

BMJ Open Strategies to evaluate healthcare provider trainings in shared decision-making (SDM): a systematic review of evaluation studies

Evamaria Müller,¹ Alena Strukava,¹ Isabelle Scholl,¹ Martin Härter,¹ Ndeye Thiab Diouf,² France Légaré,² Angela Buchholz¹

To cite: Müller E, Strukava A, Scholl I, *et al*. Strategies to evaluate healthcare provider trainings in shared decision-making (SDM): a systematic review of evaluation studies. *BMJ Open* 2019;**9**:e026488. doi:10.1136/bmjopen-2018-026488

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-026488>).

Received 4 September 2018
Revised 27 March 2019
Accepted 6 June 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

²Department of Family Medicine and Emergency Medicine, Laval University, Quebec, Canada

Correspondence to

Evamaria Müller;
e.mueller@uke.de

ABSTRACT

Design and objectives We performed a systematic review of studies evaluating healthcare provider (HCP) trainings in shared decision-making (SDM) to analyse their evaluation strategies.

Setting and participants HCP trainings in SDM from all healthcare settings.

Methods We searched scientific databases (Medline, PsycInfo, CINAHL), performed reference and citation tracking, contacted experts in the field and scanned the Canadian inventory of SDM training programmes for healthcare professionals. We included articles reporting data of summative evaluations of HCP trainings in SDM. Two reviewers screened records, assessed full-text articles, performed data extraction and assessed study quality with the integrated quality criteria for review of multiple study designs (ICROMS) tool. Analysis of evaluation strategies included data source use, use of unpublished or published measures and coverage of Kirkpatrick's evaluation levels. An evaluation framework based on Kirkpatrick's evaluation levels and the Quadruple Aim framework was used to categorise identified evaluation outcomes.

Results Out of 7234 records, we included 41 articles reporting on 30 studies: cluster-randomised (n=8) and randomised (n=9) controlled trials, controlled (n=1) and non-controlled (n=7) before-after studies, mixed-methods (n=1), qualitative (n=1) and post-test (n=3) studies. Most studies were conducted in the USA (n=9), Germany (n=8) or Canada (n=7) and evaluated physician trainings (n=25). Eleven articles met ICROMS quality criteria. Almost all studies (n=27) employed HCP-reported outcomes for training evaluation and most (n=19) additionally used patient-reported (n=12), observer-rated (n=10), standardised patient-reported (n=2) outcomes or training process and healthcare data (n=10). Most studies employed a mix of unpublished and published measures (n=17) and covered two (n=12) or three (n=10) Kirkpatrick's levels. Identified evaluation outcomes covered all categories of the proposed framework.

Conclusions Strategies to evaluate HCP trainings in SDM varied largely. The proposed evaluation framework maybe useful to structure future evaluation studies, but international agreement on a core set of outcomes is needed to improve evidence.

PROSPERO registration number CRD42016041623.

Strengths and limitations of this study

- A strength of this study is the fact that we sought all types of evaluation strategies for healthcare provider trainings in shared decision-making and included all types of study designs from post-test studies to qualitative and cluster-randomised controlled studies.
- A strength of this study is the fact that we developed an evaluation framework, which may be a first step towards agreement on a core set of evaluation outcomes and measures to improve evidence on healthcare provider trainings in shared decision-making.
- A limitation of this study is the fact that we did not analyse which measures are useful to evaluate healthcare provider trainings in shared decision-making.
- A limitation of this study is the fact that we focused on the analysis of evaluation outcomes, but did not analyse evaluation strategies with regard to appropriate measurement points or the match between training contents and use of evaluation outcomes.
- A limitation of the proposed evaluation framework is that it focuses on evaluation outcomes, but does not take into account aspects like appropriate study designs.

INTRODUCTION

Healthcare policies, clinical guidelines and a growing body of research strongly advocate for the implementation of shared decision-making (SDM) as a central element of patient-centred care.¹ Policy makers are interested in SDM, because it tackles overuse, underuse and misuse of healthcare interventions all at the same time.² In SDM, the patient and at least one clinician share information and values, deliberate the next step and arrive at a jointly made decision.³ Patients who experienced SDM reported less decisional conflict and improved satisfaction,⁴ but evidence regarding health-related outcomes is limited.⁴⁻⁶ To date, the most

conclusive argument for SDM is ethical. Patients have the right to learn about available treatment options and their implications, and to participate in decision-making regarding their health.^{4,7} Despite multiple implementation initiatives⁸ and widespread support, SDM is not yet implemented in routine care.^{7,9}

Interventions to foster the implementation of SDM usually target healthcare providers (HCPs), patients or both.¹⁰ They may include the distribution of written educational material or patient decision aids, patient coaching, audit and feedback for HCPs or HCP trainings in SDM.¹¹ HCP trainings in SDM are group or online courses that address HCP SDM attitudes, knowledge or skills. They include the use of lectures, case studies, role play, group discussion or didactic materials.¹² HCP trainings in SDM are considered key to implement SDM in healthcare, but it is unclear what kind of trainings are most effective and which outcomes they affect.¹⁰⁻¹³ The lack of consensus on an evaluation framework for HCP trainings in SDM partly accounts for this lack of evidence.¹⁴

Evaluation frameworks support practitioners and researchers in the design of coherent evaluation strategies.¹⁵ Kirkpatrick's four-level training evaluation model¹⁶ is the most established and feasible model for training evaluation and can be applied to the context of HCP professional development.¹⁷ Kirkpatrick's four levels are: 1) reaction, 2) learning, 3) behaviour and 4) results. The reaction level includes participant reactions to the training and can be assessed with attendance levels or subjective training impressions. The learning level covers participant changes in attitudes, knowledge or skills after the training. The behaviour level covers changes in participant behaviours or transference of training content to the workplace. The results level describes more tangible trainings results, for example, system effects or patient health outcomes.^{4,17,18}

Elwyn *et al*⁹ argue that SDM research has neglected investigation of diverse long-term consequences on the results level. They postulate that widespread implementation of SDM leads to safer and more cost-effective decisions, to reduce utilisation rates and to improve patient health outcomes, but evidence is lacking.⁹ The influential Quadruple Aim framework aims to improve the experience of care, the health of populations, the per capita cost of healthcare and the work life of HCPs,¹⁹ and may be useful to structure evaluation of HCP trainings in SDM on the results level.

In this review, we aimed to analyse how the diversity of evaluation strategies and the quality of published evaluations contributes to the current lack of evidence on HCP trainings in SDM. Thus, we aimed to investigate the quality of published evaluations of HCP trainings in SDM, and to analyse their evaluation strategies. We aimed to analyse evaluation strategies regarding 1) use of data sources, 2) use of unpublished or self-developed and published or psychometrically tested measures and 3) coverage of Kirkpatrick's four levels. We aimed to categorise identified outcomes in an evaluation framework for

HCP trainings in SDM based on Kirkpatrick's four-level evaluation model¹⁶ and the Quadruple Aim framework¹⁹ to guide future research and to initiate discussion about a core set of evaluation outcomes for this purpose.

