2020 UPDATE CHAPTER L: ESTABLISHING THE EFFECTIVENESS OF PATIENT DECISION AIDS

SECTION 1: AUTHORS/AFFILIATIONS

Karen Sepucha (co-lead)	Harvard Medical School, Boston MA	USA
Richard Thomson (co-lead)	Newcastle University, Newcastle Upon Tyne	UK
Jennifer Blumenthal-Barby	Baylor College of Medicine, Houston TX	USA
Jesse Jansen	Maastricht University, Maastricht	The Netherlands
Mirjam Körner	Albert-Ludwigs University Freiburg, Freiburg	GER
Joanne Lally	Newcastle University, Newcastle upon Tyne	UK
Daniel D. Matlock	University of Colorado, Denver CO	USA
Lilisbeth Perestelo-Perez	Canary Islands Health Service, Tenerife	Spain
Mary Ropka	University of Virginia. VA	USA
Christine Stirling	University of Tasmania, Hobart	Australia
Logan Trenaman	University of California, Davis CA	USA
Kathrene Valentine	Harvard Medical School, Boston MA	USA
Ha Vo	Massachusetts General Hospital	USA
Celia E. Wills	Ohio State University, Columbus OH	USA

Suggested citation: Sepucha K, Thomson R, Blumenthal-Barby J, Jansen J, Korner M, Lally J, Matlock D, Perestelo-Perez L, Ropka M, Stirling C, Trenaman L, Valentine K, Vo H, Wills C. Chapter L: Evaluating Effectiveness of Patient Decision Aids. International Patient Decision Aids Standards (IPDAS). 2020. https://decisionaid.ohri.ca/ipdas/resources.html. Accessed on DD/MM/YYYY.

SECTION 2:
CHAPTER SUMMARY

This chapter updates the previous version as published in 2013 (see Sepucha et al. 2013)

What is this Dimension?

To establish the effectiveness of a patient decision aid (PtDA), it is critical to provide evidence that the PtDA improves: i) the quality of the decision-making process; and ii) the quality of the choice that is made (i.e., "decision quality").

There is broad recognition that appropriate measures of a good decision are separate from traditional clinical/health outcomes (although it is important that these are also measured). The *quality of the decision-making process* refers to the extent to which a PtDA helps patients to: recognize that a decision needs to be made; feel informed about the options and their characteristics; be clear about what matters most to them in this decision; discuss goals, concerns, and preferences with their health care providers; and be involved in decision making. *The quality of the decision choice that is made* (often called "decision quality") is defined as the extent to which a patient's eventual choice is consistent with their informed values and is actually implemented.

What is the Theoretical Rationale for Including this Dimension?

There are good reasons for including a chapter on the dimension of effectiveness of PtDAs. This is distinct from the implementation of PtDAs, which is covered in another chapter. In this chapter, we have organized this rationale into scientific, ethical, conceptual, and health policy sub-sections. We have also clearly indicated the different rationales for measuring how PtDAs impact the quality of the decision-making process and the quality of the choice that is made.

What is the Evidence to Support Including or Excluding this Dimension?

This updated overview of the evidence incorporates: a) the results of the 2017 Cochrane Collaboration's review of the 105 randomized controlled trials of PtDAs (Stacey et al 2017); b) an examination of the reported performance of instruments used to measure primary process and outcome variables in studies included in the 2014 and 2017 Cochrane Collaboration's reviews; and c) good practice examples of how investigators should report on the performance of instruments used to measure decision-making process and outcomes.

The evidence from the Cochrane review for effectiveness is strong and from a large body of trials. In contrast, the use of measurement in the trials is variable, and the quality of reporting of the instruments and measures used could be much improved. The further

analysis of more recently published trials demonstrates continued shortcomings in this reporting.

This overview raises issues to be addressed in future conceptual and empirical investigations. Key questions include, what are the best measures of decision quality? How much, and what type of, patient knowledge is needed to evaluate or support high quality decisions? What is the impact of PtDAs on health inequalities and health literacy? Can we develop a minimum (core) dataset of measures of decision quality and PtDA effectiveness? What is the contribution of the patient-provider relationship to PtDA effectiveness? What are the (most) active ingredients of PtDAs? Which components are core to effectiveness? How can we better measure the cost-effectiveness of PtDAs? Do we have evidence that the decision-making process variables are predictive of decision quality? A full list of questions is included in Section 6.

Suggested Changes to IPDAS Criteria.

Since the publication of the last chapter, an international consensus process has led to robust guidance on reporting of PtDA evaluation that supports this quality dimension. Those reviewing the effectiveness of specific PtDAs should also refer to the Standards for UNiversal reporting of patient Decision Aid Evaluation (SUNDAE) (Sepucha et al 2017; Hoffman et al 2018) in order to appraise the quality of the evaluation undertaken.

In order to improve the IPDAS checklist, we recommend that an additional criterion be added:

"IV. Evaluation" - "Was any evaluation of the PtDA conducted and reported with attention to SUNDAE guidelines."

SECTION 3: DEFINITION (CONCEPTUAL/OPERATIONAL) OF THIS QUALITY DIMENSION

a) Updated Definition

To establish the effectiveness of a patient decision aid (PtDA), it is critical to provide evidence that the PtDA improves: a) the quality of the decision-making process; and b) the quality of the decision choice (or decision quality), i.e. the quality of the choice that is made. Further, to promote high quality evidence, it is important that evaluation studies of PtDAs are conducted and reported with attention to the SUNDAE guidelines (Sepucha et al 2017; Hoffman et al 2018).

There is broad recognition that appropriate outcomes of a good decision are separate from traditional clinical/health outcomes (although it may be important that these are also measured). The *quality of the decision-making process* refers to the extent to which a

PtDA helps patients to: recognize that a decision needs to be made; feel informed about the options and their characteristics; be clear about what matters most to them in this decision; discuss goals, concerns, and preferences with their health care providers; and be involved in decision making. *The quality of the decision choice that is made* (often called "decision quality") is defined as the extent to which a patient's eventual choice is consistent with their informed values and is implemented.

