt GH, Manja V, *et al*. Patient values and preferences on transcatheter or surgical aortic valve replacement therapy for aortic stenosis: a systematic review. *BMJ Open* 2016;6:e014327.
6 Gleason TG, Reardon MJ, Popma JJ, *et al*. 5-year outcomes of self-expanding transcatheter versus surgical aortic valve replacement in high-risk patients. *J Am Coll Cardiol* 2018;72:2687–96.
7 Søndergaard L, Ihlemann N, Capodanno D, *et al*. Durability of Transcatheter and Surgical Bioprosthetic Aortic Valves in Patients at Lower Surgical Risk. *J Am Coll Cardiol* 2019;73:546–53.
8 Selva A, Solà I, Zhang Y, *et al*. Development and use of a content search strategy for retrieving studies on patients' views and preferences. *Health Qual Life Outcomes* 2017;15:126.
9 Zhang Y, Coello PA, Brożek J, *et al*. Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. *Health Qual Life Outcomes* 2017;15:52.
10 Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;8:45.
11 Heen AF, Vandvik PO, Brandt L, *et al*. A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines. *J Clin Epidemiol* 2021;129:104–13.
12 Critical. APS. CASP qualitative checklist. Available: https://casp-uk.net/casp-tools-checklists/2019 [Accessed 1 Oct 2019].
13 Zhang Y, Alonso-Coello P, Guyatt GH, *et al*. GRADE guidelines: 19. assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. *J Clin Epidemiol* 2019;111:94–104.
14 Zhang Y, Coello PA, Guyatt GH, *et al*. GRADE guidelines: 20. assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. *J Clin Epidemiol* 2019;111:83–93.
15 Lewin S, Bohren M, Rashidian A, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 2: how to make an overall CERQual assessment of confidence and create a summary of qualitative findings table. *Implement Sci* 2018;13:10.
16 Colvin CJ, Garside R, Wainwright M, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 4: how to assess coherence. *Implement Sci* 2018;13:13.
17 MAGIC, 2018. Available: http://magicproject.org/
18 Siemieniuk RA, Agoritsas T, Macdonald H, *et al*. Introduction to BMJ Rapid Recommendations. *BMJ* 2016;354:i5191.
19 Agoritsas T, Heen AF, Brandt L, *et al*. Decision AIDS that really promote shared decision making: the PACE quickens. *BMJ* 2015;350:g7624.
20 Frank D, Kennon S, Bonaros N, *et al*. Trial protocol for the validation of the 'Toronto Aortic Stenosis Quality of Life (TASQ) Questionnaire' in patients undergoing surgical aortic valve replacement (SAVR) or transfemoral (TF) transcatheter aortic valve implantation (TAVI): the TASQ registry. *Open Heart* 2019;6:e001008.
21 Olsson K, Näslund U, Nilsson J, *et al*. Experiences of and coping with severe aortic stenosis among patients waiting for transcatheter aortic valve implantation. *J Cardiovasc Nurs* 2016;31:255–61.
22 Ontario Health Technology Assessment Series. Transcatheter aortic valve implantation in patients with severe, symptomatic aortic valve stenosis at intermediate surgical risk. health quality Ontario, 2018. Available: https://www.hqontario.ca/evidence-to-improve-care/health-technology-assessment/reviews-and-recommendations/transcatheter-aortic-valve-implantation-in-patients-with-severe-symptomatic-aortic-valve-stenosis-at-intermediate-surgical-risk
23 Skaar E, Ranhoff AH, Nordrehaug JE, *et al*. Conditions for autonomous choice: a qualitative study of older adults' experience of decision-making in TAVR. *J Geriatr Cardiol* 2017;14:42–8.
24 Coylewright M, Palmer R, O'Neill ES, *et al*. Patient-defined goals for the treatment of severe aortic stenosis: a qualitative analysis. *Health Expect* 2016;19:1036–43.
25 Marsh K, Hawken N, Brookes E, *et al*. Patient-Centered benefit-risk analysis of transcatheter aortic valve replacement. *F1000Res* 2019;8:394.
26 Hussain AI, Garratt AM, Brunborg C, *et al*. Eliciting patient risk willingness in clinical consultations as a means of improving decision-making of aortic valve replacement. *J Am Heart Assoc* 2016;5:e002828.
27 Lauck SB, Baumbusch J, Achtem L, *et al*. Factors influencing the decision of older adults to be assessed for transcatheter aortic valve implantation: an exploratory study. *Eur J Cardiovasc Nurs* 2016;15:486–94.
28 Mack MJ, Leon MB, Thourani VH, *et al*. Transcatheter aortic-valve replacement with a balloon-expandable valve in low-risk patients. *N Engl J Med* 2019;380:1695–705.
29 Popma JJ, Deeb GM, Yakubov SJ, *et al*. Transcatheter aortic-valve replacement with a self-expanding valve in low-risk patients. *N Engl J Med* 2019;380:1706–15.
30 Lewin S, Booth A, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci* 2018;13:2.
31 van Beek-Peeters JJAM, van Noort EHM, Faes MC, *et al*. Shared decision making in older patients with symptomatic severe aortic stenosis: a systematic review. *Heart* 2020;106:647–55.
32 Heen AF, Vandvik PO, Agoritsas T, *et al*. Including practical issues and patient perspectives in rapid recommendations. *BMJ* 2018;363:k4871.
33 Heen AF, Vandvik PO, Brandt L, *et al*. A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines. *J Clin Epidemiol* 2021;129:104–13.
34 Marsh K, Hawken N, Brookes E. Patient-centered benefit-risk analysis of transcatheter aortic valve replacement [version 4; peer review: 2 approved]. *F1000Res* 2020;8:394.
35 Styra R, Dimas M, Svitak K, *et al*. Toronto aortic stenosis quality of life questionnaire (TASQ): validation in TAVI patients. *BMC Cardiovasc Disord* 2020;20:209.