METHODS

Registration and search strategy

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for systematic reviews²⁰ in most parts (see online supplementary file S1). We made the following changes to the protocol: we adapted the PICOS (P: patient, problem or population, I: intervention, C: comparison, control or comparator, O: outcomes, S: study type) criteria to meet our research purpose, we did not remove duplicates in the secondary search, we did not assess risk of bias across studies and we did not use any summary measures or additional analyses as we limited our work to qualitative synthesis only. We performed an electronic database search employing Medline, CINAHL and PsycInfo databases (via OVID) on 26 June 2016. For this purpose, we developed a detailed search strategy for each database. We adapted the PICOS criteria²⁰ and considered a combination of the following aspects appropriate: population AND intervention AND construct AND outcome OR study design. Terms and keywords were adapted for each database and searches in Medline and PsycInfo were limited to publications concerning humans. We updated the electronic database search on 30 January 2019. Full insight in the electronic database search strategy is attainable in online supplementary file S2. Moreover, we performed a secondary search including reference and citation tracking of included full-text articles, consultation of experts in the field of research via a shared decision-making facebook group and a screening of the Canadian inventory of SDM training programmes for healthcare professionals (<http://www.decision.chaire.fmed.ulaval.ca/en/list-of-sdm-programs>). Additionally, we screened references of two reviews on SDM interventions for HCPs.^{11,12} We registered details of the protocol for this systematic review on PROSPERO website accessible via www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016041623.

Article selection

We aimed to include articles reporting on summative evaluations (outcome or study design) of HCP (population) trainings (intervention) in SDM (construct) and developed inclusion and exclusion criteria, accordingly. We aimed to exclude articles reporting on formative evaluations or interventions that do not have the main aim to teach SDM skills to HCPs (Table 1). Following the first database search (26 June 2016), two reviewers independently screened titles and abstracts of a random sample of 300 (>5%) records identified in the electronic database to ensure sufficient inter-rater reliability. We discussed any differences until we reached consensus.

Table 1 Descriptive data of included study articles

Study	Country of origin	Study design	Healthcare provider sample*	Patient sample*
Bernhard <i>et al</i> ⁵⁴	AUS, NZ, CHE, GER, AUT	RCT	Medical, surgical, radiation and gynaecological oncologists (n=62)	n=769
Bieber <i>et al</i> ²³	GER	RCT	Specialists in internal medicine (n=13)	n=111
Bieber <i>et al</i> ²²	GER	RCT	Specialists in internal medicine (n=10)	n=85
Bieber <i>et al</i> ²⁴	GER	NCBA	Physicians with direct patient contact (n=123)	n/a
Bieber <i>et al</i> ³⁰	GER	RCT	Physicians treating patients with cancer (n=27)	n=107
Butow <i>et al</i> ⁵⁵	AUS	RCT	Medical, surgical, radiation and gynaecological oncologists (n=62)	n=158
Cohen <i>et al</i> ⁶⁰	UK	CRCT	General practitioners (n=20)	n=n/r
Davis <i>et al</i> ⁵¹	UK	QUAL	General practitioners (n=21)	n=38
Dion <i>et al</i> ⁴¹	CAN	NCBA	Family medicine residents (n=247)	n/a
Edwards <i>et al</i> ⁵²	UK	RCT	General practitioners (n=20)	n=n/r
Edwards <i>et al</i> ⁵²	UK	CRCT	General practitioners (n=21)	n=747
Edwards <i>et al</i> ⁵³	UK	QUAL	General practitioners (n=18)	n/a
Edwards <i>et al</i> ⁵²	UK	CRCT	General practitioners (n=20)	n=352
Feng <i>et al</i> ⁴²	USA	RCT	Primary care physicians (n=118)	n=n/r
Geiger <i>et al</i> ³¹	GER	RCT	Physicians (n=38)	n=152
Härter <i>et al</i> ²⁵	GER	CRCT	Physicians treating patients with cancer (n=33)	n=160
Jo and An ⁵⁹	KOR	CBA	Female intensive care unit nurses (n=41)	n/a
Kasper <i>et al</i> ³²	GER	NCBA	Physicians working in outpatient clinics (n=10)	n=40
Körner <i>et al</i> ²⁶	GER	CRCT	Healthcare provider executives of different occupational backgrounds (n=74)	n/a
Körner <i>et al</i> ²⁷	GER	CRCT	Healthcare provider executives of different occupational backgrounds (n=74)	n=n/r
LeBlanc <i>et al</i> ³³	CAN	CRCT	Family medicine group physicians (n=39)	n=544
Légaré <i>et al</i> ³⁶	CAN	CRCT	Family medicine group physicians (n=33)	n=459
Légaré <i>et al</i> ³⁵	CAN	CRCT	Family physicians (n=306)	n=449
Légaré <i>et al</i> ³⁴	CAN	CRCT	Family physicians, residents and nurse practitioners (n=250)	n=250
Loh <i>et al</i> ²⁸	GER	Post	General practitioners (n=20)	n/a
McCallister <i>et al</i> ⁴³	USA	NCBA	Pulmonary and critical care medical fellows (n=16)	n/a
Metcalfe <i>et al</i> ⁵⁶	AUS	NCBA	General practitioners (n=63)	n/a
Murray <i>et al</i> ³⁷	CAN	RCT	Oncology or palliative care nursing and allied healthcare providers (n=88)	n/a
Price-Haywood <i>et al</i> ⁴⁴	USA	CRCT	Primary care physicians (n=18)	n=161
Sanders <i>et al</i> ⁵⁷	NL	CRCT	General practitioners (n=42)	n=175
Sanders <i>et al</i> ⁵⁸	NL	CRCT	General practitioners (n=47)	n=226
Simmons <i>et al</i> ⁶⁰	USA	NCBA	Internal medicine residents (n=98)	n/a
Stacey <i>et al</i> ³⁸	CAN	RCT	Call centre nurses (n=39)	n/a
Stacey <i>et al</i> ³⁹	CAN	NCBA	Oncology medical residents (n=11)	n/a
Sullivan <i>et al</i> ⁴⁶	USA	RCT	Internal medical residents, attendings (n=45)	n/a
Sullivan <i>et al</i> ⁴⁵	USA	RCT	Internal medical residents (n=213)	n/a
Tinsel <i>et al</i> ²⁹	GER	CRCT	General practitioners (n=36)	n=1120
Towle <i>et al</i> ⁴⁰	CAN	QUAL	Family physicians (n=6)	n=198
Volk <i>et al</i> ⁴⁷	USA	Post	Clinicians from diverse specialties (n=49)	n/a
Wilkes <i>et al</i> ⁴⁸	USA	CRCT	Primary care physicians (n=120)	n=712
Yuen <i>et al</i> ⁴⁹	USA	Post	Internal medicine residents (n=29)	n/a

*Participants who provided data for analysis.

†Articles report data from one study.

AUS, Australia; AUT, Austria; CAN, Canada; CBA, controlled before-after study; CHE, Switzerland; CRCT, cluster-randomised controlled trial; GER, Germany; KOR, Korea; n/a, not applicable; NCBA, non-controlled before-after study; NZ, New Zealand; NL, The Netherlands; n/r, not reported; QUAL, qualitative study; Post, post-test only study; RCT, randomised controlled trial.

Records identified in the electronic database search were then split in half to be assessed for possible inclusion in the study by one of two reviewers. Following the update of the database search (30 January 2019), two reviewers independently screened all identified records and discussed any differences until consensus was reached. Two reviewers independently assessed full-text articles for eligibility by applying inclusion and exclusion criteria (box 1). We resolved differences by discussion until we reached consensus. If consensus could not be reached, the final decision was made by discussion with two other reviewers.

Data extraction, quality assessment and analysis of evaluation strategies

We used data extraction sheets to collect descriptive data of included articles, for example, country of origin of the study, study design, characteristics of HCP and patient samples. Furthermore, we extracted data on evaluation outcomes reported in included articles and all data relevant to assess study quality of included articles. Data extraction sheets were pilot-tested and adjusted accordingly. We assessed study quality of included articles with the integrated quality criteria for review of multiple study designs (ICROMS) tool.²¹ Two reviewers independently performed data extraction and quality evaluation and discussed any differences until consensus was reached. One reviewer performed analysis of evaluation strategies in discussion with the team. As study results are repeatedly published in more than one article, we will present results on two levels: study and article, if applicable.