The Quality of the Decision-Making Process

Core areas/domains about the decision-making process that should be captured, and some details about how each area/domain might be measured, include the extent to which PtDAs help patients to:

- Recognize that a decision requiring their input needs to be made (e.g., one approach is to assess whether the patient understands that there is more than one reasonable approach to the situation and that her/his participation is important, e.g. as measured by items in the Preparation for Decision Making Scale (PMDS)) (Bennett et al., 2010)

- Feel informed about the options and the risks, benefits, and consequences of the options (e.g., the uninformed subscale of the Decisional Conflict Scale (DCS)) (O'Connor, 1995)

- Be able to identify what matters to them in this decision (e.g., the unclear values subscale of the DCS) (O'Connor, 1995)

- Feel able to state a clear preference.

- Discuss goals, concerns, and preferences with their health care providers (e.g., as measured by items in the Perceived Involvement in Care Scale (PICS)) (Lerman et al., 1990)

- Be involved in decision making (e.g., adaptation of Degner's Control Preferences Scale (CPS) that assesses who made the decision (Degner, 1992)).

The Quality of the Decision Choice That is Made (Decision Quality)

The quality of the choice that is made following use of a PtDA (often referred to as "decision quality") was defined in the prior round of IPDAS voting as the extent to which a patient's decision is consistent with their informed values (Elwyn 2006). Other definitions of decision quality emphasize the additional need for the choice to be implemented (O'Connor 1997, Sepucha 2004). It follows from the definition that the impact of PtDAs on the following outcomes should be measured whenever possible:

- *Informed patient*:
  This is most commonly measured by assessing a patient's knowledge of the options and outcomes. Note that this is not about a patient's perceptions of their knowledge; instead, knowledge questions are used to assess their understanding of necessary and important information (objective knowledge). This may include an assessment of accurate risk/benefit perceptions and realistic expectations, when applicable.

- *Concordance between a patient's values and goals (what matters to the patient) and the chosen option*:
  Despite considerable variation in assessment of this construct, most approaches require the following: (1) elicitation of the patient's values and/or treatment preferences; (2) identification of the patient's chosen or implemented option; and (3) calculation of the extent to which the chosen option aligns with the stated values or preferences.

Other Measures

Many other process and outcome variables have been used to evaluate PtDAs. For example, these include decision self-efficacy, decision regret, and patient satisfaction with decision making and choice of option. Furthermore, many surveys and scales cover some, but not all, of the process and outcome domains presented above. The definition presented here focuses on the core set of domains and some examples of scales that have been used; it does not imply that these are the only aspects that should or could be measured. For example, we may be interested in assessment of biases (e.g., sunk cost bias, focusing effects, optimism bias) that an effective PtDA could help minimize.

It is also important to assess the cost-effectiveness of PtDAs, as well as other impacts such as harms or unintended consequences, even though these are not direct measures of effectiveness per se. Effectiveness of a PtDA is measured centrally by improvements in the decision-making process and in the quality of choice that is made.

b) Changes from the Original Definition and Suggestions for Review of the IPDAS Criteria

The core of the original definition has been retained and the content of this section updated. Further, the definition still reflects the original IPDAS criteria that similarly focused on decision-making process and decision quality.

In order to improve the IPDAS checklist, we recommend that an additional criterion be added:

"IV. Evaluation" - "Was any evaluation of the PtDA conducted and reported with attention to SUNDAE guidelines"

c) Emerging Issues/Research Areas in Definition

See Section 6.

SECTION 4: THEORETICAL RATIONALE FOR INCLUSION OF THIS QUALITY DIMENSION

a) Updated Theoretical Rationale

Scientific Rationale

Establishing the effectiveness of any health care intervention, including PtDAs, is critical. There is a strong consensus and scientific evidence that PtDAs: a) improve the quality of the decision-making process; and b) increase the likelihood that individuals choose and/or receive health care interventions that are most consistent with their informed and considered values (Briss et al., 2004; Elwyn et al., 2006; O'Connor et al., 1997; Ratliff et al., 1999; Stacey et al., 2003; Stacey et al., 2011; Sepucha et al., 2004; Sepucha 2008). Further, the Ottawa Decision Support Framework, one of the more common frameworks used to guide decision aid development, has recently been updated and includes informed, values-based decisions as a key outcome (Stacey 2020). Arguably, the best study design, when applicable, for determining efficacy of health care interventions is the double-blind randomized controlled trial (RCT) that attempts to minimize risk of bias between groups to allow for head-to-head comparisons. This is the inclusion criterion for the Cochrane review whose results we draw upon in this chapter.

Ethical Rationale

Whilst this chapter focuses on effectiveness, PtDAs are commonly seen as a means of supporting a shift from paternalism to informed and shared decision making (SDM). SDM is a process by which a decision is made between the patient, with or without his/her family and/or significant others, and one or more healthcare professionals. It offers a model to approach patient engagement, particularly in preference-sensitive decisions – decisions in which there is more than one reasonable option and in which the choice among those options is likely to be influenced by patient preferences and values.

Furthermore, not only is there a very clear ethical imperative to engage patients in decisions about their own care, but also PtDAs can support better informed consent. King

et al (2006) have argued that traditional informed consent methods are inadequate to engage and inform patients about treatment options in preference-sensitive decisions, and that SDM, because it extends beyond information-giving to supporting the formulation and communication of informed preferences, may offer an ethically and legally supported means for fostering informed choice. Research shows that, when usual care is compared to use of PtDA, current approaches without PtDAs are inadequate for ensuring that patients are informed and have realistic expectations (Stacey et al., 2017).

