MATRIX - development and feasibility of a guide for quality assessment of patient decision aids

MATRIX - Entwicklung und Verwendbarkeit eines Leitfadens zur Qualitätsbeurteilung von Entscheidungshilfen für Patienten

Abstract

Decision aids (DAs) are interventions designed to help people make specific and deliberative choices among options by providing information about the options and outcomes that is relevant to a person's health status.

There is an ongoing discussion about the quality of DAs. The present article provides an overview on systematic approaches using various quality criteria. However, these evaluation guides are not yet implemented. Up to now quality assessment of DAs is often limited to the evidence on efficacy through controlled trials using single-outcome measures. Since DAs are multi-component interventions, single-outcome trials are not sufficient for complete quality assessment. Consideration of theoretical founding and the development process is required. In an earlier paper we proposed a novel concept of quality to meet this challenge. We introduced MATRIX a guide for quality assessment of DAs aimed at disclosing the rationale behind underpinning theories, methods, and goals of a DA.

The present paper reports how the development of MATRIX progressed including results of pre-testing and a feasibility study. We present the revised version of MATRIX, explain its basic concept, and describe the way to use it.

Keywords: decision aids, decision support techniques, methods, health care evaluation mechanisms, patient information, evidence based medicine

Matthias Lenz[1]
Jürgen Kasper[1,2]

1 Unit of Health Sciences and Education, University of Hamburg, Hamburg, Germany

2 Institute of Neuroimmunology and Clinical MS-Research (INiMS), University of Hamburg, Hamburg, Germany

Zusammenfassung

Entscheidungshilfen (Decision Aids, DAs) werden entwickelt, um Menschen zu unterstützen, gesundheits- bzw. krankheitsspezifische abwägende Entscheidungen zu treffen, indem die relevanten Informationen über die Entscheidungsoptionen und Outcomes bereitgestellt werden. Die Beurteilung der Qualität von DAs ist Gegenstand wissenschaftlicher Diskussionen. Die vorliegende Arbeit bietet einen Überblick über verschiedene systematische Ansätze der Qualitätsbeurteilung. Indes hat sich keiner dieser Ansätze bislang durchgesetzt. Die Qualitätsbeurteilung von DAs beschränkt sich auf Wirksamkeitsnachweise über kontrollierte Studien anhand einzelner Ergebnisparameter. Da DAs jedoch komplexe Interventionen sind, die aus multiplen Komponenten bestehen, kann deren Qualität mit diesem Vorgehen nicht vollständig beurteilt werden. Theoretische Fundierung und Entwicklungsprozess müssen Berücksichtigung finden. In einer früheren Arbeit haben wir ein neuartiges Konzept vorgeschlagen, das diesen Anforderungen entsprechen soll. Wir stellten MATRIX vor, eine systematische Anleitung zur Evaluation von DAs, die darauf zielt den Zusammenhang zwischen Theoriebasierung, Methoden und Zielen des DAs offen zu legen.

Mit der vorliegenden Arbeit berichten wir über Fortschritte bei der Entwicklung von MATRIX, einschließlich Pre-Test und Machbarkeitsstudien.

Wir stellen die überarbeitete Version von MATRIX vor, erklären das Grundkonzept und beschreiben, wie das Instrument verwendet wird.

Introduction

Decision aids (DAs) are interventions designed to help people making specific and deliberative choices among options by providing information about the options and outcomes that is relevant to a person's health status [31]. Detailed definitions of DAs have been published by the Cochrane Collaboration [31] and the International Patient Decision Aid Standards (IPDAS) Collaboration [15]. DAs address a variety of health decisions [30], e.g. on preventative measures, on diagnostic procedures or on treatment options. DAs appear in various forms, e.g. decision boards, booklets, interactive software and videos. They aim to achieve various goals, e.g. to enhance knowledge and to strengthen patient autonomy.

Finally, decisions are made even without a DA. So, what does a DA contribute? DAs are expected to improve the quality of decisions by strengthening patients' autonomy and comprehension [15], [31]. Following the concepts of shared decision making and informed choice a "good decision" is characterised by the extent to which it is informed and in consent with patients' personal values [10], [27], [31], [33]. DAs aim to enhance knowledge, to generate realistic expectations and satisfaction with the decisions, and to reduce decisional conflict [29].

However, since efficacy studies on DAs do not provide sufficient evidence, further research is needed [1], [27], [31].

In an earlier paper we systematically reviewed current methods of development and quality assessment of DAs [16]. We identified a number of systematic approaches [1], [2], [4], [5], [6], [10], [14], [17] covering important quality criteria. However, the approaches are not yet implemented.

The current scientific discussion about quality assessment of DAs addresses four central issues.

1.) Since DAs are complex interventions, studies using single-outcome measures alone are not sufficient for quality assessment [28], [16], [23]. The U. K. Medical Research Council (UKMRC) proposes a phased evaluation approach including both qualitative and quantitative methods [12]. Beyond efficacy and effectiveness the modelling of the intervention and the underlying theory are to be analysed [3], [12].

