Assessing the conceptual clarity and evidence base of quality criteria∕standards developed for evaluating decision aids

Heather McDonald MSc,* Cathy Charles PhD† and Amiram Gafni PhD‡

*PhD Candidate, Health Research Methodology (HRM) Programme, McMaster University, Hamilton, ON, †Professor, Department of Clinical Epidemiology and Biostatistics, Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, ON, and ‡Professor, Department of Clinical Epidemiology and Biostatistics, Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, ON, Canada

Correspondence
Heather McDonald, MSc
Department of Clinical Epidemiology and Biostatistics
Centre for Health Economics and Policy Analysis
McMaster University (CRL-208)
1280 Main Street West
Hamilton
ON L8S4K1
Canada
E-mail: heatherpauline2000@yahoo.com

Abstract

Context Promoting patient participation in treatment decision making is of increasing interest to researchers, clinicians and policy makers. Decision aids (DAs) are advocated as one way to help achieve this goal. Despite their proliferation, there has been little agreement on criteria or standards for evaluating these tools. To fill this gap, an international collaboration of researchers and others interested in the development, content and quality of DAs have worked over the past several years to develop a checklist and, based on this checklist, an instrument for determining whether any given DA meets a defined set of quality criteria.

Objective∕Methods In this paper, we offer a framework for assessing the conceptual clarity and evidence base used to support the development of quality criteria∕standards for evaluating DAs. We then apply this framework to assess the conceptual clarity and evidence base underlying the International Patient Decision Aids Standards (IPDAS) checklist criteria for one of the checklist domains: how best to present in DAs probability information to patients on treatment benefits and risks.

Conclusion We found that some of the central concepts underlying the presenting probabilities domain were not defined. We also found gaps in the empirical evidence and theoretical support for this domain and criteria within this domain. Finally, we offer suggestions for steps that should be undertaken for further development and refinement of quality standards for DAs in the future.

Introduction

Promoting patient participation in treatment decision making in the medical encounter is of increasing interest to researchers, clinicians and policy makers.[1–5] Wennberg notes that 'More effort is needed to develop models for implementing shared decision making (SDM) into everyday practice and measure the quality of the patient decision-making'.[6] He suggests that

'Policy makers need to undertake experiments in reimbursement reform that support the necessary infrastructure and reward clinicians who successfully implement SDM'.[6] The use of decision aids (DAs) is advocated as one way to facilitate patient participation in treatment decision making.

Implementing shared decision making in the medical encounter requires effective communication between physicians and patients, a process that can be challenging and complex. There may be a variety of concepts and information items that a physician needs to communicate to her patients and which patients need to absorb, understand and retain, such as: (i) a description of treatment options and their benefits and side effects and (ii) the concept of uncertainty in treatment outcomes. Typically DAs convey such information with the intent to facilitate patient understanding and involvement in the treatment decision-making process to the extent desired by the patient.

The growth in the number of DAs has been dramatic. According to Elwyn *et al.*,[7] 'By 1999, approximately 15 patient DAs had been developed in academic institutions. More than 500 now exist, produced largely by a mix of not for profit and commercial organizations'. The design of a DA (i.e. the way in which information on treatment options, their potential benefits and side effects is communicated to a patient) can have a marked impact on a patient's understanding of treatment options and on her treatment decision. However, there has been a lack of agreement on the most credible criteria by which to evaluate these tools. To help fill this gap, a criteria-based checklist was recently developed by the International Patient Decision Aids Standards (IPDAS) Collaboration for evaluating 'the development, content and effectiveness of decision aids'.[7] An instrument derived from this criteria-based checklist and designed to measure the quality of patient DAs (the IPDASi) has also been developed.

Although the developers acknowledge that the IPDASi is in a pilot stage and will be subject to further evaluation, the instrument is being promoted as ready and available for use as a quality assessment tool for evaluating DAs in various settings. Training courses on how to use the IPDASi are available, and the IPDAS Collaboration offers their services to conduct formal assessments of DAs using the IPDASi.[8] In addition, the IPDAS Collaboration is also considering an accreditation process for DAs.[9]

Given the widespread development and use of DAs in a variety of clinical settings and the recent policy recommendations advocating increased use, the need to develop standardized criteria to guide the development and evaluate the quality of DAs is clearly important. The early recognition by the IPDAS Collaboration of the need for such standards deserves positive recognition, as does their efforts to be inclusive in the development process. However, there are many complex conceptual and methodological issues involved in developing quality criteria to be used as standards. These include ensuring the clarity and appropriateness of goals set for DAs (which will determine the content and scope of criteria developed to assess attainment of these goals), defining and operationalizing key constructs underlying quality criteria and identifying the theoretical and/or empirical links between quality criteria and goal attainment. The perceived sense of urgency to develop and use quality criteria for evaluating DAs needs to be balanced against the challenges involved in identifying criteria that are clear, transparent and rigorous.