As in any area of the evaluation of interventions, we should be concerned with addressing not only the benefits but also any adverse effects. In the context of PtDAs, the literature is relatively light on the exploration of adverse effects, although such effects have been posited. Adverse effects might include, for example, an increase in inequalities (through being more accessible to, or used by, well-educated patients) or selection of "less-effective" interventions with impact on population health, particularly in preventive health choices (Thomson, 2005). For example, in the past well informed patients may have selected aspirin rather than (the more effective) warfarin for prevention of stroke in atrial fibrillation as a result of balancing how they value the benefits and risks of these alternatives (Thomson, 2007). Recent guidance and availability of novel oral anticoagulants has changed this decision (NICE, 2004). Other adverse effects might include increased patient anxiety when patients are faced with clinical uncertainty (Politi et al., 2011), or are offered an unexpected role in decision making and are initially wary of engaging in decision making (Say et al., 2006). Patients may feel unsupported or "abandoned" if decision making is not actually shared but is unduly delegated to patients (Quill & Cassel, 1995). Adverse effects may also occur if PtDAs are neither well developed nor kept up to date, or if developed by those with significant conflict of interest, which may therefore bias decisions. Therefore, standards such as these are important.

Conceptual Rationale

Measures of the quality of the decision process and outcomes highlighted in this chapter have underpinnings in theories of decision making. Normative theories of decision making, such as subjective expected utility theory, are based on the ideal that all patients can approach decisions rationally and are able to weigh the risks and benefits of various interventions (von Neumann and Morgenstern, 1947). Descriptive theories of decision making, such as prospect theory, demonstrate that humans are subject to cognitive biases that cause decision making to deviate from the normative/rational (Kahneman and Tversky, 1979). For example, a well-known cognitive bias has to do with the effects of framing, where people tend to be risk-averse when statistics are presented as gains and risk-seeking when they are presented as losses (Tversky and Kahneman, 1974). These kinds of biases can threaten a person's ability to acquire accurate knowledge or to make a decision concordant with their values, thus threatening the quality of their decision. The Dual-Process Theory of decision making (Kahneman, 2003) argues that people make decisions either 'intuitively' (i.e., quickly drawing on past experiences), or 'reasonably' (i.e., using a thoughtful, analytic approach), with the latter being less subject to many of the cognitive biases. PtDAs are designed to encourage a more thoughtful or reasoned process in decision making, thus minimizing the influence of cognitive bias. If a more reasoned, normative approach is pursued, the actual choice will more likely to be informed and value concordant – the key domains of "decision quality".

Although conceptually many of these theories share similar underpinnings (e.g., an emphasis on information, and the use of deliberative processes to align choices with goals), there is considerable debate about how that gets translated into specific measures (see section 5). Debate involves not only how to measure the construct (e.g., measuring patients' preferences using the standard gamble in formal decision analysis versus using attitude scales), but also the timing of the measurement. For example, some have argued that it is only legitimate to measure the quality of decision making before and immediately after exposure to the PtDA, because later measurements may be prone to potential biases . This is especially of concern after treatments have been experienced because they may be influenced by hindsight bias based on health outcomes (Elywn and Miron-Shatz, 2010).

A key goal of health care is to improve health outcomes. However, many outside the field question the impact of PtDAs on health outcomes. There are several challenges to the use of clinical health outcomes (such as quality of life, pain, or mortality) to assess the effectiveness of PtDAs (e.g. Rutherford et al., 2019). First, the nature of the situations addressed by PtDAs requires that there are multiple reasonable options; thus, by definition, there is usually not one clearly superior treatment or intervention. Second, many of these decisions are made under uncertainty. Patients are essentially making a bet; thus, evaluation of the bet depends on the odds not the outcome. For example, a patient may choose to have surgery, feeling that the benefits outweigh the harms, and yet may suffer a severe, unanticipated complication during the procedure. Using health

outcomes to evaluate that decision would suggest it was poor, whereas using decision-making process or decision quality measures would suggest it was strong. A third challenge of using health outcomes as a measure is that studies have shown that patients vary in their willingness to trade off quality of life and quantity of life. For example, patients may elect to forego chemotherapy if their desire to avoid short-term severe side effects outweighs their desire for increasing short-term survival. To the extent that patients feel differently about the potential health outcomes, it is necessary to measure the effectiveness of PtDAs by the extent to which they enable patients to achieve the outcomes they most desire (and avoid those they most dislike).

Policy Rationale

There are several recent, widespread health policy drivers that align with the implementation of PtDAs and SDM, hence emphasizing the need for a robust evidence base. This is an international movement that is driving changes across the world. Further description on policy developments internationally can be found in a recent international review of different countries' progress undertaken by Angela Coulter (Coulter, 2018)

Policy drivers internationally include influential reports, health service policies, legislation and regulation, professional bodies, research organizations and patient organizations.

Examples of influential reports include the inclusion of SDM (often facilitated by PtDAs) as one of top five priority areas for measurement gaps in the US National Quality Forum (NQF) identified (National Quality Forum, 2011. Table 2).

Examples of national health policies include inclusion of SDM ("Nothing about me, without me") in UK Government health policy (Department of Health, 2010); the fourth Danish national cancer plan (Kræftplan IV; "The patients' cancer plan" 2016) including an initiative that "The patient will be part of the decision-making process" (http://cphpost.dk/news/new-cancer-plan-put-in-motion.html); and in 2015 in Taiwan, the Ministry of Health and Welfare launch of a programme to promote SDM across Taiwan under the leadership of the Joint Commission of Taiwan (JCT) (see Coulter, 2018).

Examples of legislation and regulation include Washington State, USA passing legislation in 2007 promoting SDM and the use of PtDAs (King, 2017); recent medical negligence case law in the UK Supreme Court (Montgomery v Lanarkshire. UK Supreme Court, 2017) that has embedded the general and particular patient perspective into processes of consent, effectively requiring upstream SDM prior to consent; the Australian Commission on Safety and Quality in Healthcare National Safety and Quality Healthcare Standards (which underpin their accreditation process) have 'Partnering with Consumers' as a key standard that requires health care organizations to partner with patients to share decisions and care planning as one key strategy (ACSQHC, 2017)); in

Germany, the Law of Patient Rights (§§ 630c-e BGB/German Civil Code) supports the promotion of adequate information and participation in decision-making.