2.) Goal setting deserves closer attention [3], [4], [29]. It is criticised that the rationale for particular goals targeted by a DA as well as the evaluation concepts are not identifiable in the related background literature [4], [16]. Since "goal setting drives measurement" [4] outcome parameters for efficacy proof have to be precisely operationalised.

3.) It is challenged that commonly used outcome parameters are appropriate surrogates for informed choice [2], [17]. Thus, the validity of available efficacy proofs is questionable [2], [17].

4.) Although several DAs achieve desirable effects, it is not yet clear, why. It is hardly possible to interpret results from controlled efficacy trials without being able to draw conclusions about the mediating mechanisms. Developers of DAs ought to make transparent and explicit their analytical reasoning for making predictions about how DAs can be expected to achieve their goals [1], [4], [16], [27]. Consequently, a concept of quality should focus the mediating mechanisms between goals and methods.

The present article gives an overview on current quality assessment of DAs. Based on a critical appraisal of identified evaluation systems we introduce a new concept of quality. Results of pre-testing and of a feasibility trial of this novel assessment guide are reported.

Quality assessment of decision aids

Three evaluation systems were identified, designed to structure the issue of DA quality implying both, guidance for development processes and criteria for quality appraisal.

The 'CREDIBLE'-criteria

The *Ottawa Health Research Institute* (OHRI) provided 'CREDIBLE', a checklist for quality assessment of DAs (Table 1) [31]. The evaluation guide was used by the OHRI as a standardized reporting format to communicate the quality of DAs [31], until the OHRI recently switched to the IPDAS criteria. The CREDIBLE-criteria were used for the quality assessment of DAs within the Cochrane review [31].

Table 1: The 'C.R.E.D.I.B.L.E.' criteria

The 'C.R.E.D.I.B.L.E.' criteria:	
C	- **C**ompetent developers and development;
R	- **R**ecent update;
E	- **E**vidence-based;
DI	- **D**evoid of conflicts of **I**nterest;
BL	- **BaL**anced presentation of options; benefits and harms; and
E	- **E**fficacy at improving decision making.

The IPDAS criteria

Aiming to establish international standards for development and evaluation of DAs the IPDAS collaboration intends to achieve a consensus on a set of minimal criteria for quality assessment [13]. Beyond an appraisal of effectiveness further evaluation criteria are included, e.g. characteristics of the development process and methods

to present evidence based information. The quality criteria, proposed by the IPDAS collaboration [7] resulted from a broad validation procedure, initially based on the adoption of the 'CREDIBLE'- items. These were completed and met in an international voting on the importance of each criterion. By this procedure the collection of relevant criteria became more comprehensive. The resulting guide contains 80 criteria in 12 quality domains (Table 2). Furthermore, the IPDAS collaboration published an exhaustive collection of evidence on their quality aspects [14]. Each quality aspect is theoretically explained and reviewed by a group of experts.

Table 2: IPDAS - Categories of quality items

IPDAS - Categories of quality items [7]
1. *Using a systematic development process*
2. *Providing information about options*
3. *Presenting probabilities*
4. *Clarifying and expressing values*
5. *Using patient stories*
6. *Guiding / coaching in deliberation and communication*
7. *Disclosing conflicts of interests*
8. *Delivering patient decision aids on the Internet*
9. *Balancing the presentation of options*
10. *Using plain language*
11. *Basing information on up-to-date scientific evidence*
12. *Establishing effectiveness*

'CREDIBLE' and IPDAS criteria are item checklists. They cover important aspects of quality in some particular contexts but the underlying concept of quality is not explained. Which criteria are essential for a high quality DA? Does comprehensive performance in all 80 items strongly indicate high quality? The IPDAS criteria guide a reviewer to investigate and to describe the DA in detail but they do not provide instructions how to proceed if criteria are more ore less important for the critical appraisal of a particular DA.

The Ottawa workbook

The '*Workbook on Developing and Evaluating Patient Decision Aids*' (in the following *workbook*) published by O'Connor and Jacobsen is a detailed and systematic manual to develop and evaluate DAs (Table 3) [29].

Table 3: The Ottawa workbook – seven steps of development and evaluation of decision aids

The Ottawa workbook – seven steps of development and evaluation of decision aids:
1) Assess need
2) Assess feasibility
3) Define the objectives of the aid
4) Identify the framework of decision support
5) Select the methods of decision support to be used in the aid
6) Select the designs and measures to evaluate the aid
7) Plan dissemination

The first sequence (steps 1, 2 and 3) guides the developer as well as the reviewer to analyse the decision context, e.g. the specific difficulties of the decision making, the needs of the target group or the availability of evidence. The goals have to be clear, specific, and measurable. The guidance includes the consideration, that the goals determine the outcome parameters of the efficacy trial.

The second sequence of evaluation (steps 4 and 5) attends to the selection of the specific decision support methods referring to the goals of the DA. In step 4 the developer and the reviewer, respectively, are recommended to consider a *decision support framework* as a suitable theoretical background to deduce the methods of the DA. Examples of frameworks are suggested: the concept of '*Shared Decision Making*' [5]; the concept of '*Evidence Based Patient Choice*' [8], [9]; the '*Motivation Theory*' and '*Theory of Reasoned Action*' [32]; the '*Ottawa Decision Support Framework*' (ODSF) [29]. A concrete procedure how to use a *decision support framework* to arrive at the methods is not outlined.