Appendix Table of Content

Appendix 1. MOOSE Checklist for Meta-analyses of Observational Studies

Item No	Recommendation	Reported on Page No
Reporting of background should include		
1	Problem definition	3
2	Hypothesis statement	N/A
3	Description of study outcome(s)	4
4	Type of exposure or intervention used	4
5	Type of study designs used	4
6	Study population	4
Reporting of search strategy should include		
7	Qualifications of searchers (e.g., librarians and investigators)	3,4
8	Search strategy, including time period included in the synthesis and key words	3, 25-27
9	Effort to include all available studies, including contact with authors	3,4
10	Databases and registries searched	4
11	Search software used, name and version, including special features used (e.g., explosion)	N/A
12	Use of hand searching (e.g., reference lists of obtained articles)	4
13	List of citations located and those excluded, including justification	16, 28-30
14	Method of addressing articles published in languages other than English	N/A
15	Method of handling abstracts and unpublished studies	4
16	Description of any contact with authors	4
Reporting of methods should include		
17	Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	5
18	Rationale for the selection and coding of data (e.g., sound clinical principles or convenience)	5
19	Documentation of how data were classified and coded (e.g., multiple raters, blinding and interrater reliability)	5

20	Assessment of confounding (e.g., comparability of cases and controls in studies where appropriate)	N/A
21	Assessment of study quality, including blinding of quality assessors, stratification or regression on possible predictors of study results	5
22	Assessment of heterogeneity	N/A
23	Description of statistical methods (e.g., complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	N/A
24	Provision of appropriate tables and graphics	16-22
Reporting of results should include		
25	Graphic summarizing individual study estimates and overall estimate	N/A
26	Table giving descriptive information for each study included	17
27	Results of sensitivity testing (e.g., subgroup analysis)	N/A
28	Indication of statistical uncertainty of findings	N/A
Reporting of discussion should include		
29	Quantitative assessment of bias (e.g., publication bias)	N/A
30	Justification for exclusion (e.g., exclusion of non-English language citations)	N/A
31	Assessment of quality of included studies	33, 34
Reporting of conclusions should include		
32	Consideration of alternative explanations for observed results	12-14
33	Generalization of the conclusions (i.e., appropriate for the data presented and within the domain of the literature review)	14
34	Guidelines for future research	14
35	Disclosure of funding source	16

From: Stroup DF, Berlin JA, Morton SC, et al, for the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group. Meta-analysis of Observational Studies in Epidemiology. A Proposal for Reporting. *JAMA*. 2000;283(15):2008-2012. doi: 10.1001/jama.283.15.2008

Appendix 2. Search strategy example – MEDLINE.

Database searched = OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

1. *Attitude to Health/
2. *Patient Participation/
3. preference*.ti,ab.
4. *Patient Preference/
5. choice.ti.
6. choices.ti.
7. value*.ti.
8. health state values.ti,ab.
9. valuation*.ti.
10. expectation*.ti,ab.
11. attitude*.ti,ab.
12. acceptab*.ti,ab.
13. knowledge.ti,ab.
14. point of view.ti,ab.
15. user participation.ti,ab.
16. users participation.ti,ab.
17. users' participation.ti,ab.
18. user's participation.ti,ab.
19. patient participation.ti,ab.
20. patients' participation.ti,ab.
21. patients participation.ti,ab.
22. patient's participation.ti,ab.
23. patient perspective*.ti,ab.
24. patients perspective*.ti,ab.
25. patients' perspective*.ti,ab.

26. patient's perspective*.ti,ab.
27. patient perce*.ti,ab.
28. patients perce*.ti,ab.
29. patients' perce*.ti,ab.
30. patient's perce*.ti,ab.
31. health perception*.ti,ab.
32. user view*.ti,ab.
33. users view*.ti,ab.
34. users' view*.ti,ab.
35. user's view*.ti,ab.
36. patient view*.ti,ab.
37. patients view*.ti,ab.
38. patients' view*.ti,ab.
39. patient's view*.ti,ab.
40. or/1-39
41. patient*.ti.
42. user*.ti.
43. men.ti.
44. women.ti.
45. or/41-44
46. exp *Decision Making/
47. decision mak*.ti,ab.
48. decisions mak*.ti,ab.
49. decision*.ti.
50. mak*.ti.
51. 49 and 50
52. avoidance learning/
53. 46 or 47 or 48 or 51 or 52
54. 45 and 53
55. discrete choice.ti,ab.
56. decision board*.ti,ab.