In Canada, there is a long history of academic research. For example, the Ottawa Hospital Research Institute web site has hosted the IPDAS content and a database of PtDAs for many years, and has a significant international profile and impact (OHRI, https://decisionaid.ohri.ca). Ottawa also holds the lead for the Cochrane review of PtDAs.

A different pressure for measurement comes from the policy perspective —costs. If health systems are to fund access to PtDAs, then they want to know the intervention is not only effective but also cost-effective. The impact of PtDAs on utilization has been demonstrated in a small number of decisions, but, as of yet, there is not strong evidence that PtDAs are cost-effective or reduce costs. Their impact on cost seems to depend on baseline utilisation (NICE, 2011). As implementation efforts expand, examining the impact on costs/cost-effectiveness will be an important outcome to include in those efforts (Treneman et al.,2014; Butt, 2019;Walsh et al., 2014).

b) Changes from the Original Theoretical Rationale

The previous section has been updated, in particular to capture examples of the growing policy initiatives.

c) Emerging Issues/Research Areas in Theoretical Rationale

See Section 6.

SECTION 5:
EVIDENCE BASE UNDERLYING THIS QUALITY DIMENSION

a) Updated Evidence

Cochrane Collaboration Review

The 2017 update of the Cochrane Collaboration's Systematic Review includes 105 RCTs comparing individual PtDAs for treatment or screening decisions to usual care, incorporating 18 new studies and removing 28 previously included studies that had an active comparator. The trials were identified by searching through to April 2015 on CENTRAL, MEDLINE, EMBASE, PsycINFO, grey literature and Cochrane databases, and through to September 2008 for CINAHL. Seventy-one trials measured knowledge (67.6%), though 19 were not able to be included in the pooled analyses due to the way their data were presented. Twenty-five (23.8%) measured accurate risk perceptions (8 of which were unable to be pooled) and 16 (15.3%) measured agreement between values

and choice (6 of which were unable to be pooled). Sixty-three studies (60%) used the Decisional Conflict Scale, 23 studies reported on measured feeling clear about values, and 27 measured feeling informed (Stacey et al., 2017).

Evidence about the Quality of the Decision-Making Process
PtDAs resulted in:

a) reduction in feeling uninformed (MD -9.28 of 100; 95% CI -12.20 to -6.36, n=27);
b) reduction in feeling unclear about personal values (MD -8.81; 95% CI -11.99 to -5.63, n=23); and
c) fewer people passive in decision making (RR 0.68; 95% CI 0.55 to 0.83, n=16).

No studies evaluated decision-making process attributes relating to helping patients to recognize that a decision needs to be made or understand that their values affect the choice (Stacey et al. 2017).

Evidence about Decision Quality - The Quality of the Choice that is Made
The latest results indicate that: PtDAs improve knowledge by 13% (mean difference 13.27 out of 100, 95% CI 11.32-15.23, n=52), and improve accurate risk perception (relative risk 2.10; 95% CI 1.66-2.66, n=17). PtDAs also improve agreement between informed values and choice (RR 2.06; 95% CI 1.46-2.91, n=10). There was no difference in values-choice congruence in studies that did not use informed values (n=6).

Other Outcomes
PtDAs also appear to have a positive effect on patient-practitioner communication (n=10) and a positive (n=4) or neutral (n=15) effect on satisfaction. PtDAs do not appear to be different from usual care in terms of anxiety (n = 20), and general health outcomes (n = 7) or condition-specific health outcomes (n = 9). The effects of PtDAs on adherence (16 studies) and costs/resource use were inconclusive (Stacey et al., 2017). Overall, decisions aids extended the length of the visit by a median of 2.6 minutes (24 vs 21; 7.5% increase, n=10). Studies did not report adverse events associated with use of decision aids.

Review of the Reported Performance of Instruments/Measures

There are many different survey instruments and measures used to assess decision-making process and decision quality. It is important to clearly document and appraise the performance of the instruments used to evaluate the effectiveness of PtDAs. A few recent studies exist that examine the performance of instruments such as:

- Kryworuchko et al.'s 2008 appraisal of selected "primary outcome" measures used in decision aid trials, which included some decision-making process and decision quality measures. The study concluded that most publications provided inadequate detail for appraising the included instruments.

- Sepucha and Ozanne's 2009 systematic review of methods to calculate value concordance, which found considerable variability in definitions and methodology and minimal reports of validation of instruments or approaches.
- Scholl et al.'s 2011 review of SDM instruments (an update of Simon 2007), which showed that, whilst reliability of most scales is good, they differ in their extent of validation.
- Gartner et al.'s 2018 review of SDM instruments, which included 40 instruments and, as in prior reviews, found limited evidence on psychometrics.

In the previous version of this chapter, we identified all reported outcomes from the 86 studies in the 2011 Cochrane Collaboration's review and determined whether they measured one or more of the elements of the "quality of the decision-making process" or "decision quality" as defined earlier. We extracted data on the reported development and performance of each instrument, focusing on the reporting of acknowledged key criteria for high-performing patient-reported measures including reliability, validity, responsiveness, precision, interpretability, feasibility, and acceptability. The abstraction was limited to the details provided within the published trial papers – based on how a reader might evaluate the measures as described by the trial authors.

Previously, we found that most studies included measurement of at least one aspect of decision quality or decision-making process (80/86 trials); yet few provided details on psychometric properties of the individual measures (Sepucha et al 2014). This work informed the subsequent development and publication of reporting guidelines for evaluations of patient decision aids (Sepucha et al 2017).