Quality assessment with the ICROMS tool

The ICROMS tool appraises the quality of multiple study designs and stems from an iterative process over 2 years that included review of existing quality criteria,

pilot testing and expert consensus. It aims to establish criteria critically appraising the quality of multiple study designs, in order to broaden the database for systematic reviews and to inspire rigorous research.²¹ The ICROMS tool comprises 7 dimensions and defines 33 specific criteria for these dimensions applicable only to some study designs. ICROMS dimensions are 1) clear aims and justification, 2) managing bias in sampling between groups, 3) managing bias in outcome measurements and blinding, 4) managing bias in follow-up, 5) managing bias in other study aspects, 6) analytical rigour, 7) managing bias in reporting/ethical considerations. The ICROMS tool is applicable for cluster-randomised and randomised controlled trials, controlled and non-controlled before-after studies, controlled and non-controlled interrupted times series, cohort studies and qualitative studies. As the ICROMS tool is not applicable to post-test studies, we did not assess study quality for articles reporting on this study type. ICROMS-specific criteria are answerable with yes (2 points), no (0 points) or unclear (1 point). The ICROMS tool defines mandatory criteria and minimum scores for different study types to distinguish if studies are fit for inclusion in a systematic review. Minimum scores vary per study type and range from 16 for qualitative studies over 18 for controlled before-after studies to 22 for non-controlled before-after studies or cluster-randomised and randomised controlled trials. Detailed information on the ICROMS tool is attainable in the original publication.²¹ We analysed quality assessment results on article level.

Analysis of evaluation strategies

One reviewer analysed evaluation strategies regarding use of data sources (HCPs, patients, standardised patients, observers, training process and healthcare data), use of

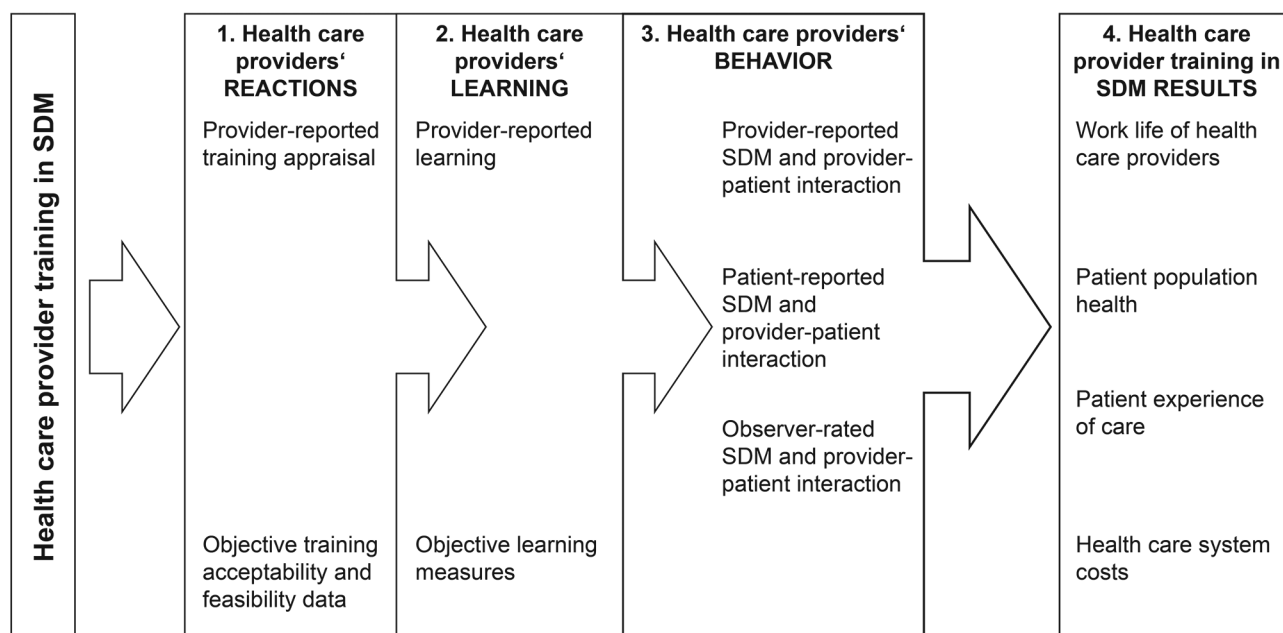


Figure 1 Evaluation framework for healthcare provider trainings in shared decision-making (SDM).

Box 1 Inclusion and exclusion criteria

Inclusion criteria* Inclusion criteria*

1. Article reports data on a health care provider training in SDM
2. Article reports data on summative evaluation of a SDM training
3. Article has the aim to evaluate a SDM training

Exclusion criteria Exclusion criteria

1. Article is a study protocol or clinical trial report
2. Article reports on a study in which participants were medical or health care students
3. Article only reports on formative evaluation of a SDM training
4. Article reports on an intervention that is limited to or has the main aim to teach the use of a decision-making tool
5. Article reports on an intervention that does not have the main aim to teach SDM communication skills
6. Article reports on a complex intervention, in which health care provider training is only a component and no separate summative evaluation data on the training are reported
7. Article is written in a language other than English, German, French or Dutch (languages spoken by team members)
8. Full text of the article is not available

*Articles were excluded if inclusion criteria were not met.

unpublished or self-developed and published or psychometrically tested measures and coverage of Kirkpatrick's four levels of reaction, learning, behaviour and results.¹⁶ One reviewer categorised identified evaluation outcomes in the proposed evaluation framework for HCP trainings in SDM (Figure 1) that is based on the Kirkpatrick's four-level evaluation model¹⁶ and the Quadruple Aim framework.¹⁹ One reviewer developed comprehensive subcategories of evaluation outcomes based on the measures identified in the review and categorised evaluation outcomes accordingly. The study team supervised this process and provided feedback in team discussions. As study results are repeatedly published in more than one article, we will present results on two levels: study and article, if applicable.

Patient and public involvement

We did not involve patients in the conduction of this study.

RESULTS

Literature search and article selection

The electronic database search on 26 June 2016 identified 5317 records. After removal of duplicates, 4543 records remained. We found an additional number of 1636 records through the secondary search. The electronic database search on 30 January 2019 identified additional 1222 records. After removal of duplicates, 1055 records remained. We finally screened 7234 records, of which some are likely to be unidentified duplicates due to our complex search strategy. We excluded 7137 records based on title and abstract screening and assessed 97 full-text articles for eligibility. Of the remaining 97 full-text articles, we excluded 56 full-text articles by applying

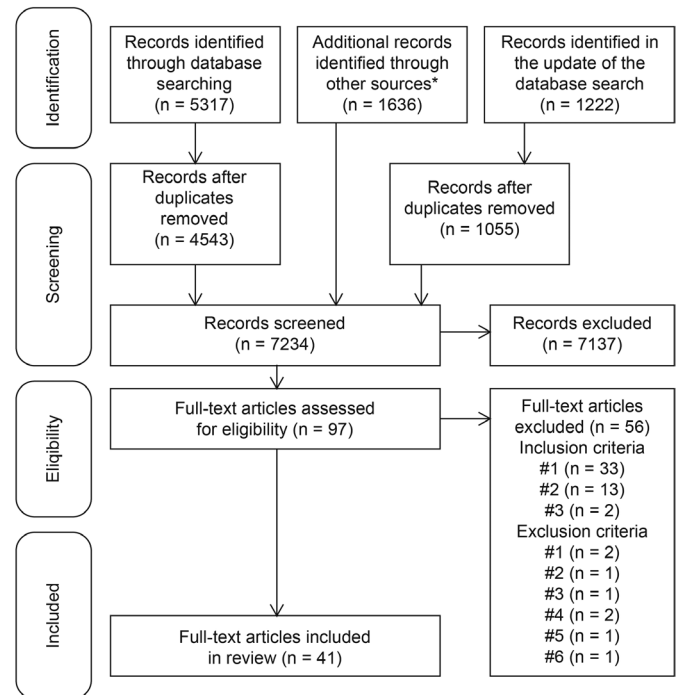


Figure 2 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart. *Due to multiple search strategies, duplicates were not removed.

inclusion and exclusion criteria (Table 1). We excluded the majority of full-text articles because they did not meet the first inclusion criterion and did not report data on an SDM training for HCPs. We included 41 articles in this review. Figure 2 shows the process of article selection.