57. decision analy*.ti,ab.
58. decision-support.ti,ab.
59. decision tool*.ti,ab.
60. decision aid*.ti,ab.
61. discrete-choice*.ti,ab.
62. decision*.ti,ab.
63. 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62
64. 45 and 63
65. 54 or 64
66. decision support techniques/
67. (health and utilit*).ti.
68. gamble*.ti,ab.
69. prospect theory.ti,ab.
70. preference score.ti,ab.
71. preference elicitation.ti,ab.
72. health utilit*.ti,ab.
73. (utility and (value* or score* or estimate*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
74. health state.ti,ab.
75. feeling thermometer*.ti,ab.
76. best-worst scaling.ti,ab.
77. best worst scaling.mp.
78. best worst.ti,ab.
79. TTO.ti,ab.
80. time trade-off.ti,ab.
81. probability trade-off.ti,ab.
82. or/66-81
83. Choice Behavior/

84. or/66-83
85. preference based.ti,ab.
86. preference score.ti,ab.
87. multiattribute.ti,ab.
88. multi attribute.mp.
89. EuroQoL 5D.mp.
90. EuroQoL5D.ti,ab.
91. EQ5D.mp.
92. EQ 5D.ti,ab.
93. SF6D.ti,ab.
94. SF 6D.ti,ab.
95. HUI.ti,ab.
96. 15D.ti,ab.
97. or/85-96
98. SF36.ti,ab.
99. SF 36.ti,ab.
100. SF12.ti,ab.
101. SF 12.mp.
102. HRQoL.ti,ab.
103. QoL.ti,ab.
104. quality of life.ti,ab.
105. "Quality of Life"/
106. or/98-105
107. 40 or 65 or 84 or 97 or 106
108. Aortic Stenosis.mp. or exp Aortic Valve Stenosis/
109. (aortic valve implantation or TAVR or transcatheter or transfemoral or transapical or transaxillary or SAVR or heart valve replacement or surgical aortic valve replacement or surgical AVR or SAVR or TAVI or aortic valve replacement or transvascular).af.
110. 107 and 108 and 109
111. limit 110 to humans

Appendix Table 1. Excluded studies, with reasons.

#	Title	First author	Year	Reason for exclusion
1	Toronto Aortic Stenosis Quality of Life Scale (TASQ): Development and quality of life in aortic stenosis and TAVI patients	Styra	2019	Abstract only
2	Rapid-cycle development of decision support tools for patients with symptomatic aortic stenosis	Knoepke	2018	Abstract only
3	Risk willingness and survival in patients with severe aortic stenosis	Hussain	2019	Abstract only
4	A learning curve for shared decision making: The impact of physician experience on decision aid efficacy in severe aortic stenosis	Coylewright	2018	Abstract only
5	Subjective preferences and goal among the patients treated with transaortic valvular replacement	Sugiura	2019	Abstract only
6	Patients and their physicians do not agree on shared decision making in transcatheter aortic valve replacement	Coylewright	2016	Not about values and preferences elicitation
7	Is it worth it? A collaborative clinical decision making exercise using an old-school debate	Wright	2016	Not about values and preferences elicitation
8	The medically managed patient with severe symptomatic aortic stenosis in the TAVR era: Patient characteristics, reasons for medical management, and quality of shared decision making at heart valve treatment centers	Dharmarajan	2017	Not about values and preferences elicitation
9	Patients' Decision Making About Undergoing Transcatheter Aortic Valve Implantation for Severe Aortic Stenosis	Olsson	2016	Not about values and preferences elicitation
10	Determinants and Outcome of Decision Making Among Patients with Severe Aortic Stenosis	Hussain	2017	Not about values and preferences elicitation

11	Patients' self-reported function, symptoms and health-related quality of life before and 6 months after transcatheter aortic valve implantation and surgical aortic valve replacement	Olsson	2017	Not about values and preferences elicitation
12	Self-reported health status, treatment decision and survival in asymptomatic and symptomatic patients with aortic stenosis in a Western Norway population undergoing conservative treatment: a cross-sectional study with 18 months follow-up	Oterhals	2017	Not about values and preferences elicitation
13	[ANMCO/SIC/SICI-GISE/SICCH Consensus document: Risk stratification in elderly patients undergoing cardiac surgery and transcatheter aortic valve implantation]	Pulignano	2016	Not about values and preferences elicitation
14	Patients and informal caregivers' experience of surgical and transcatheter aortic valve replacement: Real-world data contributing to establish value-based medicine in Denmark	Rosseel	2019	Not about values and preferences elicitation
15	Current decision making and short-term outcome in patients with degenerative aortic stenosis: the Pooled-RotterdAm-Milano-Toulouse In Collaboration Aortic Stenosis survey	Van Mieghem	2016	Not about values and preferences elicitation
16	Factors influencing the choice between transcatheter and surgical treatment of severe aortic stenosis in patients younger than 80 years: Results from the OBSERVANT study	Tarantini	2020	Not about values and preferences elicitation
17	Transforming the experience of aortic valve disease in older patients: A qualitative study	Kirk	2019	Not about values and preferences elicitation
18	Long-term outcomes of transcatheter versus surgical aortic valve replacement in low risk, elderly patients with severe aortic stenosis	Kang	2019	Not about values and preferences elicitation