As part of this update, we abstracted outcomes from 49 new studies included in the 2014 and 2017 Cochrane reviews. Most studies (44/49) included one or more measure of decision quality or process. We abstracted 109 measures from the 44 studies that covered knowledge (n=48), concordance (n=12), and decision process (n=55) (of note: 6 measures covered both knowledge and concordance). Studies included limited information on psychometric properties of the measures such as reliability (23% of measures), validity (6% of measures), responsiveness (1% of measures) and precision (0%). Studies also did not include much on clinical sensibility of the measures such as feasibility (1%), acceptability (6.5%) and interpretability (5%). The details of this review of outcomes is published in Trenaman et al (2021).

Despite the general lack of reporting on the performance of instruments, there are a few examples of high quality reporting. One instrument that had some of the best reporting on its performance was the Decision Conflict Scale (DCS) (O'Connor, 1995). It is also one of the most widely-used measures (in 72 of the 135 trials) (Garvelink et al 2019a and 2019b). The DCS measures two of the decision-making process domains—whether patients feel that they understand testing/treatment options and outcomes, and whether they feel clear about what matters most to them.

Tables 1 and **2** illustrate notable examples of how investigators included fairly comprehensive but parsimonious reporting on the performance of instruments.

Table 1 Reports on the Decisional Conflict Scale.

TABLE 1: Extract from Vodermaier A et al., 2009.

"Measures
Patients were assessed with the following instrumentation:
Primary outcome variable
Decisional conflict. The Decisional Conflict Scale (DCS; O'Connor, 1995) measures patients' uncertainty about which treatment to choose, factors contributing to uncertainty (believing to be uninformed, unclear values, and unsupported in decision making), and perceived effectiveness of decision making. Questions have to be answered on a 5-point Likert scale [*from strongly agree to strongly disagree*]. Higher scores on the scale or subscales reflect higher decisional conflict, uncertainty, and a less effective choice. The German version of the scale demonstrated subscale and total score internal consistencies in the present sample between 0.73 and 0.94. The scale discriminates between patients who make and those who delay decisions (O'Connor, 1995; Bunn and O'Connor, 1996) and is sensitive to change (O'Connor et al., 1998)."

Table 2 Reports on a knowledge measure that was developed for the featured study.

TABLE 2. Extracts from Barry et al., 1997.

"The quality of the treatment decisions was examined using a multifaceted approach. First, we determined whether subjects were better informed through a twenty question test of BPH Knowledge … developed by a panel including a general internist, a urologist, a survey researcher, and a lawyer with a special interest in informed consent. Correct responses were scored +1, incorrect responses -1, and "not sure" responses were scored 0 (total range -20 to +20). The test of knowledge was administered two weeks after exposure to either the [patient decision aid]) or the control brochure."

and

"Validation of New Outcome Measures
Cronbach's alpha statistic for the items testing BPH knowledge was 0.68. The criterion validity of this test was assessed by comparing scores for a convenience sample of 12 urologic nurses with the scores of the 167 BPH patients enrolled in the baseline period. The nurses had a mean score of 14.8 [out of 20], compared to 5.6 for the patients ($p <$

0.001). Nurses answered an average of 85% of the questions correctly, compared to 48% for the patients ($p < 0.001$). Furthermore, a modest correlation between these patients' knowledge scores and their educational levels was seen, r = 0.23 ($p < 0.001$)."

Other work that has advanced the measurement of decision quality is not yet reflected in the Cochrane Collaboration's Review. Sepucha and colleagues have published psychometric analyses of three decision quality instruments (for osteoarthritis of the knee or hip, herniated disc, and breast cancer surgery) that assess the extent to which patients are informed and receive treatments that match their goals (Sepucha, Stacey et al., 2011; Sepucha, Feibelmann et al., 2012; Sepucha, Belkora, et al., 2012). In general, these instruments met several criteria for patient-reported outcomes, including test-retest reliability, content and discriminant validity, acceptability, and feasibility. It will be important to further evaluate these as outcome measures in trials of decision aids to achieve the outcome of decision quality as defined in this chapter and in previous IPDAS publications (Elwyn et al., 2006).

Limitations

Given that we are focused on papers included within the Cochrane review, we recognise that there are some limitations to our conclusions, some of which will be addressed within the implementation chapter. The RCT design attempts to evenly distribute factors other than the intervention between groups to allow for head-to-head comparisons and minimize bias. However, the limitations of RCTs and their inability to capture appropriate information for emergent and complex interventions are increasingly recognised (Glasgow, 2012). RCTs by design are reductionist, and so do not capture data about patient values and complexities (Greenhalgh et al., 2016) nor do they necessarily capture important sub-group differences, unless considered a priori. Whilst most RCTs have some design flaws, with RCTs of PtDAs it is also challenging (or impossible) to blind participants (patients or providers) to the intervention. Perhaps a greater issue is that even an *ideal* RCT translates poorly into practice because the ideal circumstances, limited participant base, and the heavily supported implementation of the efficacy trial rarely mirror usual service. Nor do they capture the necessary information needed for implementation into a variety of contexts with diverse client groups (Curran et al., 2012). RCTs of PtDAs can also be limited because they tend to include a fairly homogeneous sample of well-educated, white, middle income patients, and less is known about the impact PtDAs have on disadvantaged patients (e.g., individuals with low literacy, or low incomes). In short, many RCTs do not help answer the realist questions about which treatments work for which individuals in which contexts. Some of these issues are being addressed in other chapters, for example on health literacy and implementation.

b) Changes to Original Evidence

The above summary of evidence has been updated, incorporating: a) the results of the 2014 and 2017 Cochrane Collaboration's reviews of the 105 RCTs of PtDAs; and b) an examination of the reported performance of instruments used to measure primary process and outcome variables in 49 new studies included in the 2014 and 2017 Cochrane Collaboration's review (updating the previous analysis).

c) Emerging Issues/Research Areas in Evidence

See Section 6.