The ODSF is described in detail in the *workbook*. It proposes ascertained strategies to solve three defined categories of decisional problems. E.g. *unrealistic expectations (perceived probabilities of outcomes) may be realigned by presenting probabilities of outcomes that are tailored to the patient's clinical risk and by describing outcomes that they are easy to imagine and identify with* [29].

Step 5 provides an overview of common contents and methods of DAs. Concrete strategies are exemplarily described, e.g. to present contents and probabilities of outcomes, methods of value clarification, coaching and communication in decision making, methods of delivering DAs and preparing practitioners. The developer of a DA gets access to the wide practical experience of the OHRI, but the theoretical founding of the methods described remains unclear.

Within the third sequence (steps 6 and 7) procedures to develop and methods to evaluate a DA are recommended. The authors of the *workbook* emphasise that development and evaluation depend on the objectives of DAs. De-

velopers need to decide on the sampling and design architecture, the criteria for evaluation and the measurement tools that will be used to operationalise the criteria. In contrast to the checklist approach of 'CREDIBLE' and the IPDAS criteria the *workbook* focuses the development procedure. With the *workbook* another concept of quality was introduced. It allows appraisal of a variety of formats of DAs. Furthermore, it provides a structure for the evaluation procedure comparable to the framework of the UKMRC [12], considering that DAs are complex interventions.

However, it is not explained how to use a theoretical framework for decision support when deciding on methods and mediating mechanisms.

Introduction of a new evaluation guide

In an earlier paper we introduced MATRIX as a first version of a new guide for evaluation of DAs [16], based on the *workbook* of O'Connor and Jacobsen [29]. The sequence of selecting methods proposed in the *workbook* was restructured, and the UKMRC approach of evaluation of complex interventions [12] was implemented. With the new approach we aimed at supporting the reviewer to retrace the methodological decisions made by the developer. A reflection matrix was developed to support this reviewing process by correlating the goals of a DA with decisions made on effect mediators. Effect mediators are design features (e.g. the contents, the setting, the presentation form, and the media) mediating the intended effects of the DA [16].

The reflection matrix was embedded into a systematic evaluation guide, designed to support the reviewer exploring the rationale of goals and methods and critically appraising the evidence on efficacy and on effectiveness of a DA.

We already discussed the underpinning theory and described the modelling of its components [16]. In the following we report how the development of MATRIX progressed, finally leading to an evaluation guide. MATRIX is primarily designed to support quality appraisal, which is emphasised in the present evaluation trials. However, since quality appraisal and development are closely interrelated in the MATRIX concept, it is intended for guidance of structured development of DAs as well.

Develoment of MATRIX

The current version of MATRIX is based on a five steps evaluation procedure (Table 4), which is also reported in the following. The first version of MATRIX underwent pre-testing (September to October 2004) and two feasibility studies (May to July 2005 and March to September 2006). The CREDIBLE guide was chosen as reference standard.

Table 4: Development procedure of MATRIX

Development procedure of MATRIX
1. Review and summary of current research
2. Theoretical phase, development of a concept of quality (expert panel)
3. Modelling phase and pre-testing of the first version of MATRIX (by use of in depth interviews)
4. Revision and feasibility studies of the MATRIX guide (by questionnaire)
5. Revision and development of the current MATRIX guide

1. Pre-testing and first revision of MATRIX

In the first study (pre-testing), two *Ottawa DAs* [19] were quality assessed by four external experts in the field of shared decision making. The first version of MATRIX [16] and CREDIBLE were used as evaluation guides. Background publications of the DAs were provided [18], [24], [25], that had been identified in systematic database searches [22]. To prevent order effects, two experts were instructed to use the MATRIX version first and then the CREDIBLE-Criteria, two were instructed in the opposite order. In-depth interviews were conducted analysing the feasibility and differences in the evaluation procedures. Following the experts' comments MATRIX underwent a revision focused on reduction of complexity and textual simplification.

2. First feasibility trial and second revision of MATRIX

One *Ottawa DA* [19] was quality assessed by 15 experts in the field of shared decision making using the revised version of MATRIX and CREDIBLE. The objective of this study was to investigate whether MATRIX is understandable and feasible. The guide used in the study was in German language. The participating experts were recruited by personal telephone contact to German experts associated with research in the field of shared decision making, scientists, consultants and policymakers. Inclusion criteria were English language competence, knowledge about DAs, ability to respond in a given two week's period. No participants were recruited out of the authors' research group. The background publication [18] of the Ottawa DA was provided for assessment. To prevent order effects the order to use the evaluation guides was randomised. Characteristics of the evaluation procedure were surveyed by a 13-item questionnaire, which the experts were asked to complete after reviewing the DA. The first 7 items addressed the "*comprehensibility*" and "*feasibility*" of the guides. Furthermore, in order to inform the interpretation of these feasibility ratings, a set of 6 items was added to the questionnaire. The participants were asked to express their attitudes towards underpin-

ning quality concepts and the quality of assessments generated by use of each evaluation guide. Answering format of the questionnaire was a four point Likert-scale combined with an open answering space (Table 5). Participants were asked to rate their agreement to a given statement ranging from 'very' to 'not at all'. Mean differences between the ratings lower than 0.3 were regarded as equal.