In this paper, we offer a framework for assessing the conceptual clarity and evidence base used to support the development of quality criteria/standards for evaluating DAs. We then apply this framework to assess the IPDAS checklist criteria on how best to present in DAs probability information to patients on treatment benefits and risks. Based on our results, we conclude by offering some suggestions for steps that we think should be undertaken in the further development and refinement of quality standards for DAs in the future. The IPDAS Collaboration defines DAs as 'tools designed to help people participate in decision making about health care options'.[10] As our paper focuses on the writings of the IPDAS Collaboration, we use their definition of DAs in this paper.

Table 1 The 12 domains of the IPDAS Quality Criteria Checklist[12]

Using a systematic development process
Providing information about (treatment) options
Presenting probabilities
Clarifying and expressing values
Using personal stories
Guiding ∕ Coaching in deliberation and communication
Disclosing conflicts of interest
Delivering DAs on the internet
Balancing the presentation of options
Using plain language
Basing information on up-to-date scientific evidence
Establishing the effectiveness of DAs

Background: development of the IPDAS checklist

The IPDAS Collaboration was developed in 2005 to 'establish an internationally approved set of criteria to determine the quality of patient decision aids'.[11] The process used to develop these criteria has been described by Elwyn *et al.*[7] First, the IPDAS Collaboration identified 12 broad quality domains for DAs (see Table 1). These domains are considered by the IPDAS Collaboration to be important components for evaluating the development, content and effectiveness of DAs. According to Elwyn *et al.* 'Members of the shared decision making electronic listserve, composed of 181 interested academics and practitioners first discussed the validity of the domains'.[7]

Next, 12 panels prepared 'background evidence reports' for each domain. These reports were to include 'definitions of key concepts; theoretical links between the domain and decision quality and evidence to support the inclusion or exclusion of suggested domain criteria, including fundamental studies and results from the systematic review of 34 randomised trials'.[7] Based on the information collected in these reports, the IPDAS Collaboration developed a set of initial quality criteria for each domain. The criteria were reviewed and edited by the Steering, Methods and Evidence Review groups and by a plain language expert and were then assessed in a pilot test.

In the final stage of the development process, individuals from four stakeholder groups (patients, health practitioners, policymakers and DA developers and researchers) were invited to participate in a voting process in which the importance of each criterion was rated on a scale of 1 ('not important') to 9 ('very important') in two consecutive rounds. Participants could add free text comments and, in the second round, raters received a summary of the results from the previous round. Criteria receiving an overall equimedian rating of 7–9 in the second round were included in the final checklist. The IPDAS Collaboration also considered the level of disagreement among voters. If, for a given criterion, 30% or more of the ratings in the second round scored between 1 and 3 (i.e. at the 'not important' end of the scale) and 30% or more scored between 7 and 9 (i.e. at the 'important' end of the scale), then the IPDAS Collaboration considered this to be evidence that voters disagreed on the criterion, and the criterion was consequently eliminated from the checklist.

A checklist comprised of 74 individual quality criteria spread across 12 quality domains was produced through this development process. As noted earlier, the IPDAS Collaboration are now using an instrument derived from the checklist called the IPDASi to evaluate various DAs and are training others to use the IPDASi, either for evaluating existing DAs or for developing new ones.[8]

An analytic framework for assessing the conceptual clarity and evidence base of quality standards ∕ criteria for evaluating decision aids

As noted above, the conceptual and methodological issues involved in developing quality criteria ∕ standards for DAs are complex. First, ideally, the goals of DAs (which the quality criteria are intended to facilitate achieving) need to be clearly stated and a rationale needs to be provided for their importance. Some developers of quality standards for DAs, including the IPDAS Collaboration, also define a variety of specific domains for DAs to include. In such

cases, the intended contribution of these domains to achieving the overarching goals of DAs needs to be specified. Specific criteria for achieving higher order goals (either domain or more directly overarching goals) then need to be defined, described and justified. Second, the key terms and constructs identified at each of these levels of analysis need to be defined and operationalized. Third, the theoretical and/or empirical mechanisms by which quality criteria (specific DA design features) link to goal attainment (desired outcomes) need to be identified to justify setting specified criteria as quality standards. All of the above steps contribute to building a transparent process for developing quality criteria standards.