19	Reasons for choosing conservative management in symptomatic patients with severe aortic stenosis - Observations from the CURRENT AS registry	Ishii	2019	Not about values and preferences elicitation
20	Patient disposition and clinical outcome after referral to a dedicated TAVI clinic	Gorecka	2019	Not about values and preferences elicitation
21	Validation of a Heart Team Performance for Patients with Severe Aortic Stenosis	D'Aronco	2019	Not about values and preferences elicitation
22	The Learning Curve for Shared Decision-making in Symptomatic Aortic Stenosis	Coylewright	2020	Not about values and preferences elicitation
23	Pilot Study of a Patient Decision Aid for Valve Choices in Surgical Aortic Valve Replacement	Anaya	2019	Not about values and preferences elicitation
24	Exploring the experience of early discharge after transcatheter aortic valve implantation for older adults and their informal caregivers (Dissertation)	Knoll	2018	Not about values and preferences elicitation
25	Living with Aortic Stenosis: A Phenomenological Study of Patients' Experiences and Subsequent Health Choices (Dissertation)	Hagen-Peter	2015	Not about values and preferences elicitation
26	Low Gradient Aortic Stenosis: Who Benefits From Intervention?	Baumgartner	2019	Not primary study
27	TAVR in Patients With End-Stage Renal Disease and Critical Aortic Stenosis: Hard Choices	Bayliss	2019	Not primary study
28	Quality of life after transcatheter aortic valve replacement	Bonow	2017	Not primary study
29	TAVR: A Good Fix, But It Cannot Fix Everything	Carabello	2016	Not primary study
30	Treating of aortic valve stenosis in real-life: A multifaceted decision-making challenge	Franken	2017	Not primary study
31	Are transcatheter procedures the treatment of choice for all patients with severe aortic stenosis?	Hernandez-Vaquero	2017	Not primary study

32	The less complex the case is, the more complex is it to choose? The case of lower risk patients with aortic valve stenosis	Lemos	2018	Not primary study
33	Elevating Aortic Stenosis Treatment?	McCabe	2018	Not primary study
34	Transcatheter aortic valve implantation in patients with severe aortic stenosis: Does lower-risk profile mean a young patient?	Michel	2019	Not primary study
35	Transcatheter aortic valve replacement: Suitable for all?	Minakata	2018	Not primary study
36	Aortic stenosis: Treat the patient not the numbers	Otto	2018	Not primary study
37	Surgical or transcatheter aortic-valve replacement	Reyes	2017	Not primary study
38	From knowledge to wisdom	Sousa-Uva	2018	Not primary study
39	TAVR - The future of aortic stenosis management	Ullah	2016	Not primary study

Appendix Table 2. Additional study and participant demographics.

Study	Data collection period	Setting	Funding source	Conflicts of interest
Quantitative studies				
Marsh 2020	July-August 2018	Not applicable (online survey)	Edwards Lifesciences	Two authors are employees of Edwards Lifesciences. Three studies are employees of Evidera. Evidera received funding from Edwards Lifesciences to conduct the study and develop the manuscript.
Hussain 2016	May 2010-April 2014	Large university hospital	Norwegian Health Association and Inger and John Fredriksen	No conflict of interest
Qualitative studies				
Coylewright 2015	June 2012-August 2014	Tertiary academic medical institution	No funding sources	No conflict of interest
Olsson 2016	May 2010-June 2011	Large university hospital	Vasterbotten's County Council, Umea˚University, and The Heart Foundation of Northern Sweden	No conflict of interest
Skaar 2017	February 2014-April 2015	Large university hospital	Grieg Foundation, Department of Heart Disease, Haukeland University Hospital and Kavli Research Centre for Geriatrics and Dementia, Haraldsplass Deaconess Hospital, Bergen.	NR
Lauck 2016	NR	Provincial cardiac TAVI center	Providence Health Care Nursing Research Competition	Four authors are consultants to Edward Lifesciences. One author is a consultant for Edward Lifesciences, St. Jude Medical and Abbott Inc., HearthWare, and Norvasc.
Ontario Health Technology Assessment Series 2018	NR	Not applicable (phone interview)	Health Quality Ontario	NR

Frank 2019/Styra 2019	2015-2017	Tertiary academic medical institution	Partially funded from Edwards Lifesciences (manufacturer of TAVI valves)	NR

NR = Not reported.

Appendix Table 3. Qualitative study quality.

Study	Coylewright 2016	Ontario Health Technology Assessment Series 2018	Lauck 2015	Olsson 2019	Skaar 2019	Styra/Frank 2019
1. Was there a clear statement of the aims of the research?	Yes	Yes	Yes	Yes	Yes	Yes
2. Is a qualitative methodology appropriate?	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the research design appropriate to address the aims of the research?	Yes	Yes	Yes	Yes	Yes	Yes
4. Was the recruitment strategy appropriate to the aims of the research?	No	Can't tell	No	Yes	Yes	Yes
5. Was the data collected in a way that addressed the research issue?	Yes	Yes	Yes	Yes	Yes	Can't tell
6. Has the relationship between researcher and participants been adequately considered?	Can't tell	No	No	No	Yes	No
7. Have ethical issues been taken into consideration?	Yes	Can't tell	Yes	Yes	Yes	Yes
8. Was the data analysis sufficiently rigorous?	No	Can't tell	Yes	Yes	Yes	Can't tell
9. Is there a clear statement of findings?	Yes	Yes	Yes	Yes	Yes	Yes
Overall methodological limitations	Moderate	Serious	Moderate	No or very minor	No or very minor	No or very minor

Appendix Table 4. Quantitative study quality.