Table 5: Answering format of the questionnaire items

Answering format of the questionnaire items							
CREDIBLE-Criteria					MATRIX		
Not at all			Very		Not at all			Very
☐	☐	☐	☐		☐	☐	☐	☐
Comment:							

Out of 15 experts 14 completed the questionnaire. The participants' expertise in DA-development was rated on a four point scale between 1 = newcomer to 4 = developer. The expertise scored by 2.5 in mean (mean=2.5, SD=1) in the middle of the scale. The quantitative results are presented in Table 6.

In this study, CREDIBLE was rated higher compared to MATRIX guide concerning "comprehension of the contents", "answerability of the questions" and "effort benefit relation". Many comments were made regarding the "answerability of the questions" and the "availability of background information" on the DA. Some found fault with MATRIX as they were not able to identify the required information. Others commented this phenomenon by mentioning MATRIX would "uncover a lack of background information in the literature" (n=6). Some participants (n=6) mentioned "the higher effort of MATRIX was adequate regarding its advantages".

No differences in mean values were found in "structural comprehension", "clarity of the basic concept of DA quality" and "utility of the systematic approach". Nevertheless, polarised statements were found, representing the controversial discussion on these issues. Some experts (n=6) complain "difficulties to apply the reflection matrix" (Section B of MATRIX). Other comments (n=5) referred to CREDIBLE, as its implicit structure was perceived as neither traceable nor exhaustive.

In sum, the MATRIX approach was expected to result in more "systematic" and "complete" reviews and to improve "accuracy", "transparency" and "validity" of the reviews. A part group appreciated the new quality concept mentioning its accuracy and exhaustive approach. One participant expected MATRIX to "stimulate development of highly adaptive DAs in an evolutionary sense".

To address remaining difficulties in comprehension of the terminology, the underpinning quality concept, the MATRIX guide was revised in terms of simplification.

3. Second feasibility trial

To retest the revised version of the MATRIX guide a second feasibility trial was undertaken. Again, comprehension and feasibility were objects of research. 24 scientists and health care stake holders from Germany, the Netherlands and the UK were asked for participation. Finally 15 participants - 10 of them already participants in the previous trial - completed the task. One Ottawa DA [21] had to be appraised supported by MATRIX and the CREDIBLE guide, which both were provided in English language. The MATRIX translation had been revised by a native speaking scientist and validated by retranslation. The application sequence of the evaluation guides was randomised. Background publications [11], [20] of the Ottawa DA were provided for appraisement, which had previously been identified in systematic database searches [22]. To evaluate the appraisal procedures the same questionnaire as in the first trial was used. The questionnaire was provided online. The participants' expertise was assessed. Two of them described themselves as developers, six as newcomers in the field. Since our study particularly focused the evaluation processes guided by CREDIBLE and MATRIX, the reviews themselves or the assessments` results concerning the evaluated DAs were not collected.

The pattern of results (Table 7) was similar to the first feasibility trial. The comments revealed two main difficulties when using the MATRIX guide: firstly, that information defined as essential for evaluating the DA was not available; secondly, to understand and use the reflection matrix (Section B of MATRIX). The latter was commented e.g.: "Why does MATRIX not consider basic criteria as 'conflict of interest' or 'update'". However, the reflection matrix is designed to appraise such criteria.

The subgroup analysis of the 10 reviewers participating in both studies shows that revision of MATRIX did not change "comprehension" and "feasibility". Focussing the attitudes of the 10 reviewers a shift to disagreement with the MATRIX concept becomes apparent. In particular, they sceptically commented the impact of MATRIX on "accuracy" and "traceability" of the reviews. The attitudes towards the validity of the quality concept remained unchanged compared to CREDIBLE.