Descriptive data of included studies and articles

Identified articles (n=41) report on studies (n=30) conducted in a limited number of countries (n=10). Most studies were conducted in the USA (n=9), Germany (n=8) and Canada (n=7). Eleven articles report on studies from Germany,^{22–32} nine articles report on studies from Canada^{33–41} and eight articles report on studies from the USA.^{42–49} Six articles depict one study from the UK^{50–53} and four articles present studies conducted either multinationally,⁵⁴ in Australia,^{55 56} the Netherlands^{57 58} or Korea.⁵⁹ The majority of included articles (n=27) report on cluster-randomised^{25–27 29–31 33–36 44 48 50 52 57 58} and randomised^{22 23 30 31 37 38 42 45 46 50 54 55} controlled trials. Further articles report on one controlled⁵⁹ and seven non-controlled^{24 32 39 41 43 56 60} before-after studies, three qualitative^{40 51 53} and three post-test^{28 47 49} studies. Most articles (n=34) report on the evaluation of physician trainings,^{22–25 28–33 35 36 39–46 48–58 60} two articles report on trainings for nurses^{38 59} and five articles report on trainings for diverse HCPs.^{26 27 34 37 47} Overall, identified articles report on HCP samples ranging from 6 to 306, and n=25 articles^{22 23 25 27 29–36 40 42 44 48 50–52 54 55 57 58} report on the use of patient samples ranging from 38 to 1120. Table 1 illustrates descriptive data of included studies and articles.

Quality results of the ICROMS tool

Assessment of the quality of included articles with the ICROMS tool was applicable to 38 of the included articles (Table 2). Three articles were post-test studies,^{28 47 49} which could not be assessed with the ICROMS tool. Of the 22 articles that met the minimum score, 7 reported on randomised controlled trials,^{22 30 31 37 38 42 55} 12 on cluster-randomised controlled trials^{25 29 33 35 36 44 48 50 52 57 58} and 3 reported on qualitative studies.^{40 51 53} Looking in detail at the 16 articles^{23 24 26 27 32 34 39 41 43 45 46 50 54 56 59 60} that did not meet the minimum score, most of them failed to meet criterion 3E (blinded assessment of primary outcome), 3F (reliable primary outcome measures) and 7D (free of other bias). For detailed results regarding ICROMS criteria, see online supplementary file S3.

Most of the included studies (n=30) and articles (n=41) report use of more than one type of data source to evaluate training effects. Of the studies employing HCP-reported data (n=27), eight studies^{24 28 45-47 49 56 59} relied solely on HCPs for training evaluation. The remaining 19 studies additionally employed other types of outcomes, for example, patient-reported,^{22 23 27 31-36 40 44 48 51 54 55} observer-rated,^{31 32 37-40 42 43 52 60} standardised patient-reported outcomes^{44 48} or training process and healthcare data.^{32 33 36-39 41 44 60} The three studies not relying on HCPs as data source^{25 29 30 57 58} combined patient-reported data with observer-rated measures^{25 57} or training process and healthcare data.^{29 57} Table 3 presents an overview of the data sources used for training evaluation in identified studies and articles.

All but one study⁴⁰ employed quantitative evaluation strategies. Of the 29 remaining studies, articles of 17 studies^{22 23 26 27 29 32-39 44-46 48 52 55 57-59} reported use of a combination of unpublished or self-developed and published or psychometrically tested measures. Nine studies^{24 28 41-43 47 49 56 60} employed only unpublished or self-developed measures and three studies^{25 30 31 54} applied only published or psychometrically tested outcomes.

Looking at studies' evaluation strategies with regard to coverage of Kirkpatrick's four levels, most studies covered two (n=12) or three (n=10) levels. One study⁵⁴ measured only on the results level and seven studies covered all four evaluation levels (table 4). All but one³² of the studies measuring on the reaction level also used outcomes on the learning level.^{24 26 28 33 37-39 41 45-48 53 55 56 60} Articles of five studies^{33 36 49 54 55 59} indicate measurement on the results level without covering the evaluation level of behaviour.

Evaluation outcomes identified in the articles of this review were categorised in comprehensive subcategories of the proposed evaluation framework (table 5). The first level of HCPs' reactions to the training includes provider-reported appraisal of the training and objective acceptability and feasibility data. The second level of HCPs' learning includes provider-reported learning like subjective knowledge gain, attitudes and intentions to engage in SDM, confidence in SDM communication skills and medical competencies. Objective learning measures are knowledge tests like multiple-choice or open questions on

contents covered in the training. The third level of HCPs' behaviour includes provider-reported or (standardised) patient-reported and observer-rated SDM performance and assessment of the patient-provider interaction. The fourth-level HCP training in SDM results reflects the Quadruple Aim framework including the work life of HCPs, patient population health, patient experience of care and healthcare system costs. Table 5 presents the subcategories of evaluation outcomes and how frequently they were used in included articles. Detailed information on the outcomes used in respective articles can be found in the online supplementary file S4.

DISCUSSION

Our review aimed to investigate how the diversity of evaluation strategies and the quality of published evaluations contributes to the current lack of evidence on HCP trainings in SDM. Thus, we analysed the quality of published articles on HCP trainings in SDM, and analysed their evaluation strategies regarding 1) use of data sources, 2) use of unpublished or self-developed and published or psychometrically tested measures and 3) coverage of Kirkpatrick's four levels. We found 41 articles reporting on 30 studies that met our inclusion criteria. Most of these studies were cluster-randomised and randomised controlled trials that evaluated SDM trainings for physicians and were conducted in high-income countries like Canada, the USA, the UK or Germany. Sample sizes varied largely. Of the 38 articles eligible for assessment with the ICROMS tool, only 11 articles met ICROMS quality criteria. Diverse strategies were used to evaluate HCP trainings in SDM, but most studies relied on provider-reported outcomes, covered two or three of Kirkpatrick's levels and combined published and unpublished measures. The proposed evaluation framework based on Kirkpatrick's four-level evaluation model¹⁶ and the Quadruple Aim framework¹⁹ appears useful for the design or analysis of strategies to evaluate HCP trainings in SDM.

The poor quality of identified publications indicates that researchers should aim to design more methodologically sound studies to evaluate HCP trainings in SDM. The ICROMS tool is a decision matrix to evaluate the robustness of studies for inclusion in a review²¹ and present results could inspire researchers to be more rigorous in their study. Since measurement bias was a common problem of many included studies, it would be good to use more objective training acceptability and feasibility data, more objective learning and observer-rated measures and healthcare data for evaluation. However, assessment of specific learning objectives may require application of self-developed measures. Combined with psychometrically sound primary outcomes, this may be the ideal evaluation approach.