SECTION 6: EMERGING ISSUES AND RESEARCH AREAS IN THIS QUALITY DIMENSION

Inevitably, the revision of this chapter has also identified several questions that remain to be fully answered or settled. A large number emerged in our discussions, which we have reviewed and summarized as high-level questions that we suggest are most worthy of pursuit in future research.

Measuring the Decision-Making Process and Decision Quality

- Current measures are almost exclusively patient-reported measures; what advantages/disadvantages would be involved in using provider-reported measures or measures of the patient-provider interaction?
- What are the best measures of decision quality?
- How much, and what type of patient knowledge, is needed to evaluate or support high quality decisions?
- What is the value of decision quality measurement at an individual and aggregate level?
- What is the difference between risk perception and risk understanding?
- What methods are most appropriate for determining value concordance? Is it best to focus on the chosen option or the implemented option?

What Other Outcomes Should be Used to Evaluate Patient Decision Aids?

- What is the role, if any, for including measurement of health outcomes?
- How can we better measure the cost-effectiveness of PtDAs? What outcomes should be used to measure their cost-effectiveness?

- What is the value (or limitation) of measuring adherence to chosen treatment option (e.g. medication adherence)?
- What is the impact of PtDAs on treatment rates, adherence, service utilization, and service costs?
- How do we measure impact of PtDAs on patient-provider interaction?
- What is the impact of PtDAs on health inequalities and health literacy?
- Should harms/adverse effects of PtDAs be explicitly measured and, if so, how? (e.g., bias, cognitive burden, decision regret)

Designing Evaluation Studies of Decision Aids

- What are the best research designs for measuring effectiveness of PtDAs? What is the role of effectiveness-implementation hybrid designs?
- Can we develop a minimum (core) dataset of measures of decision quality and PtDA effectiveness?

Contextual Considerations

- What is the contribution of the patient-provider relationship to PtDA effectiveness?
- How does the affective status of patients impact upon the effectiveness of PtDAs?
- How (and through which methods) can PtDAs best support values elicitation?
- Do the benefits (and potential harms) of PtDAs differ between countries and cultures?

Decision Aid Design and the Impact on Effectiveness

- What is the role of brief PtDAs compared to complex PtDAs?
- What are the (most) active ingredients of PtDAs? Which components are core to effectiveness?

When Should we Measure Impact?

- Measuring knowledge or patients' preferences too long after the decision has been made may lead to hindsight bias in the appraisal of that decision.
- Measuring immediately after a PtDA but before the decision has been made is also problematic as provider consults will likely influence both knowledge and patients' preferences.

Theoretical Issues

- Multiple theories are relevant to the development of PtDAs and improving decision quality (decision making theories, information processing theories, communication theories, etc.) – and each may suggest different outcomes. How to reconcile and/or link to each?

- A growing body of research suggest that defining what a good medical decision is, and how to measure it is more complicated than what is often assumed in theoretical decision making frameworks (Hamilton 2017). Real life decision making is influenced by e.g. interpersonal factors, structural constraints and affect/emotions. It provides an argument for consideration of how these factors (and potentially others) contribute to the definition of good medical decision making and a tailored approach to the measurement of decision quality.
- Increasingly, one or more options in situations covered in PtDAs involve a large behavior change component (e.g., surgery versus diet and exercise for obesity/weight management). In what ways does this behavior change component change our strategies (if at all) for the evaluation of PtDAs (e.g., do we need to assess levels of self-efficacy and motivation in addition to knowledge and concordance)?
- Do we have evidence that the decision-making process variables are predictive of decision quality?

REFERENCES

Australian Commission on Safety and Quality in Healthcare, 2017 *National Safety and Quality Healthcare Standards*, 2nd Edition ACSQHC online https://www.safetyandquality.gov.au/standards/nsqhs-standards/partnering-consumers-standard (Accessed 17 Jan 2020)

Barry MJ, Cherkin DC, et al. A randomized trial of a multimedia shared decision-making program for men facing a treatment decision for benign prostatic hyperplasia. Disease Management and Clinical Outcomes. 1997; **1**(1): 5-14.

Bennett C, Graham I, Kristjansson E, Kearing S, Clay KF, O'Connor A: Validation of a preparation for decision making scale. Patient Educ Couns 2010, 78:130-133

Briss P, Rimer B, Reilley B, Coates R, Lee N, Mullen P, et al. Promoting informed decisions about cancer screening in communities and healthcare systems. Am J Prev Med. 2004; 26:67-80.

Butt T. Measuring the Benefits of Decision Aids for Economic Evaluation. Pharmacoecon Open. 2019 Jun;3(2):143-150. doi: 10.1007/s41669-018-0087-y.

Coulter A. 2018. National Strategies for Implementing Shared Decision Making. Bertelsmann Stiftung

Curran, G.M., M. Bauer, B. Mittman, et al., *Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact.* Med Care, 2012 **50**(3):217-226

Degner LF, Sloan JA. Decision making during serious illness: What role do patients really want to play? J Clin Epidemiol. 1992; 45: 941-950.

Department of Health. Equity and excellence: liberating the NHS . London: 2010. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_117353 (accessed 17 Jan 2020)

Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, Thomson R, Barratt A, Barry M, Bernstein S, Clarke A, Entwistle V, Feldman-Stewart D, Holmes-Rovner M, Llewellyn-Thomas H, Moumjid N, Mulley A, Ruland C, Sepucha K, Whelan T. Developing a quality criteria framework for patient decision aids: an online international Delphi consensus process. BMJ, Aug 2006; 333: 417 ; doi:10.1136/bmj.38926.629329.