Table 6: Results of feasibility trial 1

	Total entries					no entry
	0 (not at all)	**1**	**2**	**3** (very)	**mean**	
Feasibility						
1. Could you understand the content of the guide?						
- CREDIBLE	0	1	4	9	2.6	0
- MATRIX	0	5	9	0	1.6	0
2. Could you follow the structure of the guide?						
- CREDIBLE	0	3	4	7	2.3	0
- MATRIX	0	3	6	5	2.1	0
3. Is the basic concept of quality of decision aid noticeable?						
- CREDIBLE	0	2	6	6	2.3	0
- MATRIX	0	2	7	5	2.2	0
4. Are the items/questions answerable?						
- CREDIBLE	0	4	7	3	1.9	0
- MATRIX	0	9	4	1	1.4	0
5. You received both, the decision aid and the article in which it was published. Were these two resources sufficient to allow you to judge its quality?						
- CREDIBLE	0	2	0	10	1.8	2
- MATRIX	0	4	8	0	1.7	2
6. Does the systematic approach of the guide facilitate the quality assessment?						
- CREDIBLE	1	2	7	4	2.0	0
- MATRIX	1	2	6	5	2.1	0
7. Is the effort to complete the guide balanced by the value of completing it?						
- CREDIBLE	0	1	6	7	2.4	0
- MATRIX	0	6	4	4	1.9	0
Attitudes towards the evaluation concepts						
8. By using the guide, was your quality assessment thorough as could be?						
- CREDIBLE	1	3	7	2	1.8	1
- MATRIX	0	4	3	6	2.2	1
9. By using the guide, was your quality assessment exhaustive as it could be?						
- CREDIBLE	2	3	8	1	1.6	0
- MATRIX	1	3	5	5	2.0	0
10. By using the guide, was your quality assessment accurate as possible?						
- CREDIBLE	0	7	4	2	1.6	1
- MATRIX	0	2	5	6	2.3	1
11. By using the guide, are you able to justify the reasons for your assessment?						
- CREDIBLE	2	0	11	1	1.8	0
- MATRIX	0	2	4	8	2.4	0
12. By using the guide, would different reviewers achieve the same rating of decision aid quality (interrater-reliability)?						
- CREDIBLE	0	3	6	2	1.9	3
- MATRIX	1	3	6	1	1.6	3
13. Did the guide allow you to assess the decision aid quality?						
- CREDIBLE	0	3	4	7	2.3	0
- MATRIX	0	2	5	7	2.4	0

Table 7: Results of feasibility trial 2

	Total entries					no entry
	0 (not at all)	**1**	**2**	**3** (very)	**mean**	
Feasibility						
1. Could you understand the content of the guide?						
- CREDIBLE	0	1	7	7	2.4	4
- MATRIX	0	8	6	1	1.5	4
2. Could you follow the structure of the guide?						
- CREDIBLE	0	1	7	7	2.4	4
- MATRIX	0	4	4	7	2.2	4
3. Is the basic concept of quality of decision aid noticeable?						
- CREDIBLE	0	0	12	2	2.1	5
- MATRIX	0	3	5	6	2.2	5
4. Are the items/questions answerable?						
- CREDIBLE	0	3	4	8	2.3	4
- MATRIX	0	8	6	1	1.5	4
5. You received both, the decision aid and the article in which it was published. Were these two resources sufficient to allow you to judge its quality?						
- CREDIBLE	0	2	8	5	2.2	4
- MATRIX	2	3	6	3	1.7	5
6. Does the systematic approach of the guide facilitate the quality assessment?						
- CREDIBLE	0	2	9	4	2.1	4
- MATRIX	0	3	7	4	2.1	5
7. Is the effort to complete the guide balanced by the value of completing it?						
- CREDIBLE	0	3	7	4	2.1	5
- MATRIX	1	5	5	3	1.7	5
Attitudes towards the evaluation concepts						
8. By using the guide, was your quality assessment thorough as could be?						
- CREDIBLE	0	5	8	2	1.8	4
- MATRIX	0	5	6	4	1.9	4
9. By using the guide, was your quality assessment exhaustive as it could be?						
- CREDIBLE	0	8	6	1	1.5	4
- MATRIX	0	4	7	4	2.0	4
10. By using the guide, was your quality assessment accurate as possible?						
- CREDIBLE	0	6	7	2	1.7	4
- MATRIX	0	7	7	1	1.6	4
11. By using the guide, are you able to justify the reasons for your assessment?						
- CREDIBLE	0	2	8	5	2.2	4
- MATRIX	0	7	3	3	1.9	4
12. By using the guide, would different reviewers achieve the same rating of decision aid quality (interrater reliability)?						
- CREDIBLE	2	2	4	6	2.0	5
- MATRIX	1	5	8	0	1.5	5
13. Did the guide allow you to assess the decision aid quality?						
- CREDIBLE	0	2	11	2	2.0	4
- MATRIX	0	6	6	3	1.8	4

A. Appropriateness of goals

'Goal setting ought to drive measurement and not the other way around '

Consider publications referring to the background theory, modelling of the DA and exploratory trials, to identify the goals of the DA.

▪ Importance within the particular decision making context:

A1: Does an explicit rationale for the selection of the goals exist? (personal, socio-cultural, and clinical context)

▪ Precision and measurability:

A2: Are the goals precisely worded?

A3: Are the goals defined in terms of measurable dimensions? (E.g. if the goal is empowerment, measurable dimensions may be social- and self-competence, knowledge, and sense of responsibility)

B. Appropriateness of effect mediators

'Effect mediators (e.g. methods, media and design) refer to the goals'. What are the mechanisms mediating to the *goals?* Use the matrix to explore the correlation between the particular goals and the effect mediators of the DA.

B1: List the particular goals of the DA into the left column.

B2: Identify the effect mediators and allocate them into the upper row.

B3: Within each cell of the matrix, frame questions to explore the mediating mechanisms by which the effect mediators are expected to achieve the goals.