Risk of bias criteria		**Hussain 2016**	**Marsh 2020**
Selection of participants into the study	Was an appropriate study sample selected from the sampling frame?	Moderate risk of bias	Serious risk of bias
Completeness of data	Was the attrition sufficiently low to minimize the risk of bias?	Moderate risk of bias	Serious risk of bias
Measurement Instrument: Choice of the methodology	Was the instrument used for eliciting relative importance of outcomes valid and reliable?	Moderate risk of bias	Low risk of bias
Measurement Instrument: Administration of the methodology	Was the instrument administered in the intended way?	Low risk of bias	Moderate risk of bias
Measurement Instrument: Outcome presentation	Was a valid representation of the outcome (health state) utilized?	Moderate risk of bias	Serious risk of bias
Measurement Instrument: Understanding of the methodology	Did the researchers check the understanding of the instrument?	Moderate risk of bias	Low risk of bias
Data analysis	Were the results analyzed appropriately to avoid influence of bias and confounding?	Moderate risk of bias	Serious risk of bias

Appendix Table 5. Qualitative results - CERQual Summary of Findings

Summary of Qualitative Review Findings	Reference	Explanation of CERQual assessment
Values and preferences concerning perioperative mortality risk of procedure		
Patients with severe aortic stenosis viewed declining treatment to be worse than accepting the risk related to the procedure	[23]	Limited, thin data to support this finding, only one study that did address both TAVI and SAVR
Risk willingness varied considerably, but many patients were generally willing to accept a high perioperative mortality risk when undergoing aortic valve replacement	[21 23]	Limited, thin data to support this finding, not enough studies, not enough settings, and studies did not address both TAVI and SAVR.
Values and preferences concerning health-related quality of life when deciding on treatment		
Function/ activities of daily living		
Patients aimed for improved body function, better health and activities of daily living when deciding on treatment.	[21 27 23 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate
Patient-defined goals central to decision-making included specific activities and hobbies.	[21 27 23 24]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR.
Patients emphasized importance of managing by oneself or be independent as reasons to undergo treatment.	[21 27 24 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR.
Improve quality of life		
Patients hoped the procedure would improve their quality of life, and spoke of their desire to get back to 'normal', have a 'good life' or have a 'new lease on life' when deciding on treatment.	[27 22 23 24]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies and studies did not address both TAVI and SAVR
Maintaining independence/ obligations		
Patients reported their sense of responsibility to maintain the best possible health condition to be able to fulfill their obligations, including on financial management, maintaining one's home and participating in day-to-day activities.	[21 27 24 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate

Some patients reported that they felt an obligation to their relatives to accept a treatment that was recommended.	[27] [23]	Studies with methodological limitations, limited, thin data to support this finding, uncommon, but important finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate
Values and preferences concerning pain and risk of stroke		
Some patients were concerned about pain or getting a stroke after the procedure.	[22]	Study with methodological limitations, uncommon, but important finding, only one study and TAVI and SAVR was reported separately
Values and preferences related to decision-making guidance on treatment and practical issues		
Patients stressed the importance of a trusting relationship with their physician(s) as essential sources of information, decision-making guidance and facilitators of referral and decision-making	[21] [27] [23]	Studies with methodological limitations, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR, and when it was reported it was separate
There was a high degree of variability on the reliance on informal social support provided by family, friends and community members on their decision making.	[21] [27] [23]	Studies with methodological limitations, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR, and when it was reported it was separate
Patients and caregivers noted that the costs involved with travel and a longer hospital stay were an additional burden and a potential barrier to receiving SAVR.	[22] [21] [23] [27]	All but one study did not address both TAVI and SAVR, and when it was reported it was separate

Appendix Table of Content

Appendix 1. MOOSE Checklist for Meta-analyses of Observational Studies

Item No	Recommendation	Reported on Page No
Reporting of background should include		
1	Problem definition	3
2	Hypothesis statement	N/A
3	Description of study outcome(s)	4
4	Type of exposure or intervention used	4
5	Type of study designs used	4
6	Study population	4
Reporting of search strategy should include		
7	Qualifications of searchers (e.g., librarians and investigators)	3,4
8	Search strategy, including time period included in the synthesis and key words	3, 25-27
9	Effort to include all available studies, including contact with authors	3,4
10	Databases and registries searched	4
11	Search software used, name and version, including special features used (e.g., explosion)	N/A
12	Use of hand searching (e.g., reference lists of obtained articles)	4
13	List of citations located and those excluded, including justification	16, 28-30
14	Method of addressing articles published in languages other than English	N/A
15	Method of handling abstracts and unpublished studies	4
16	Description of any contact with authors	4
Reporting of methods should include		
17	Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	5
18	Rationale for the selection and coding of data (e.g., sound clinical principles or convenience)	5
19	Documentation of how data were classified and coded (e.g., multiple raters, blinding and interrater reliability)	5

20	Assessment of confounding (e.g., comparability of cases and controls in studies where appropriate)	N/A
21	Assessment of study quality, including blinding of quality assessors, stratification or regression on possible predictors of study results	5
22	Assessment of heterogeneity	N/A
23	Description of statistical methods (e.g., complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	N/A
24	Provision of appropriate tables and graphics	16-22
Reporting of results should include		
25	Graphic summarizing individual study estimates and overall estimate	N/A
26	Table giving descriptive information for each study included	17
27	Results of sensitivity testing (e.g., subgroup analysis)	N/A
28	Indication of statistical uncertainty of findings	N/A
Reporting of discussion should include		
29	Quantitative assessment of bias (e.g., publication bias)	N/A
30	Justification for exclusion (e.g., exclusion of non-English language citations)	N/A
31	Assessment of quality of included studies	33, 34
Reporting of conclusions should include		
32	Consideration of alternative explanations for observed results	12-14
33	Generalization of the conclusions (i.e., appropriate for the data presented and within the domain of the literature review)	14
34	Guidelines for future research	14
35	Disclosure of funding source	16

From: Stroup DF, Berlin JA, Morton SC, et al, for the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group. Meta-analysis of Observational Studies in Epidemiology. A Proposal for Reporting. *JAMA*. 2000;283(15):2008-2012. doi: 10.1001/jama.283.15.2008

Appendix 2. Search strategy example – MEDLINE.