The above normative statements can be reworded as questions and linked together to form an analytic framework that can be used to assess the conceptual clarity and evidence base of quality criteria/standards for evaluating DAs. The specific questions included at each level of analysis (goals, domains and criteria) are as follows: (i) What is the overarching goal or goals of the DA and each domain (if specified) and specific criteria? (ii) Why is the overarching goal or goals of the DA and each domain (if specified) and specific criteria seen as important? (iii) Are key terms and constructs underlying the overarching goal or goals of DA and each domain (if specified) and specific criteria clearly

defined and operationalized so that the relationship between the features of a DA and achievement of specific goals can be measured? (iv) What support is there, either theoretical, empirical or both, for making predictions about the mechanisms by which particular design features of a given DA can be expected to produce a particular outcome at each of the three levels of analysis? This framework is depicted as a diagram in Fig. 1.

We suggest that the framework is useful for several reasons. First, it offers a transparent approach to critically assessing how standards for DAs are developed and, hence, the credibility of the resulting criteria. Second, the framework helps the reader identify important steps that might be missing in descriptions of any given standards development process for DAs. Finally, the framework provides a bounded and uniform context for thinking through the rationale for inclusion or exclusion of specific DA goals and design features and provides a structure for documenting decisions made and the rationale for these decisions (an audit trail) at each step of the overall process. Admittedly, this framework reflects an ideal that is likely difficult to achieve in practice. However, given the tension between a policy imperative to 'do something', i.e. to develop a set of quality standards for the development and content of DAs as quickly as possible and a scientific norm of
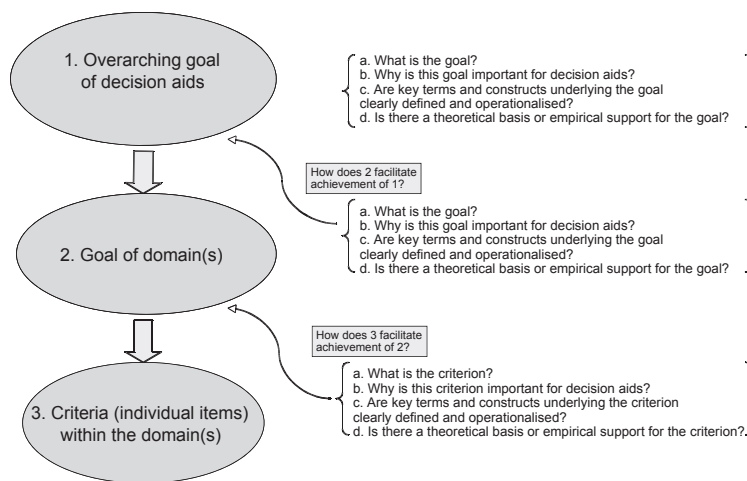

Figure 1 An analytic framework for assessing the appropriateness of goals (domains and criteria) defined for decision aids.

promoting rigor in the development of DAs, we hope that this framework can be used constructively to identify conceptual and methodological issues that still need to be addressed in future work in this field.

Application of the framework

To apply our framework to the presenting probabilities domain of the IPDAS checklist, we reviewed key IPDAS documents and references cited by IPDAS. A more detailed description of our document review is provided in Appendix A.

Specification by IPDAS of the overarching goal for decision aids

The IPDAS Collaboration states that the overarching goal of DAs is to improve the quality of decisions and defines decision quality as 'the extent to which patients choose and/or receive health care interventions that are congruent with their informed and considered values'.[7,10] Applying our framework at this level of analysis, we can raise several issues. For example, while the overarching goal for DAs is clearly defined, key concepts underlying this goal such as 'informed' and 'considered' values are not, raising concerns about how one would know whether these goals have been achieved. To our knowledge, despite additional work undertaken by the IPDAS group, these concepts have still not been defined.