   E.g. *"in consideration of the particular goal…*

   ...is the used presentation form justified by theory/by evidence out of studies/by ethical considerations...?"

   ...is the complexity of the DA justified by theory/by evidence out of studies/by ethical considerations...?"

   ...is the used medium justified by theory/by evidence out of studies/by ethical considerations...?"

B4: Answer your questions and ascertain the mediating mechanisms.

Goals (precise and measurable)	Categories of *effect mediators*					
	Contents: ⋮	Structure: ⋮	Complexity: ⋮	Setting: ⋮	Presentation: ⋮	Media: ⋮
1. ...						
2. ...						
3. ...						

C. Efficacy and effectiveness

Consider randomised controlled trials and implementation trials to appraise the effectiveness of the DA under controlled and uncontrolled conditions.

▪ Outcome measures:

C1. Are the outcome measures patient relevant and do they represent the particular goals of the DA?

C2. Are the dimensions of goals operationalised appropriately?

C3. Are the outcome measures considered completely (e.g. side effects)?

▪ Study design:

C4. Are the designs and methods (randomisation, allocation, blinding etc.) of the studies evidentiary?

▪ Effectiveness and implementation:

C5. To which extent and referring to which outcomes is the DA effective?

C6. Is the DA effective under uncontrolled conditions (reproducibility, external validity)?

▪ Interpretation:

C7. Do the data on effectiveness of the DA support the intended mediating mechanisms?

Figure 1: MATRIX - systematic evaluation of decision aids

Quality assessment using MATRIX

In the following we present the current version of MATRIX as it was revised after the first feasibility study and re-tested in the second study. The evaluation procedure with MATRIX follows three steps (A, B, and C) of system-atic appraisal (Figure 1). MATRIX supports the reviewer to collect and appraise information relevant for quality judgement comprehensively. However, the guide does not provide an instruction for the reviewer how to proceed when weighting and integrating the appraisals to come to a judgement. An example for the use of MATRIX to review a decision aid is provided in Attachment 1.

A. Appropriateness of the goals

The quality of a DA is limited by the appropriateness of its goals. They are appropriate, if there is a rationale to see them as important in a particular decision making context. Usually, important outcome measures are the operationalised goals [4].

The goals should be defined in terms of measurable dimensions. For example: if the goal is empowerment, measurable dimensions can be social- and self-competence, knowledge, and sense of responsibility. This allows operationalisation into one or more outcome measures [4]. If the goal is to support "informed choice" in terms of the approach of Marteau et al., a multidimensional outcome measure can be used [26]. This measure comprises an eight-item scale of knowledge, a four-item scale assessing attitudes towards undergoing the screening test and a record of test uptake.

To identify the goals of the DA, publications referring to the underlying theory, modelling of the DA and exploratory trials are to be considered.

B. Appropriateness of effect mediators - the use of the reflection matrix

The reflection matrix unfolds between the goals of a DA and categories of effect mediators (contents, structure, complexity, setting, presentation form, and used media). The rationale underlying the relation between goals and effect mediators should be appraised critically, by reflecting on the mechanisms by which a developer expects to achieve the goals. It should be proved whether these mechanisms are explained and whether they are traceable and supported empirically.

The effect mediators should be assessed considering the founding theory, plausibility (e.g. time-frame, target group), and ethical standards and pre-studies conducted by the developers in the target group. Practically, the reviewer is guided from cell to cell of the reflection matrix to assess, whether the decisions to use certain effect mediators are justified (Figure 1). Within each cell, questions can be framed to explore the mechanisms by which the effect mediators are expected to achieve the goals. Three examples: 1) "In consideration of the particular goal, is the used presentation form justified by theory or by evidence out of studies or by ethical considerations?" 2) "In consideration of the particular goal, is the degree of complexity of the DA justified by theory or by evidence out of studies or by ethical considerations?" 3) "In consideration of the particular goal, is the used medium justified by theory or by evidence out of studies or by ethical considerations?". The comprehensive analysis of the goal-methods interrelation which is represented in the full matrix provides the basis for the reviewer's judgement.

C. Efficacy and effectiveness

The DA is effective if the goals defined by the authors of the DA have been achieved. Therefore, randomised controlled trials and implementation trials needed to be conducted [12]. The reviewer should critically appraise studies on efficacy, effectiveness and implementation. In particular, it should be proved, whether the outcome measures used in the trials are patient relevant and whether they represent the particular goals of the DA, whether the dimensions of goals were operationalised appropriately (e.g. if the goal is empowerment, measurable dimensions may be social- and self-competence, knowledge, and sense of responsibility), and whether the data on effectiveness support the intended mediating mechanisms.

Discussion

Quality assessment of DAs is a challenging endeavour. DAs address a variety of health decisions, appear in various forms, and aim to achieve various goals. Quality assessment needs to consider the specificity of DA concepts. The CREDIBLE guide [31] is easy to use, but does not consider all important quality criteria. The *IPDAS-evaluation criteria* [7] and the *Ottawa workbook* [29] include important quality criteria for DAs but do not provide a transparent step by step strategy to systematically analyse the mediating mechanisms intended by the methods of a DA.