Database searched = OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

1. *Attitude to Health/
2. *Patient Participation/
3. preference*.ti,ab.
4. *Patient Preference/
5. choice.ti.
6. choices.ti.
7. value*.ti.
8. health state values.ti,ab.
9. valuation*.ti.
10. expectation*.ti,ab.
11. attitude*.ti,ab.
12. acceptab*.ti,ab.
13. knowledge.ti,ab.
14. point of view.ti,ab.
15. user participation.ti,ab.
16. users participation.ti,ab.
17. users' participation.ti,ab.
18. user's participation.ti,ab.
19. patient participation.ti,ab.
20. patients' participation.ti,ab.
21. patients participation.ti,ab.
22. patient's participation.ti,ab.
23. patient perspective*.ti,ab.
24. patients perspective*.ti,ab.
25. patients' perspective*.ti,ab.

26. patient's perspective*.ti,ab.
27. patient perce*.ti,ab.
28. patients perce*.ti,ab.
29. patients' perce*.ti,ab.
30. patient's perce*.ti,ab.
31. health perception*.ti,ab.
32. user view*.ti,ab.
33. users view*.ti,ab.
34. users' view*.ti,ab.
35. user's view*.ti,ab.
36. patient view*.ti,ab.
37. patients view*.ti,ab.
38. patients' view*.ti,ab.
39. patient's view*.ti,ab.
40. or/1-39
41. patient*.ti.
42. user*.ti.
43. men.ti.
44. women.ti.
45. or/41-44
46. exp *Decision Making/
47. decision mak*.ti,ab.
48. decisions mak*.ti,ab.
49. decision*.ti.
50. mak*.ti.
51. 49 and 50
52. avoidance learning/
53. 46 or 47 or 48 or 51 or 52
54. 45 and 53
55. discrete choice.ti,ab.
56. decision board*.ti,ab.

57. decision analy*.ti,ab.
58. decision-support.ti,ab.
59. decision tool*.ti,ab.
60. decision aid*.ti,ab.
61. discrete-choice*.ti,ab.
62. decision*.ti,ab.
63. 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62
64. 45 and 63
65. 54 or 64
66. decision support techniques/
67. (health and utilit*).ti.
68. gamble*.ti,ab.
69. prospect theory.ti,ab.
70. preference score.ti,ab.
71. preference elicitation.ti,ab.
72. health utilit*.ti,ab.
73. (utility and (value* or score* or estimate*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
74. health state.ti,ab.
75. feeling thermometer*.ti,ab.
76. best-worst scaling.ti,ab.
77. best worst scaling.mp.
78. best worst.ti,ab.
79. TTO.ti,ab.
80. time trade-off.ti,ab.
81. probability trade-off.ti,ab.
82. or/66-81
83. Choice Behavior/

84. or/66-83
85. preference based.ti,ab.
86. preference score.ti,ab.
87. multiattribute.ti,ab.
88. multi attribute.mp.
89. EuroQoL 5D.mp.
90. EuroQoL5D.ti,ab.
91. EQ5D.mp.
92. EQ 5D.ti,ab.
93. SF6D.ti,ab.
94. SF 6D.ti,ab.
95. HUI.ti,ab.
96. 15D.ti,ab.
97. or/85-96
98. SF36.ti,ab.
99. SF 36.ti,ab.
100. SF12.ti,ab.
101. SF 12.mp.
102. HRQoL.ti,ab.
103. QoL.ti,ab.
104. quality of life.ti,ab.
105. "Quality of Life"/
106. or/98-105
107. 40 or 65 or 84 or 97 or 106
108. Aortic Stenosis.mp. or exp Aortic Valve Stenosis/
109. (aortic valve implantation or TAVR or transcatheter or transfemoral or transapical or transaxillary or SAVR or heart valve replacement or surgical aortic valve replacement or surgical AVR or SAVR or TAVI or aortic valve replacement or transvascular).af.
110. 107 and 108 and 109
111. limit 110 to humans

Appendix Table 1. Excluded studies, with reasons.

#	Title	First author	Year	Reason for exclusion
1	Toronto Aortic Stenosis Quality of Life Scale (TASQ): Development and quality of life in aortic stenosis and TAVI patients	Styra	2019	Abstract only
2	Rapid-cycle development of decision support tools for patients with symptomatic aortic stenosis	Knoepke	2018	Abstract only
3	Risk willingness and survival in patients with severe aortic stenosis	Hussain	2019	Abstract only
4	A learning curve for shared decision making: The impact of physician experience on decision aid efficacy in severe aortic stenosis	Coylewright	2018	Abstract only
5	Subjective preferences and goal among the patients treated with transaortic valvular replacement	Sugiura	2019	Abstract only
6	Patients and their physicians do not agree on shared decision making in transcatheter aortic valve replacement	Coylewright	2016	Not about values and preferences elicitation
7	Is it worth it? A collaborative clinical decision making exercise using an old-school debate	Wright	2016	Not about values and preferences elicitation
8	The medically managed patient with severe symptomatic aortic stenosis in the TAVR era: Patient characteristics, reasons for medical management, and quality of shared decision making at heart valve treatment centers	Dharmarajan	2017	Not about values and preferences elicitation
9	Patients' Decision Making About Undergoing Transcatheter Aortic Valve Implantation for Severe Aortic Stenosis	Olsson	2016	Not about values and preferences elicitation
10	Determinants and Outcome of Decision Making Among Patients with Severe Aortic Stenosis	Hussain	2017	Not about values and preferences elicitation