Garvelink MM , Boland L, Klein K, Vu Nguyen D, Menear M, Bekker HL, Eden KB, LeBlanc A, O'Connor AM, Stacey D, Légaré F. Decisional Conflict Scale Use over 20

Years: The Anniversary Review. Med Decis Making. 2019a DOI: 10.1177/0272989X19851345

Garvelink MM , Boland L, Klein K, Vu Nguyen D, Menear M, Bekker HL, Eden KB, LeBlanc A, O'Connor AM, Stacey D, Légaré F. Decisional Conflict Scale Findings among Patients and Surrogates Making Health Decisions: Part II of an Anniversary Review. Med Decis Making. 2019b DOI: 10.1177/0272989X19851346

German Journal for Quality in Health Care - Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen. Policy and practice developments in the implementation of shared decision making: An international perspective. Edited by Martin Harter, Trudy van der Weijden and Glyn Elwyn. 2011 Volume 105, Issue 4, Pages 227-326.

German Law of Patient Rights (§§ 630c-e BGB/German Civil Code) https://www.gesetze-im-internet.de/englisch_bgb/englisch_bgb.html#p2645 (Accessed 17 Jan 2020)

Glasgow, R.E. and D. Chambers, *Developing Robust, Sustainable, Implementation Systems Using Rigorous, Rapid and Relevant Science.* Clin Transl Sci, 2012 **5**(1):48-55

Greenhalgh, T., Annandale, E., Ashcroft, R., Barlow, J., Black, N., Bleakley, A. Ziebland, S. (2016). An open letter to the BMJ editors on qualitative research. *BMJ, 352*, i563. doi:10.1136/bmj.i563

Hamilton JG, Lillie SE, Alden DL, Scherer L, Oser M, Tanaka M, Baleix J, Brewster M, Craddock Lee S, Goldstein MK, Jacobson RM, Myers RE, Zikmund-Fisher BJ, Waters EA. What is a good medical decision? A research agenda guided by perspectives from multiple stakeholders. J Behav Med. 2017 Feb;40(1):52-68.

Hoffman AS, Sepucha KR, Abhyankar P, Sheridan S, Bekker HL, LeBlanc A, Levin C, Ropka M, Shaffer V, Stacey D, Stalmeier P, Vo H, Wills C, Thomson R. Explanation and elaboration of the Standards for Universal reporting of patient Decision Aid Evaluations (SUNDAE) guidelines: examples of reporting SUNDAE items from patient decision aid evaluation literature BMJ Qual Saf 2018;**0**:1–24. doi:10.1136/bmjqs-2017-006985

Institute of Medicine. Crossing the quality chasm: A New Health System for the 21st Century. IOM, 2001.

Kahneman D. A perspective on judgement and choice. American Psychologist. 2003; 58, 697-720.

Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk, Econometrica. 1979; XLVII, 263-291.

King JS, Eckman MH, et al. (2011). The potential of shared decision making to reduce health disparities. The Journal of Law, Medicine & Ethics. 2011; 39: 30-33.

King J, Moulton B. Group Health's participation in a shared decision-making demonstration yielded lessons, such as role of culture change. Health Aff (Millwood). 2013 Feb;32(2):294-302. doi: 10.1377/hlthaff.2012.1067. PubMed PMID: 23381522.

Kræftplan IV; "The patients' cancer plan" 2016

Kryworuchko, J., D. Stacey, et al. Appraisal of primary outcome measures used in trials of patient decision support. Patient Education and Counseling. 2008; 73(3): 497-503.

Lerman C, Brody D, Caputo G, Smith D, Lazaro C, Wolfson H: Patients' perceived involvement in care scale: relationship to attitudes about illness and medical care. J Gen Intern Med 1990, 5:29-33.

NICE (National Institute for Health and Clinical Excellence) (2011). Patient experience in adult NHS services quality standard. https://www.nice.org.uk/guidance/qs15 (accessed 17 Jan 2020)

NICE (National Institute for Health and Clinical Excellence) 2014 Atrial fibrillation: management. Clinical guideline [CG180]

National Priorities Partnership. 2010 Input to the Secretary of Health and Human Services on Priorities for the 2011 National Quality Strategy

National Quality Forum. Measure Prioritization Advisory Committee Report: Measure development and endorsement agenda. 2011 http://www.qualityforum.org/Publications/2011/01/Measure_Prioritization_Advisory_Committee_Report.aspx (Accessed 28th May 2020)

O'Connor AM. Validation of a decision conflict scale. Med Decis Making. 1995; 15: 25–30.

O'Connor A, Llewellyn-Thomas H, Sawka C, Pinfold S, To T, Harrison D. Physicians' opinions about decision aids for patients considering systemic adjuvant therapy for axillary-node negative breast cancer. Patient Education & Counseling. 1997;30:143-53.

Ottawa Health Research Institute. Ref Web site https://decisionaid.ohri.ca (Accessed 17 Jan 2020)

Politi MC et al. Communicating uncertainty can lead to less decision satisfaction: a necessary cost of involving patients in shared decision making? Health Expectations. 2001; 14(1), 84-91.

Quill T, Cassel C. Non-abandonment: a central obligation for physicians. Annals of Internal Medicine. 1995; 122, 368-74.

Ratliff A, Angell M, Dow R, Kupperman M, Nease R, Fisher R, et al. What is a good decision? Effective Clinical Practice. 1999;2:185-97.

Rutherford C, King MT, Butow P, Legare F, Lyddiatt A, Souli I, Rincones O, Stacey D. Is quality of life a suitable measure of patient decision aid effectiveness? Sub-analysis of a Cochrane systematic review. Qual Life Res. 2019 Mar;28(3):593-607

Say R, Murtagh M, Thomson, R. Patients' preference for involvement in medical decision making: a narrative review. Patient Education and Counseling. 2006; 60(2), p. 102.

Scholl, I., M. K. Loon, et al. Measurement of shared decision-making-a review of instruments. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2011;105: 313-34.