The identification of the above overarching goal for DAs (achievement of a quality decision) is one of several that could have been chosen. This goal, for example, is different from the goal of educational interventions that are aimed at presenting a broader range of information and do not aim to assist decision-making processes. Furthermore, even if there is widespread agreement that the overarching goal of DAs should be to improve the quality of decisions, others may have different ideas about what counts as a quality decision. For example, advocates of evidence-based medicine may argue that a critical review of the quality and strength of the clinical research evidence supporting a given treatment relative to others is the best way to improve the quality of treatment decisions. Criteria for defining a quality decision, under this assumption, relate primarily to the positioning of the research evidence in terms of a hierarchy of research designs. For the purposes of this paper, however, we do not challenge the IPDAS Collaboration's overarching goal. Instead, we accept the goal proposed by the IPDAS Collaboration and apply the rest of our analytical framework accordingly.

When developing quality criteria for DAs, it is important that developers identify the underlying theory of decision making under conditions of uncertainty that they subscribe to and that should inform the choice of relevant quality criteria and how these should be measured. The IPDAS Collaboration does not provide in the written documents we reviewed a theoretical framework of decision making to which they subscribe and from which both the goals defined for DAs and the criteria developed for evaluating the achievement of these goals derive. There are many theories of how people make decisions under conditions of uncertainty and we do not know which, if any, of these theories provided the underlying framework for the development of the quality criteria produced by the IPDAS Collaboration. We did not think it would be appropriate for us to pick a decision-making theory and then evaluate the IPDAS criteria on the extent to which they fit the theory, which we have chosen but which the IPDAS Collaboration has never mentioned as one they support. Hence, we have not adopted a particular theoretical perspective in our review but rather use the analytic framework we developed as a guide for our analysis.

IPDAS Domain on how to present probability information on treatment options and their benefits and side effects to patients

Because of time and space constraints, we focus below on only one domain, although the framework can be used to assess checklist criteria included in any of the 12 domains identified by

the IPDAS Collaboration. We chose the 'presenting probabilities' domain to discuss in relation to our framework because the concept of uncertainty of treatment outcomes (both good and bad) at the individual level is an important yet complex idea that physicians need to communicate to patients, both for the purposes of achieving informed consent and/or participating in the decision making process. The issue is complex because, while randomised controlled trials provide information on average patient outcomes from different treatments, patients are interested in what will happen to them, as individuals. At the individual patient level, there are four possible outcomes of a particular treatment: a patient may experience both the benefits and the side effects (both to varying degrees of magnitude), the benefits with no side effects, no benefits but still side effects or no benefits and no side effects. However, we cannot say with certainty in advance of the treatment which outcomes and/or side effects any given patient will experience.

In terms of our first framework question, the IPDAS Collaboration, in the presenting probabilities section of their Background Document state that 'A key objective of patient DAs is to provide information to help patients understand the possible benefits and harms of their choice, and the chances that these will occur'.[12] From this statement, one can infer that the goal of the presenting probabilities domain is to help patients understand the chances, or uncertainty, associated with the potential outcomes (i.e. the potential benefits, side effects and their magnitude) of their treatment options. In terms of our second framework question, the rationale given for this goal is provided: 'Since no intervention is 100% effective in all patients without harms (including side effects), probabilities must be presented in DAs'.[13] The IPDAS Collaboration indicates that event rates should be used to present probabilities. However, event rates are based on population averages, and no rationale is provided for why presenting statistical probabilities about the *average* clinical outcomes (both good and bad) of specified treatments in a defined research group of patients is considered the *best* way to help individual patients under-

stand the uncertainty associated with the potential outcomes of their therapeutic options (the second framework question). The IPDAS Collaboration recognizes the need to address uncertainty around the averages, but this uncertainty is still based on data from a population of patients and therefore does not provide information on what will happen to an individual patient.

In terms of our third framework question, some of the key constructs underlying the goal of the presenting probabilities domain are not explicitly identified by the IPDAS Collaboration. For example, the meaning of 'patient understanding', 'chances' and 'uncertainty' are not defined. On the positive side, technical concepts (such as 'probability', 'event rate' and 'framing') are defined in the glossary of the first and second round voting documents.