The use of DAs is expanding. New concepts of DAs have been developed and various approaches of decision support are conceivable. To include potential approaches, a concept of quality needs to go beyond the scope of status quo. MATRIX meets this challenge by focusing the traceability of the development process and the inherent rationale of the DA. In this respect, high quality means, that the mediating principle between goals and methods is justified by evidence, theory or plausibility.

Based on the MATRIX approach, a guide for reviewers has been developed to analyse the effectiveness of DAs in terms of complex interventions. That requires access to information about the development process [3], [12], [16], [22], [23].

Results of the feasibility trials show that the MATRIX guide was predominantly rated to be more complicated. The comments, given by the participants were insightful to understand the problems of MATRIX. Firstly, some important quality aspects could not be assessed, because of unavailable background information. That might be unsatisfying for a reviewer. However, unavailability of information can be due to limited search strategies or due to unpublished data. Secondly, some quality criteria of MATRIX were perceived to be too abstract, e.g. "does an explicit rationale for the selection of the goals exist?" In particular, if a rationale is not identifiable, it might be difficult for the reviewer to get an idea what a good per-

formance would look like. There still seems to be a need for more guidance and explanation in the MATRIX guide. Thirdly, by focusing the development process of the DA, MATRIX seems to interfere with intuitive evaluation strategies, which primarily tend to focus the final product (the DA). This can cause cognitive dissonance: e.g. a DA can appear on high standard; however, without sufficient background information the DA would be critically appraised as insufficient due to lacking traceability.

Unfortunately, it is at present not readily possible to identify the background information of currently available DAs with commonly used database search strategies [22]. The MATRIX approach emphasizes once more that systematic access to such information is necessary.

MATRIX may also facilitate methodological research, representing a framework for existing and non-existing evidence: if frequently targeted goals (e.g.: information, value clarification, participation [29]) and categories of effect mediators (e.g. setting, content, presentation, and media) are inserted into the reflection matrix, the existing evidence could be arranged within the cells and gaps of research become apparent. E.g. while within the cell "information X presentation" comprehensive evidence might exist, the cell "value clarification X media" might appear empty. An empty cell indicates the need for research on this particular issue.

MATRIX is still under development and evaluation. It exists in a print version. We plan to develop an interactive software version, which is expected to be more suitable to unfold the reflection MATRIX. The evaluation guide then has to pass further feasibility testing. Outcome parameters need to be determined for a randomised controlled trial, to generate evidence whether MATRIX influences traceability and validity of the reviews, and inter-reviewers' reliability.

The use of MATRIX requires an intensified attention and more time than a check of standard criteria. This effort has to be valued with regard to the state of research on DA quality. If evidence indicated an existing high quality standard of current DAs in general, there would be no need for implementation of a more systematic concept. However, at the moment we do not know enough about the efficiency of methods to support lay people's medical decision making.

Notes

Authorship

Both authors have comparably contributed to this paper.

Conflicts of interest

None declared.

Attachments

Available from

http://www.egms.de/en/journals/psm/2007-4/psm000041.shtml

1. GMS-Lenz-Kasper-appendix.pdf (87.112 KB) Appendix: Example of a systematic evaluation of a decision aid using MATRIX

References

1. Bekker HL, Hewison J, Thornton JG. Understanding why decision aids work: linking process with outcome. Patient Educ Couns. 2003;50(3):323-29.

2. Bekker HL, Legare F, Stacey D, O'Connor A, Lemyre L. Is anxiety a suitable measure of decision aid effectiveness: a systematic review? Patient Educ Couns. 2003;50(3):255-62.

3. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. BMJ. 2000;321(7262):694-6.

4. Charles C, Gafni A, Whelan T, O'Brien MA. Treatment decision aids: conceptual issues and future directions. Health Expect. 2005;8(2):114-25.

5. 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.

6. Dowie J. The role of patients' meta-preferences in the design and evaluation of decision support systems. Health Expect. 2002;5(1):16-27.

7. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, Thomson R, Barratt A, Barry M, Bernstein S, Butow P, Clarke A, Entwistle V, Feldman-Stewart D, Holmes-Rovner M, Llewellyn-Thomas H, Moumjid N, Mulley A, Ruland C, Sepucha K, Sykes A, WhelanT. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ. 2006;333(7565):417.

8. Entwistle VA, Sheldon TA, Sowden A, Watt IS. Evidence-informed patient choice. Practical issues of involving patients in decisions about health care technologies. Int J Technol Assess Health Care. 1998;14(2):212-25.

9. Entwistle VA, Sowden AJ, Watt IS. Evaluating interventions to promote patient involvement in decision-making: by what criteria should effectiveness be judged? J Health ServcRes Policy. 1998;3(2):100-7.

10. Feldman-Stewart D, Brundage MD. Challenges for designing and implementing decision aids. Patient Educ Couns. 2004;54(3):265-73.

11. Grant FC, Laupacis A, O'Connor AM, Rubens F, Robblee J. Evaluation of a decision aid for patients considering autologous blood donation before open-heart surgery. CMAJ. 2001;164(8):1139-44.