Although HCPs were the main data source in included studies, reaction to the training was the least studied evaluation level. Training participants' favourable reactions

Table 2 Quality results of the ICROMS tool

Study	Study design	ICROMS score	Minimum score* met	Mandatory criteria met	Recommendation for inclusion
Bernhard <i>et al</i> ⁵⁴	RCT	20	No	No	No
Bieber <i>et al</i> ²³ †	RCT	20	No	No	No
Bieber <i>et al</i> ²² †	RCT	24	Yes	Yes	Yes
Bieber <i>et al</i> ²⁴	NCBA	19	No	No	No
Bieber <i>et al</i> ³⁰ †	RCT	23	Yes	Yes	Yes
Butow <i>et al</i> ⁵⁵	RCT	22	Yes	No	No
Cohen <i>et al</i> ⁵⁰ †	CRCT	26	Yes	No	No
Davis <i>et al</i> ⁵¹ †	QUAL	21	Yes	Yes	Yes
Dion <i>et al</i> ⁴¹	NCBA	18	No	No	No
Edwards <i>et al</i> ⁵² †	RCT	18	No	No	No
Edwards <i>et al</i> ⁵² †	CRCT	24	Yes	No	No
Edwards <i>et al</i> ⁵³ †	QUAL	20	Yes	No	No
Edwards <i>et al</i> ⁵² †	CRCT	26	Yes	Yes	Yes
Feng <i>et al</i> ⁴²	RCT	22	Yes	No	No
Geiger <i>et al</i> ³¹	RCT	23	Yes	Yes	Yes
Härter <i>et al</i> ²⁵ †	CRCT	23	Yes	Yes	Yes
Jo and An ⁵⁹	CBA	16	No	No	No
Kasper <i>et al</i> ³²	NCBA	19	No	No	No
Körner <i>et al</i> ²⁶ †	CRCT	20	No	No	No
Körner <i>et al</i> ²⁷ †	CRCT	18	No	No	No
LeBlanc <i>et al</i> ³³ †	CRCT	26	Yes	No	No
Légaré <i>et al</i> ³⁶ †	CRCT	24	Yes	No	No
Légaré <i>et al</i> ³⁵ †	CRCT	23	Yes	No	No
Légaré <i>et al</i> ³⁴ †	CRCT	19	No	No	No
Loh <i>et al</i> ²⁸	Post-test	n/a	n/a	n/a	n/a
McCallister <i>et al</i> ⁴³	NCBA	19	No	No	No
Metcalfe <i>et al</i> ⁵⁶	NCBA	8	No	No	No
Murray <i>et al</i> ³⁷	RCT	22	Yes	Yes	Yes
Price-Haywood <i>et al</i> ⁴⁴	CRCT	22	Yes	No	No
Sanders <i>et al</i> ⁵⁷ †	CRCT	22	Yes	No	No
Sanders <i>et al</i> ⁵⁸	CRCT	23	Yes	Yes	Yes
Simmons <i>et al</i> ⁶⁰	NCBA	11	No	No	No
Stacey <i>et al</i> ³⁸	RCT	23	Yes	Yes	Yes
Stacey <i>et al</i> ³⁹	NCBA	18	No	No	No
Sullivan <i>et al</i> ⁴⁶	RCT	17	No	No	No
Sullivan <i>et al</i> ⁴⁵	RCT	19	No	No	No
Tinsel <i>et al</i> ²⁹	CRCT	25	Yes	Yes	Yes
Towle <i>et al</i> ⁴⁰	QUAL	18	Yes	No	No
Volk <i>et al</i> ⁴⁷	Post-test	n/a	n/a	n/a	n/a
Wilkes <i>et al</i> ⁴⁸	CRCT	22	Yes	Yes	Yes
Yuen <i>et al</i> ⁴⁹	Post-test	n/a	n/a	n/a	n/a
No. of articles		38	22	11	11

*ICROMS minimum score for study type: CRCT and RCT: 22, CBA: 18, NCBA: 22, QUAL: 16, for further details see original publication of the ICROMS tool.²¹

†Articles report data from one study.

CBA, controlled before-after study; CRCT, cluster-randomised controlled trial; ICROMS, integrated quality criteria for review of multiple study designs; NCBA, non-controlled before-after study; Post-test, post-test only study; n/a, not applicable; QUAL, qualitative study; RCT, randomised controlled trial.

Table 3 Analysis of data source use

Data source	Healthcare providers	Patients	Observers	Standardised patients	Training process and healthcare data
Bernhard <i>et al</i> ⁵⁴	▲	▲			
Bieber <i>et al</i> ^{23*}	▲	▲			
Bieber <i>et al</i> ^{22*}	▲	▲			
Bieber <i>et al</i> ²⁴	▲				
Bieber <i>et al</i> ^{30*}		▲			
Butow <i>et al</i> ⁵⁵	▲	▲			
Cohen <i>et al</i> ^{50*}					▲
Davis <i>et al</i> ^{51*}	▲	▲			
Dion <i>et al</i> ⁴¹	▲				▲
Edwards <i>et al</i> ^{52*}	▲				
Edwards <i>et al</i> ^{52*}		▲			
Edwards <i>et al</i> ^{53*}	▲				
Edwards <i>et al</i> ^{52*}	▲		▲		
Feng <i>et al</i> ⁴²	▲		▲		
Geiger <i>et al</i> ³¹	▲	▲	▲		
Härter <i>et al</i> ^{25*}		▲	▲		
Jo and An ⁵⁹	▲				
Kasper <i>et al</i> ³²	▲	▲	▲		▲
Körner <i>et al.</i> ^{26*}	▲				
Koerner <i>et al</i> ^{27*}	▲	▲			
LeBlanc <i>et al</i> ^{33*}	▲	▲			▲
Légaré <i>et al</i> ^{36*}	▲	▲			▲
Légaré <i>et al</i> ^{35*}	▲	▲			
Légaré <i>et al</i> ^{34*}	▲	▲			
Loh <i>et al</i> ²⁸	▲				
McCallister <i>et al</i> ⁴³	▲		▲		
Metcalfe <i>et al</i> ⁵⁶	▲				
Murray <i>et al</i> ³⁷	▲		▲		▲
Price-Haywood <i>et al</i> ⁴⁴	▲	▲		▲	▲
Sanders <i>et al</i> ^{57*}			▲		▲
Sanders <i>et al</i> ^{58*}		▲			
Simmons <i>et al</i> ⁶⁰	▲		▲		▲
Stacey <i>et al</i> ³⁸	▲		▲		▲
Stacey <i>et al</i> ³⁹	▲		▲		▲
Sullivan <i>et al</i> ⁴⁶	▲				
Sullivan <i>et al</i> ⁴⁵	▲				
Tinsel <i>et al</i> ²⁹		▲			▲
Towle <i>et al</i> ⁴⁰	▲	▲	▲		
Volk <i>et al</i> ⁴⁷	▲				
Wilkes <i>et al</i> ⁴⁸	▲	▲		▲	
Yuen <i>et al</i> ⁴⁹	▲				
No. of articles (n=41)	34	20	12	2	12
No. of studies (n=30)	27	15	12	2	11

*Articles report data from one study.

▲, the article reports the use of this type of data source for training evaluation.