Sepucha KR, Belkora JK, Chang Y, Cosenza C, Levin CA, Moy B, Partridge A, Lee CN. Measuring decision quality: A psychometric evaluation of new instrument for breast cancer surgery. BMC Med Inform Decis Mak. 2012 Jun 8;12(1):51.[Epub ahead of print]

Sepucha KR, Feibelmann S, Abdu WA, Clay CF, Cosenza C, Kearing S, Levin CA, Atlas SJ. Psychometric evaluation of a decision quality instrument for treatment of lumbar herniated disc. Spine. 2012; Mar 15. [Epub ahead of print].

Sepucha, KR, Fowler, FJ, Mulley AG. Policy support for patient-centered care: The need for measurable improvements in decision quality. Health Affairs 2004. https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.var.54 (Accessed 28th May 2020)

Sepucha K, Ozanne E, Silvia K, Partridge A, Mulley A. An approach to measuring the quality of breast cancer decisions. Patient Education and Counseling. 2007; 65(2):261-9.

Sepucha K, Ozanne E. How to define and measure concordance between patients' preferences and medical treatments: A systematic review of approaches and recommendations for standardization. Patient Education and Counseling. 2010; Jan;78(1):12-23.

Sepucha K, Levin C, Uzogara E, Barry M, O'Connor A, Mulley A. Developing instruments to measure the quality of decisions: Early results for a set of symptom-driven decisions. Patient Education and Counseling. 2008; 73:504-510.

Sepucha K, Stacey D, Clay C, Chang Y, Cosenza C, Dervin G, Dorrwachter J, Feibelmann S, Katz JN, Kearing S, Malchau H, Taljaard M, Tomek I, Tugwell P, Levin

C. Decision quality instrument for treatment of hip and knee osteoarthritis: a psychometric evaluation. BMC Musculoskelet Disord. 2011; Jul 5;12(1):149.

Sepucha KR, Borkhoff CM, Lally J, Levin CA, Matlock DD, Ng C, Ropka ME, Stacey D, Joseph-Williams N, Wills CE, Thomson R. Establishing the effectiveness of patient decision aids: key constructs and measurement instruments BMC Medical Informatics and Decision Making 2013, 13(Suppl 2):S12 doi:10.1186/1472-6947-13-S2-S12

Sepucha KR, Matlock DD, Wills CE, Ropka M, Joseph-Williams N, Stacey D, Ng C, Levin C, Lally J, Borkhoff CM, Thomson R. "It's Valid and Reliable" Is Not Enough: Critical Appraisal of Reporting of Measures in Trials Evaluating Patient Decision Aids. Med Decis Making. 2014 Jul;34(5):560-6. doi: 10.1177/0272989X14528381. Epub 2014 Apr 8.

Sepucha KR, Abhyankar P, Hoffman AS, Bekker HL, LeBlanc A, Levin CA, Ropka M, Shaffer VA, Sheridan SL, Stacey D, Stalmeier P, Vo H, Wills CE, Thomson R. Standards for UNiversal reporting of patient Decision Aid Evaluation studies: the development of SUNDAE Checklist BMJ Qual Saf 2017;**0**:1–9. doi:10.1136/bmjqs-2017-006986

Simon D, Loh A, Härter M. Measuring (shared) decision making -- a review of psychometric instruments. Zärtzl. Fortbild. Qual. Gesundh.wes . 2007; 101; 259-267.

Stacey D, Légaré F, Boland L, et al. 20th Anniversary Ottawa Decision Support Framework: Part 3 Overview of Systematic Reviews and Updated Framework. *Med Decis Making*. 2020;40(3):379-398. doi:10.1177/0272989X20911870

Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Thomson R, Trevena L. Decision aids for people facing health treatment or screening decisions. Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD001431.DOI: 10.1002/14651858.CD001431.pub5

Thomson RG, Murtagh MJ, Khaw FM. Tensions in public health policy: patient engagement, evidence-based public health and health inequalities QSHC. 2005;14:398-400

Thomson R, Eccles M, Steen N, Greenaway J, Stobbart L, Murtagh M, May C. A patient decision aid to support shared decision making on antithrombotic treatment of patients with atrial fibrillation: randomised controlled trial. QSHC. 2007; 16: 216-223

Trenaman L, Bryan S, Bansback N. The cost-effectiveness of patient decision aids: A systematic review. Healthc (Amst). 2014 Dec;2(4):251-7. doi: 10.1016/j.hjdsi.2014.09.002. Epub 2014 Oct 3.

Trenaman L, Jansen J, Blumenthal-Barby J, Körner M, Lally J, Matlock D, Perestelo-Perez L, Ropka M, Stirling C, Valentine K, Vo H, Wills CE, Thomson R, Sepucha K.

Are We Improving? Update and Critical Appraisal of the Reporting of Decision Process and Quality Measures in Trials Evaluating Patient Decision Aids. Med Decis Making. 2021 Oct;41(7):954-959. doi: 10.1177/0272989X211011120. Epub 2021 May 8. PMID: 33966534; PMCID: PMC8474325.

Tversky A, Kahneman D. Judgement under uncertainty: heuristics and biases. Science. 1974; 185: 1124-1131.

UK Supreme Court. Montgomery *v* Lanarkshire Health Board. 11 Mar 2015. https://www.supremecourt.uk/cases/docs/uksc-2013-0136-judgment.pdf (Accessed 17 Jan 2020)

Vodermaier A, Caspari C, Koehm J , Kahlert S , Ditsch N, Untch M. Contextual factors in shared decision making: a randomised controlled trial in women with a strong suspicion of breast cancer. Br J Cancer. 2009; 100: 590-597

Von Neumann, J. and O. Morgenstern (1947). Theory of games and economic behavior. New York, Wiley.

Walsh T, Barr PJ, Thompson R, Ozanne E, O'Neill C, Elwyn G. Undetermined impact of patient decision support interventions on healthcare costs and savings: systematic review. BMJ. 2014 Jan 23;348:g188. doi: 10.1136/bmj.g188.