11	Patients' self-reported function, symptoms and health-related quality of life before and 6 months after transcatheter aortic valve implantation and surgical aortic valve replacement	Olsson	2017	Not about values and preferences elicitation
12	Self-reported health status, treatment decision and survival in asymptomatic and symptomatic patients with aortic stenosis in a Western Norway population undergoing conservative treatment: a cross-sectional study with 18 months follow-up	Oterhals	2017	Not about values and preferences elicitation
13	[ANMCO/SIC/SICI-GISE/SICCH Consensus document: Risk stratification in elderly patients undergoing cardiac surgery and transcatheter aortic valve implantation]	Pulignano	2016	Not about values and preferences elicitation
14	Patients and informal caregivers' experience of surgical and transcatheter aortic valve replacement: Real-world data contributing to establish value-based medicine in Denmark	Rosseel	2019	Not about values and preferences elicitation
15	Current decision making and short-term outcome in patients with degenerative aortic stenosis: the Pooled-RotterdAm-Milano-Toulouse In Collaboration Aortic Stenosis survey	Van Mieghem	2016	Not about values and preferences elicitation
16	Factors influencing the choice between transcatheter and surgical treatment of severe aortic stenosis in patients younger than 80 years: Results from the OBSERVANT study	Tarantini	2020	Not about values and preferences elicitation
17	Transforming the experience of aortic valve disease in older patients: A qualitative study	Kirk	2019	Not about values and preferences elicitation
18	Long-term outcomes of transcatheter versus surgical aortic valve replacement in low risk, elderly patients with severe aortic stenosis	Kang	2019	Not about values and preferences elicitation

19	Reasons for choosing conservative management in symptomatic patients with severe aortic stenosis - Observations from the CURRENT AS registry	Ishii	2019	Not about values and preferences elicitation
20	Patient disposition and clinical outcome after referral to a dedicated TAVI clinic	Gorecka	2019	Not about values and preferences elicitation
21	Validation of a Heart Team Performance for Patients with Severe Aortic Stenosis	D'Aronco	2019	Not about values and preferences elicitation
22	The Learning Curve for Shared Decision-making in Symptomatic Aortic Stenosis	Coylewright	2020	Not about values and preferences elicitation
23	Pilot Study of a Patient Decision Aid for Valve Choices in Surgical Aortic Valve Replacement	Anaya	2019	Not about values and preferences elicitation
24	Exploring the experience of early discharge after transcatheter aortic valve implantation for older adults and their informal caregivers (Dissertation)	Knoll	2018	Not about values and preferences elicitation
25	Living with Aortic Stenosis: A Phenomenological Study of Patients' Experiences and Subsequent Health Choices (Dissertation)	Hagen-Peter	2015	Not about values and preferences elicitation
26	Low Gradient Aortic Stenosis: Who Benefits From Intervention?	Baumgartner	2019	Not primary study
27	TAVR in Patients With End-Stage Renal Disease and Critical Aortic Stenosis: Hard Choices	Bayliss	2019	Not primary study
28	Quality of life after transcatheter aortic valve replacement	Bonow	2017	Not primary study
29	TAVR: A Good Fix, But It Cannot Fix Everything	Carabello	2016	Not primary study
30	Treating of aortic valve stenosis in real-life: A multifaceted decision-making challenge	Franken	2017	Not primary study
31	Are transcatheter procedures the treatment of choice for all patients with severe aortic stenosis?	Hernandez-Vaquero	2017	Not primary study

32	The less complex the case is, the more complex is it to choose? The case of lower risk patients with aortic valve stenosis	Lemos	2018	Not primary study
33	Elevating Aortic Stenosis Treatment?	McCabe	2018	Not primary study
34	Transcatheter aortic valve implantation in patients with severe aortic stenosis: Does lower-risk profile mean a young patient?	Michel	2019	Not primary study
35	Transcatheter aortic valve replacement: Suitable for all?	Minakata	2018	Not primary study
36	Aortic stenosis: Treat the patient not the numbers	Otto	2018	Not primary study
37	Surgical or transcatheter aortic-valve replacement	Reyes	2017	Not primary study
38	From knowledge to wisdom	Sousa-Uva	2018	Not primary study
39	TAVR - The future of aortic stenosis management	Ullah	2016	Not primary study

Appendix Table 2. Additional study and participant demographics.

Study	Data collection period	Setting	Funding source	Conflicts of interest
Quantitative studies				
Marsh 2020	July-August 2018	Not applicable (online survey)	Edwards Lifesciences	Two authors are employees of Edwards Lifesciences. Three studies are employees of Evidera. Evidera received funding from Edwards Lifesciences to conduct the study and develop the manuscript.
Hussain 2016	May 2010-April 2014	Large university hospital	Norwegian Health Association and Inger and John Fredriksen	No conflict of interest
Qualitative studies				
Coylewright 2015	June 2012-August 2014	Tertiary academic medical institution	No funding sources	No conflict of interest
Olsson 2016	May 2010-June 2011	Large university hospital	Vasterbotten's County Council, Umea˚University, and The Heart Foundation of Northern Sweden	No conflict of interest
Skaar 2017	February 2014-April 2015	Large university hospital	Grieg Foundation, Department of Heart Disease, Haukeland University Hospital and Kavli Research Centre for Geriatrics and Dementia, Haraldsplass Deaconess Hospital, Bergen.	NR
Lauck 2016	NR	Provincial cardiac TAVI center	Providence Health Care Nursing Research Competition	Four authors are consultants to Edward Lifesciences. One author is a consultant for Edward Lifesciences, St. Jude Medical and Abbott Inc., HearthWare, and Norvasc.
Ontario Health Technology Assessment Series 2018	NR	Not applicable (phone interview)	Health Quality Ontario	NR

Frank 2019/Styra 2019	2015-2017	Tertiary academic medical institution	Partially funded from Edwards Lifesciences (manufacturer of TAVI valves)	NR

NR = Not reported.