Table 4 Coverage of Kirkpatrick's evaluation levels

Study	Reaction	Learning	Behaviour	Results
Bernhard <i>et al</i> ⁵⁴				▲
Bieber <i>et al</i> ^{23*}			▲	▲
Bieber <i>et al</i> ^{22*}				▲
Bieber <i>et al</i> ²⁴	▲	▲		
Bieber <i>et al</i> ^{30*}				▲
Butow <i>et al</i> ⁵⁵	▲	▲		▲
Cohen <i>et al</i> ^{50*}				▲
Davis <i>et al</i> ^{51*}			▲	▲
Dion <i>et al</i> ⁴¹	▲	▲		
Edwards <i>et al</i> ^{52*}		▲	▲	
Edwards <i>et al</i> ^{52*}			▲	▲
Edwards <i>et al</i> ⁵³	▲	▲	▲	▲
Edwards <i>et al</i> ^{52*}			▲	▲
Feng <i>et al</i> ⁴²		▲	▲	▲
Geiger <i>et al</i> ³¹			▲	▲
Härter <i>et al</i> ²⁵			▲	▲
Jo and An ⁵⁹		▲		▲
Kasper <i>et al</i> ³²	▲		▲	▲
Körner <i>et al</i> ^{26*}	▲	▲	▲	
Körner <i>et al</i> ^{27*}			▲	
LeBlanc <i>et al</i> ^{33*}	▲	▲		▲
Légaré <i>et al</i> ^{36*}		▲		▲
Légaré <i>et al</i> ^{35*}		▲		▲
Légaré <i>et al</i> ^{34*}		▲	▲	▲
Loh <i>et al</i> ²⁸	▲	▲		
McCallister <i>et al</i> ⁴³		▲	▲	
Metcalfe <i>et al</i> ⁵⁶	▲	▲	▲	
Murray <i>et al</i> ³⁷	▲	▲	▲	▲
Price-Haywood <i>et al</i> ⁴⁴ (2014)		▲	▲	▲
Sanders <i>et al</i> ^{57*}			▲	▲
Sanders <i>et al</i> ^{58*}			▲	▲
Simmons <i>et al</i> ⁶⁰	▲	▲	▲	▲
Stacey <i>et al</i> ³⁸	▲	▲	▲	▲
Stacey <i>et al</i> ³⁹	▲	▲	▲	
Sullivan <i>et al</i> ⁴⁶	▲	▲	▲	▲
Sullivan <i>et al</i> ⁴⁵	▲	▲	▲	▲
Tinsel <i>et al</i> ²⁹			▲	▲
Towle <i>et al</i> ⁴⁰		▲	▲	▲
Volk <i>et al</i> ⁴⁷	▲	▲		
Wilkes <i>et al</i> ⁴⁸	▲	▲	▲	▲
Yuen <i>et al</i> ⁴⁹		▲		▲
No. of articles (n=41)	17	26	27	31

Continued

Table 4 Continued

Study	Reaction	Learning	Behaviour	Results
No. of studies (n=30)	17	23	21	22

*Articles report data from one study.

▲, the article reports the use of evaluation outcomes on this level for training evaluation.

are substantial for the training to be effective as participants' positive appraisal determines their motivations to learn from the training.¹⁶ Following the reaction level, researchers should assess HCP learning using objective learning measures for knowledge gain. Provider-reported learning measures are useful to establish training effects on HCP attitudes, intentions and confidence regarding SDM-related behaviour, which are the predecessors of actual behaviour.¹⁷ According to the theory of planned behaviour, a positive attitude, acquirement of relevant knowledge and improvement of skills determine HCP behavioural intentions, and thus behaviour change.¹⁷

Measurement of behaviour change is central, as change of SDM-related behaviours is usually the main aim of HCP trainings in SDM. Since there is no gold standard for measuring SDM⁶¹ and measurement from different viewpoints is inconsistent,⁶² multiperspective assessment from the viewpoints of HCPs (standardised), patients and observers appears the best approach.⁵ Ideally, validated measures should be used to ensure quality and comparability of results, but a lack of psychometrically tested SDM measures⁶¹ poses a problem. It is also difficult to assess behaviour change in clinical practice, because it is unclear when changes manifest themselves.¹⁶ However, it is critical to establish behaviour change, before measuring training effects on the results level^{4 16} to avoid the risk of interpreting random effects independent from the training.

To establish training effects on the results level relevant to multiple stakeholders,⁹ we recommend reference to the Quadruple Aim framework.¹⁹ Beneficial training effects on the work life of HCPs may increase their motivation to implement SDM in practice. Currently, HCPs often experience SDM as another burden and demand on their time, and are therefore often reluctant to implement SDM in routine practice.⁶³ Although effects of SDM on affective-cognitive aspects of patient experience of care are well established, evidence regarding patient population health is sparse.⁴ If studies showed beneficial SDM training effects on healthcare system costs, policy makers could be encouraged to initiate system changes to foster the implementation of SDM.⁹

In sum, the poor study quality and the multitude of evaluation strategies used in identified studies limit conclusive evidence on HCP trainings in SDM. The heterogeneous use of SDM and other outcome measures compromises the interpretation and integration of research results.^{4 10 11 64 65} To achieve solid empirical evidence, we need consensus on a core set of evaluation outcomes and

Table 5 Categories of evaluation outcomes integrated in the evaluation framework

	No. of articles
Healthcare providers' reactions	17
<i>Provider-reported training appraisal</i>	16
Overall training appraisal and satisfaction	11
Appraisal of training content	5
Appraisal of training materials	3
Appraisal of training didactics	2
Appraisal of training organisation and delivery	4
Appraisal of training impact	6
Ideas for training improvement	1
<i>Objective training feasibility and acceptability data</i>	4
Healthcare providers' learning	27
<i>Provider-reported learning</i>	23
Subjective knowledge gain	3
Attitude to SDM	8
Attitude to care	3
Intention to engage in SDM	7
Confidence in SDM and communication skills	10
Confidence in medical competence	1
<i>Objective learning measures</i>	7
Healthcare providers' behaviour	26
<i>Provider-reported SDM and provider-patient interaction</i>	14
<i>Patient-reported SDM and provider-patient interaction</i>	11
<i>Standardised patient-reported SDM and provider-patient interaction</i>	2
<i>Observer-rated SDM and provider-patient interaction</i>	12
Healthcare provider training in SDM results	31
<i>Work life of healthcare providers</i>	12
Provider-reported stress and burnout	2
Provider reaction to the decision	6
Provider satisfaction with care	4
Provider-reported provider-patient relationship	2
<i>Patient population health</i>	11
Patient-reported health literacy	2
Patient-reported intention to treatment adherence	3
Patient-reported adherence	2
Patient-reported health	10
Medical records	2
<i>Patient experience of care</i>	18
Patient-reported reaction to the decision	11
Patient-reported satisfaction with care	4
Patient-reported attitude to SDM and care	8

Continued

Table 5 Continued

	No. of articles
Patient-reported provider-patient relationship	3
Provider-reported patient reaction to care	4
<i>Healthcare system costs</i>	13
Provider-reported medical practice	4
Patient-reported decisional outcome	3
Standardised patient-reported physician's final recommendation	1
Observer-recorded provider recommendation or decision	1
Healthcare resource use	2
Training costs	1
Medical record review of decision-making	1
Duration of provider-patient interaction	4

Detailed information on evaluation outcomes is attainable in online supplementary file S4.

validated measures on all levels of the proposed framework for HCP trainings in SDM. In the design of evaluation studies, researchers should aim to cover all four levels of the framework and include outcomes on the results level that relate to the Quadruple Aim framework. Researchers should aim to use outcomes that are valued by multiple stakeholders like patients, HCPs as well as healthcare managers, executives and policy makers. They should also aim to use validated observer-rated measures and objective data to limit bias, whenever feasible. If researchers applied these recommendations, evaluation studies could have more impact and better support the implementation of SDM in routine practice.^{9 65}