12. Health Services and Public Health Research Board. A Framework for Development and Evaluation of RCTs for Complex Interventions to Improve Health. United Kingdom Medical Research Council. Available from: http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003372. 2000 (Accessed at 08-20-2007).

13. International Patient Decision Aid Standards (IPDAS) Collaboration. International Patient Decision Aid Standards. Available from: http://ipdas.ohri.ca /. 2005 (Accessed at 05-15-2007).

14. International Patient Decision Aid Standards (IPDAS) Collaboration. IPDAS Collaboration Background document. Available from: http://ipdas.ohri.ca /IPDAS_Background.pdf. 2005 (Accessed at 05-15-2007).

15. International Patient Decision Aid Standards (IPDAS) Collaboration. What are Patient Decision Aids. Available from: http://ipdas.ohri.ca /what.html. 2005 (Accessed at 05-15-2007).

16. Kasper J, Lenz M. Criteria for the development and evaluation of decision aids. Z Ärztl Fortbild Qualitätssich. 2005;99(6):359-65.

17. Kennedy AD. On what basis should the effectiveness of decision aids be judged? Health Expect. 2003;6(3):255-68.

18. Lalonde L, O'Connor AM, Drake E, Duguay P, Lowensteyn I, Grover SA. Development and preliminary testing of a patient decision aid to assist pharmaceutical care in the prevention of cardiovascular disease. Pharmacotherapy. 2004;24(7):909-22.

19. Lalonde L. Life changes to lower your risk of heart disease and stroke. Available from: http://decisionaid.ohri.ca/docs/das /Cardiovascular.pdf. 2002 (Accessed at 05-15-2007).

20. Laupacis A, O'Connor AM, Drake ER, Rubens FD, Robblee JA, Grant FC et al. A decision aid for autologous pre-donation in cardiac surgery--a randomized trial. Patient Educ Couns. 2006; 61(3):458-66.

21. Laupacis A, O'Connor AM, McAlister F, Grant F. Making Choices: Blood Transfusion in Heart Surgery. Available from: http://decisionaid.ohri.ca/docs/das /Blood_Transfusion.pdf. 1999 (Accessed at 05-15-2007).

22. Lenz M, Kasper J, Mühlhauser I. Searching for diabetes decision aids and related background information. Diabet Med. 2006; 23(8):912-6.

23. Lenz M, Steckelberg A, Richter B, Mühlhauser I. Systematic reviews of complex interventions - a complex issue: Analysis of systematic reviews of diabetes and hypertension self-management programmes. Diabetologia. 2007;50(7):1375-83.

24. Man-Son-Hing M, Laupacis A, O'Connor AM, Biggs J, Drake E, Yetisir E, Hart RG. A patient decision aid regarding antithrombotic therapy for stroke prevention in atrial fibrillation: a randomized controlled trial. JAMA. 1999;282(8):737-43.

25. Man-Son-Hing M, Laupacis A, O'Connor AM, Hart RG, Feldman G, Blackshear JL, Anderson DC. Development of a decision aid for atrial fibrillation who are considering antithrombotic therapy. J Gen Intern Med. 2000;15(10):723-30.

26. Marteau TM, Dormandy E, Michie S. A measure of informed choice. Health Expect. 2001;4(2):99-108.

27. Molenaar S, Sprangers MA, Postma-Schuit FC, Rutgers EJ, Noorlander J, Hendriks J, de Haes HC. Feasibility and effects of decision aids. Med Decis Making. 2000;20(1):112-27.

28. Mühlhauser I, Berger M. Patient education - evaluation of a complex intervention. Diabetologia. 2002;45(12):1723-33.

29. O'Connor AM, Jacobsen MJ. Workbook on Developing and Evaluating Patient Decision Aids. Ottawa Health Research Institute. Available from: http://decisionaid.ohri.ca/docs/develop /Develop_DA.pdf. 2003 (Accessed at 05-15-2007).

30. O'Connor AM, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, Entwistle V, Barry M. Cochrane inventory and evaluation of patient decision aids. Medical Decision Making 2001;21(6):527.

31. O'Connor AM, Stacey D, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, Entwistle V, Rostom A, Fiset V, Barry M, Jones J. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2001(3).

32. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Fiore C, Harlow LL, Redding CA, Rosenbloom D, Rossi SR, Rakowski W. Stages of change and decisional balance for twelve problem behaviors. Health Psychology. 1994;3(1):39-46.

33. Whelan T, Gafni A, Charles C, Levine M. Lessons learned from the Decision Board: a unique and evolving decision aid. Health Expect. 2000;3(1): 69-76.

Corresponding authors:

Matthias Lenz

University of Hamburg, Unit of Health Sciences and Education, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany, Tel.: +49-40-42838-7232, Fax: +49-40-42838-3732

matthias.lenz@uni-hamburg.de

Jürgen Kasper, PhD

University of Hamburg, Unit of Health Sciences and Education, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany, Tel.: +49-40-42838-7231, Fax: +49-40-42838-3732

k@sper.info