Appendix Table 3. Qualitative study quality.

Study	Coylewright 2016	Ontario Health Technology Assessment Series 2018	Lauck 2015	Olsson 2019	Skaar 2019	Styra/Frank 2019
1. Was there a clear statement of the aims of the research?	Yes	Yes	Yes	Yes	Yes	Yes
2. Is a qualitative methodology appropriate?	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the research design appropriate to address the aims of the research?	Yes	Yes	Yes	Yes	Yes	Yes
4. Was the recruitment strategy appropriate to the aims of the research?	No	Can't tell	No	Yes	Yes	Yes
5. Was the data collected in a way that addressed the research issue?	Yes	Yes	Yes	Yes	Yes	Can't tell
6. Has the relationship between researcher and participants been adequately considered?	Can't tell	No	No	No	Yes	No
7. Have ethical issues been taken into consideration?	Yes	Can't tell	Yes	Yes	Yes	Yes
8. Was the data analysis sufficiently rigorous?	No	Can't tell	Yes	Yes	Yes	Can't tell
9. Is there a clear statement of findings?	Yes	Yes	Yes	Yes	Yes	Yes
Overall methodological limitations	Moderate	Serious	Moderate	No or very minor	No or very minor	No or very minor

Appendix Table 4. Quantitative study quality.

Risk of bias criteria		**Hussain 2016**	**Marsh 2020**
Selection of participants into the study	Was an appropriate study sample selected from the sampling frame?	Moderate risk of bias	Serious risk of bias
Completeness of data	Was the attrition sufficiently low to minimize the risk of bias?	Moderate risk of bias	Serious risk of bias
Measurement Instrument: Choice of the methodology	Was the instrument used for eliciting relative importance of outcomes valid and reliable?	Moderate risk of bias	Low risk of bias
Measurement Instrument: Administration of the methodology	Was the instrument administered in the intended way?	Low risk of bias	Moderate risk of bias
Measurement Instrument: Outcome presentation	Was a valid representation of the outcome (health state) utilized?	Moderate risk of bias	Serious risk of bias
Measurement Instrument: Understanding of the methodology	Did the researchers check the understanding of the instrument?	Moderate risk of bias	Low risk of bias
Data analysis	Were the results analyzed appropriately to avoid influence of bias and confounding?	Moderate risk of bias	Serious risk of bias

Appendix Table 5. Qualitative results - CERQual Summary of Findings

Summary of Qualitative Review Findings	Reference	Explanation of CERQual assessment
Values and preferences concerning perioperative mortality risk of procedure		
Patients with severe aortic stenosis viewed declining treatment to be worse than accepting the risk related to the procedure	[23]	Limited, thin data to support this finding, only one study that did address both TAVI and SAVR
Risk willingness varied considerably, but many patients were generally willing to accept a high perioperative mortality risk when undergoing aortic valve replacement	[21 23]	Limited, thin data to support this finding, not enough studies, not enough settings, and studies did not address both TAVI and SAVR.
Values and preferences concerning health-related quality of life when deciding on treatment		
Function/ activities of daily living		
Patients aimed for improved body function, better health and activities of daily living when deciding on treatment.	[21 27 23 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate
Patient-defined goals central to decision-making included specific activities and hobbies.	[21 27 23 24]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR.
Patients emphasized importance of managing by oneself or be independent as reasons to undergo treatment.	[21 27 24 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR.
Improve quality of life		
Patients hoped the procedure would improve their quality of life, and spoke of their desire to get back to 'normal', have a 'good life' or have a 'new lease on life' when deciding on treatment.	[27 22 23 24]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and studies and studies did not address both TAVI and SAVR
Maintaining independence/ obligations		
Patients reported their sense of responsibility to maintain the best possible health condition to be able to fulfill their obligations, including on financial management, maintaining one's home and participating in day-to-day activities.	[21 27 24 22]	Studies with methodological limitations, limited, thin data to support this finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate

Some patients reported that they felt an obligation to their relatives to accept a treatment that was recommended.	[27 23]	Studies with methodological limitations, limited, thin data to support this finding, uncommon, but important finding, not enough studies, not enough settings and all but one study did not address both TAVI and SAVR, and when it was reported it was separate
Values and preferences concerning pain and risk of stroke		
Some patients were concerned about pain or getting a stroke after the procedure.	[22]	Study with methodological limitations, uncommon, but important finding, only one study and TAVI and SAVR was reported separately
Values and preferences related to decision-making guidance on treatment and practical issues		
Patients stressed the importance of a trusting relationship with their physician(s) as essential sources of information, decision-making guidance and facilitators of referral and decision-making	[21 27 23]	Studies with methodological limitations, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR, and when it was reported it was separate
There was a high degree of variability on the reliance on informal social support provided by family, friends and community members on their decision making.	[21 27 23]	Studies with methodological limitations, thin data to support this finding, not enough studies, not enough settings and studies did not address both TAVI and SAVR, and when it was reported it was separate
Patients and caregivers noted that the costs involved with travel and a longer hospital stay were an additional burden and a potential barrier to receiving SAVR.	[22 21 23 27]	All but one study did not address both TAVI and SAVR, and when it was reported it was separate