This review has some limitations. First, our primary search included only three databases and inclusion criteria were limited to studies aiming to evaluate HCP trainings in SDM. Consequently, we may have missed some studies, but we assume that our broad secondary search strategy made up for this limitation. Second, we did not analyse evaluation strategies regarding a match of training contents and evaluation outcomes. Additionally, we did not analyse which evaluation outcomes previously showed SDM training effects, which could be valuable information for the design of an evaluation study. However, previous studies investigated the relation between SDM and patient outcomes^{4 5 65} and interested researchers may obtain valuable information there. Third, our quality assessment with the ICROMS tool can be seen as a limitation as well as a strength of this review. On the one hand, the ICROMS tool is not applicable to post-test studies and considers patient-reported and provider-reported outcomes as unreliable, which introduces a negative bias to our quality results. On the other hand, we provided an overview of the quality of studies in the field, demonstrating a lack of robust evaluation

studies. This review has further strengths. First, this review comprises multiple study designs from post-test studies to qualitative and cluster-randomised controlled studies, which reflect the diversity of studies in the field. Second, this review provides an analysis of current strategies to evaluate HCP trainings in SDM and how their diversity functions as a barrier to conclusive evidence. Third, this review proposes an evaluation framework for HCP trainings in SDM that is based on the well-established Kirkpatrick's evaluation model and the Quadruple Aim framework. The framework may provide guidance in the design of coherence evaluation strategies for HCP trainings in SDM. Fourth, the proposed framework may initiate discussion and hopefully agreement on a core set of validated outcome measures useful for the purpose and meaningful to stakeholders.

Acknowledgements The authors would like to thank Janka Nölle (JN) for performing title and abstract screening, data extraction and quality assessments for the update of the review. The authors would like to thank Professor Dr Sigrild Harendza and Professor Dr Corinna Bergelt for their expert advice on the conduction of this study. The authors would also like to thank their student assistant Alice Diesing.

Contributors EM, IS, MH and AB conceived and planned the study. EM developed and conducted the electronic database search strategy. EM and NTD performed the secondary search strategy. EM and AS screened records. EM, AS and NTD screened full-text articles. EM and AS performed data extraction and quality assessment of included articles. EM, IS, MH, FL, NTD and AB contributed to the analysis and interpretation of study results and to the development of the evaluation framework for healthcare provider trainings in SDM. EM, AS, IS, MH, FL, NTD and AB contributed to the writing of the manuscript and approved submission.

Funding This work was partly funded by Mundipharma GmbH, a pharmaceutical company.

Disclaimer Mundipharma GmbH had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests EM and AS report grants from Mundipharma GmbH during the conduct of the study. EM, AS, IS, MH and AB conducted SDM communication skills trainings in a project funded by Mundipharma GmbH. NTD and FL have nothing to disclose.

Patient consent for publication None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Extracted data can be obtained from the corresponding author.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Härter M, Moumjid N, Cornuz J, et al. Shared decision making in 2017: International accomplishments in policy, research and implementation. *Z Evid Fortbild Qual Gesundheitswes* 2017;123:124:1–5.
- Coulter A. *Choosing wisely: learning from international experience*. Gütersloh: Bertelsmann Stiftung, 2017:1–40.
- 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;44:681–92.
- Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Med Decis Making* 2015;35:114–31.
- Clayman ML, Bylund CL, Chewning B, et al. The Impact of Patient Participation in Health Decisions Within Medical Encounters: A Systematic Review. *Med Decis Making* 2016;36:427–52.
- Hauser K, Koerfer A, Kuhr K, et al. Outcome-relevant effects of shared decision making. *Dtsch Arztebl Int* 2015;112:665–71.
- Stiggebout AM, Pieterse AH, De Haes JC. Shared decision making: Concepts, evidence, and practice. *Patient Educ Couns* 2015;98:1172–9.
- Légaré F, Adekpedjou R, Stacey D, et al. Interventions for increasing the use of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2018;7:CD006732.
- Elwyn G, Frosch DL, Kobrin S. Implementing shared decision-making: consider all the consequences. *Implement Sci* 2016;11:114.
- Légaré F, Stacey D, Turcotte S, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2014;9:CD006732.
- Légaré F, Ratté S, Stacey D, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2010;12:CD006732.
- Légaré F, Politi MC, Drolet R, et al. Training health professionals in shared decision-making: an international environmental scan. *Patient Educ Couns* 2012;88:159–69.
- Diouf NT, Menear M, Robitaille H, et al. Training health professionals in shared decision making: Update of an international environmental scan. *Patient Educ Couns* 2016;99:1753–8.
- Légaré F, Witteman HO. Shared decision making: examining key elements and barriers to adoption into routine clinical practice. *Health Aff* 2013;32:276–84.
- Tamkin P, Yarnall J, Kerrin M. *Kirkpatrick and beyond: a review of models of training evaluation*. Brighton: The Institute for Employment Studies, 2002:1–58.
- Kirkpatrick D, Kirkpatrick J. *Evaluating training programs: the four levels*. San Francisco: Berrett-Koehler, 2006:1–379.
- Légaré F, Freitas A, Thompson-Leduc P, et al. The majority of accredited continuing professional development activities do not target clinical behavior change. *Acad Med* 2015;90:197–202.
- Smidt A, Balandin S, Sigafos J, et al. The Kirkpatrick model: A useful tool for evaluating training outcomes. *J Intellect Dev Disabil* 2009;34:266–74.
- Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med* 2014;12:573–6.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
- Zingg W, Castro-Sanchez E, Secci FV, et al. Innovative tools for quality assessment: integrated quality criteria for review of multiple study designs (ICROMS). *Public Health* 2016;133:19–37.
- Bieber C, Müller KG, Blumenstiel K, et al. A shared decision-making communication training program for physicians treating fibromyalgia patients: effects of a randomized controlled trial. *J Psychosom Res* 2008;64:13–20.
- Bieber C, Müller KG, Blumenstiel K, et al. Long-term effects of a shared decision-making intervention on physician-patient interaction and outcome in fibromyalgia. A qualitative and quantitative 1 year follow-up of a randomized controlled trial. *Patient Educ Couns* 2006;63:357–66.
- Bieber C, Nicolai J, Hartmann M, et al. Training physicians in shared decision-making-who can be reached and what is achieved? *Patient Educ Couns* 2009;77:48–54.
- Härter M, Buchholz A, Nicolai J, et al. Shared decision making and the use of decision aids: a cluster-randomized study on the efficacy of a training in an oncology setting. *Dtsch Arztebl Int* 2015;112:672–89.
- Körner M, Ehrhardt H, Steger AK, et al. Interprofessional SDM train-the-trainer program "Fit for SDM": provider satisfaction and impact on participation. *Patient Educ Couns* 2012;89:122–8.
- Körner M, Wirtz M, Michaelis M, et al. A multicentre cluster-randomized controlled study to evaluate a train-the-trainer programme for implementing internal and external participation in medical rehabilitation. *Clin Rehabil* 2014;28:20–35.
- Loh A, Meier K, Simon D, et al. [Development and evaluation of a training program in shared decision making for primary care of depressive patients]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2004;47:977–84.
- Tinsel I, Buchholz A, Vach W, et al. Shared decision-making in antihypertensive therapy: a cluster randomised controlled trial. *BMC Fam Pract* 2013;14:135.
- Bieber C, Nicolai J, Gschwendtner K, et al. How Does a Shared Decision-Making (SDM) Intervention for Oncologists Affect

- Participation Style and Preference Matching in Patients with Breast and Colon Cancer? *J Cancer Educ* 2018;33:708–15.
31. Geiger F, Liethmann K, Reitz D, *et al.* Efficacy of the doktromitSDM training module in supporting shared decision making - Results from a multicenter double-blind randomized controlled trial. *Patient Educ Couns* 2017;100:2331–8.
 32. Kasper J, Liethmann K, Heesen C, *et al.* Training doctors briefly and in situ to involve their patients in making medical decisions- Preliminary testing of a newly developed module. *Health Expect* 2017;20:1254–63.
 33. LeBlanc A, Légaré F, Labrecque M, *et al.* Feasibility of a randomised trial of a continuing medical education program in shared decision-making on the use of antibiotics for acute respiratory infections in primary care: the DECISION+ pilot trial. *Implement Sci* 2011;6:5.
 34. Légaré F, Guerrier M, Nadeau C, *et al.* Impact of DECISION + 2 on patient and physician assessment of shared decision making implementation in the context of antibiotics use for acute respiratory infections. *Implement Sci* 2013;8:144.
 35. Légaré F, Labrecque M, Cauchon M, *et al.* Training family physicians in shared decision-making to reduce the overuse of antibiotics in acute respiratory infections: a cluster randomized trial. *CMAJ* 2012;184:E726–E734.
 36. Légaré F, Labrecque M, LeBlanc A, *et al.* Training family physicians in shared decision making for the use of antibiotics for acute respiratory infections: a pilot clustered randomized controlled trial. *Health Expect* 2011;14 Suppl 1:96–110.
 37. Murray MA, Stacey D, Wilson KG, *et al.* Skills training to support patients considering place of end-of-life care: a randomized control trial. *J Palliat Care* 2010;26:112–21.
 38. Stacey D, O'Connor AM, Graham ID, *et al.* Randomized controlled trial of the effectiveness of an intervention to implement evidence-based patient decision support in a nursing call centre. *J Telemed Telecare* 2006;12:410–5.
 39. Stacey D, Samant R, Pratt M, *et al.* Feasibility of training oncology residents in shared decision making: a pilot study. *J Cancer Educ* 2012;27:456–62.
 40. Towle A, Godolphin W, Grams G, *et al.* Putting informed and shared decision making into practice. *Health Expect* 2006;9:321–32.
 41. Dion M, Diouf NT, Robitaille H, *et al.* Teaching shared decision making to family medicine residents: A descriptive study of a web-based tutorial. *JMIR Med Educ* 2016;2:e17.
 42. Feng B, Srinivasan M, Hoffman JR, *et al.* Physician communication regarding prostate cancer screening: analysis of unannounced standardized patient visits. *Ann Fam Med* 2013;11:315–23.
 43. McCallister JW, Gustin JL, Wells-Di Gregorio S, *et al.* Communication skills training curriculum for pulmonary and critical care fellows. *Ann Am Thorac Soc* 2015;12:520–5.
 44. Price-Haywood EG, Harden-Barrios J, Cooper LA. Comparative effectiveness of audit-feedback versus additional physician communication training to improve cancer screening for patients with limited health literacy. *J Gen Intern Med* 2014;29:1113–21.
 45. Sullivan MD, Gaster B, Russo J, *et al.* Randomized trial of web-based training about opioid therapy for chronic pain. *Clin J Pain* 2010;26:512–7.
 46. Sullivan MD, Leigh J, Gaster B. Brief report: Training internists in shared decision making about chronic opioid treatment for noncancer pain. *J Gen Intern Med* 2006;21:360–2.
 47. Volk RJ, Shokar NK, Leal VB, *et al.* Development and pilot testing of an online case-based approach to shared decision making skills training for clinicians. *BMC Med Inform Decis Mak* 2014;14:95.
 48. Wilkes MS, Day FC, Srinivasan M, *et al.* Pairing physician education with patient activation to improve shared decisions in prostate cancer screening: a cluster randomized controlled trial. *Ann Fam Med* 2013;11:324–34.
 49. Yuen JK, Mehta SS, Roberts JE, *et al.* A brief educational intervention to teach residents shared decision making in the intensive care unit. *J Palliat Med* 2013;16:531–6.
 50. Cohen D, Longo MF, Hood K, *et al.* Resource effects of training general practitioners in risk communication skills and shared decision making competences. *J Eval Clin Pract* 2004;10:439–45.
 51. Davis RE, Dolan G, Thomas S, *et al.* Exploring doctor and patient views about risk communication and shared decision-making in the consultation. *Health Expect* 2003;6:198–207.
 52. Edwards A, Elwyn G, Hood K, *et al.* Patient-based outcome results from a cluster randomized trial of shared decision making skill development and use of risk communication aids in general practice. *Fam Pract* 2004;21:347–54.
 53. Edwards A, Elwyn G, Wood F, *et al.* Shared decision making and risk communication in practice: a qualitative study of GPs' experiences. *Br J Gen Pract* 2005;55:6–13.
 54. Bernhard J, Butow P, Aldridge J, *et al.* Communication about standard treatment options and clinical trials: can we teach doctors new skills to improve patient outcomes? *Psychooncology* 2012;21:1265–74.
 55. Butow P, Brown R, Aldridge J, *et al.* Can consultation skills training change doctors' behaviour to increase involvement of patients in making decisions about standard treatment and clinical trials: a randomized controlled trial. *Health Expect* 2015;18:2570–83.
 56. Metcalfe R, Russell R, McAvoy B, *et al.* Promoting shared decision making and informed choice for the early detection of prostate cancer: development and evaluation of a GP education program. *Cancer Forum* 2006;30:38–42.
 57. Sanders AR, Bensing JM, Essed MA, *et al.* Does training general practitioners result in more shared decision making during consultations? *Patient Educ Couns* 2017;100:563–74.
 58. Sanders ARJ, Bensing JM, Magnée T, *et al.* The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with nonspecific low back pain in primary care: a clustered randomised controlled trial. *BMC Fam Pract* 2018;19:102.
 59. Jo KH, An GJ. Effects of an educational programme on shared decision-making among Korean nurses. *Int J Nurs Pract* 2015;21:839–46.
 60. Simmons L, Leavitt L, Ray A, *et al.* Shared decision making in common chronic conditions: impact of a resident training workshop. *Teach Learn Med* 2016;28:202–9.
 61. Gärtner FR, Bomhof-Roordink H, Smith IP, *et al.* The quality of instruments to assess the process of shared decision making: A systematic review. *PLoS One* 2018;13:e0191747.
 62. Kasper J, Hoffmann F, Heesen C, *et al.* Completing the third person's perspective on patients' involvement in medical decision-making: approaching the full picture. *Z Evid Fortbild Qual Gesundhwes* 2012;106:275–83.
 63. Spatz ES, Elwyn G, Moulton BW, *et al.* Shared decision making as part of value based care: New U.S. policies challenge our readiness. *Z Evid Fortbild Qual Gesundhwes* 2017;123-124:104–8.
 64. Coulter A, Härter M, Moumjid-Ferdjaoui N, *et al.* European experience with shared decision making. *Int J Pers Cent Med* 2015;5:9–14.
 65. Oliveira VC, Ferreira ML, Pinto RZ, *et al.* Effectiveness of training clinicians' communication skills on patients' clinical outcomes: a systematic review. *J Manipulative Physiol Ther* 2015;38:601–16.