In terms of our fourth framework question, the IPDAS Collaboration documents do not offer a theoretical basis for making predictions about how presenting probabilities is expected to help patients understand the uncertainty associated with potential treatment outcomes and what this means to them at the individual level. Empirical support for the hypothesized link between presenting probabilities and increased patient understanding is presented in the form of a summary of a systematic review of DAs published through the Cochrane Collaboration.[13] Among other comparisons, the Cochrane review reported on eight studies evaluating the effect of DAs on 'patients' perceived probabilities of outcomes'.[13] All eight studies showed a trend towards more 'realistic expectations' in patients who received a DA that included descriptions of outcomes and probabilities. Nonetheless, the relevance of these studies as empirical support that presenting probabilities helps patients understand the uncertainty of treatment outcomes is debatable. When assessing the effects of DAs on realistic expectations 'perceived outcome probabilities were classified according to the percentage of individuals whose judgements corresponded to the scientific evidence about the chances of an outcome for similar people'.[13] It is unclear, however, whether

'realistic expectations' is simply capturing the ability of individuals to recall information on probabilities or whether it actually reflects individuals' understanding of (i) the research evidence concerning the likelihood of the benefits and harms of various treatment options, (ii) the relevance of this information to them, personally, and (iii) the uncertainty of treatment outcomes. Patient recall of probabilities and patient understanding (e.g. interpretation of what the probabilities mean for them) are the two different concepts[14] hence, the empirical evidence summarized in the Cochrane review may not offer support that presenting probabilities facilitates patient understanding.

Quality criteria within the presenting probabilities domain

The criteria proposed for the presenting probabilities domain are clearly stated and are shown in Table 2. Quality criteria are generally expressed in everyday language and seem to be self-explanatory. However, the meaning of the concept of 'uncertainty around probabilities' in criterion 3.4 may be open to different interpretations.

The IPDAS authors indicate that their recommendations on quality criteria for how to present probability information to patients are largely made on theoretical grounds, 'borrowing heavily from work in clinical epidemiology and evidence-based health care, psychology, risk communication and risk perception research and decision theory'.[12] The references cited include Tversky and Kahneman,[15,16] Loewenstein et al.[17] Slovic et al.[18] and von Neumann and Morganstern.[19] These references, however, do not provide satisfactory theoretical support for the criteria within the presenting probabilities domain. First, the theories are not complementary, as each one was developed to replace the previous one because of observed violations in predictions made by the predecessor. For example, numerous studies have shown that individuals systematically violate the predictions of Von Neumann and Morganstern's Expected Utility (EU) model.[20–22] In light of these violations, Tversky and Kahneman developed Prospect Theory as an alternative that could explain many observations that EU theory could not.[15,16] Noting that Prospect Theory was also unable to explain various phenomena related to decision making under uncertainty, Lowenstein et al.[17] and Slovic et al.[18] proposed the Risk-As-Feelings hypothesis and the Affect Heuristic, respectively. As the theories above are not comple-

Table 2 Proposed criteria within the IPDAS presenting probabilities domain[7]

3.1.	The patient DA presents probabilities using event rates in a defined group of patients for a specified time
3.2.	The patient DA compares the probabilities of options using the same denominator
3.3.	The patient DA compares probabilities of options over the same period of time
3.4.	The patient DA describes the uncertainty around the probabilities (e.g. by giving a range or by using phrases such as 'our best guess is')
3.5.	The patient DA uses visual diagrams to show the probabilities (e.g. faces, stick figures or bar charts)
3.6.	The patient DA uses the same scales in the diagrams comparing options
3.7.	The patient DA provides more than one way of explaining the probabilities (e.g. words, numbers, diagrams)
3.8.	The patient DA allows patients to select a way of viewing the probabilities (e.g. words, numbers, diagrams)
3.9.	The patient DA allows patients to see the probabilities of what might happen based on their own individual situations (e.g. specific to their age or severity of their disease)
3.10.	The patient DA places the chances of what might happen in the context of other situations (e.g. chances of developing other diseases, dying of other diseases or dying from any cause)
3.11.	The patient DA has a section that shows how the probabilities were calculated*
3.12.	If the chance of disease is provided by sub-groups, the patient DA describes the tool that was used to estimate the risks*
3.13.	The patient DA presents probabilities using both positive and negative frames (e.g. showing both survival rates and death rates)

*After two rounds of voting, these criteria were removed from the checklist.

mentary, it is unclear how the IPDAS group intended for these to be used as support for the criteria within the presenting probabilities domain. Interestingly, the theories of Lowenstein *et al.*[17] and Slovic *et al.*[18] argue that the probabilities are not as important to decision making as cognitive theories suggest, postulating that individuals may be much more 'sensitive to the possibility rather than the probability'[17] of events. Thus, the inclusion of these theories as theoretical support for the presenting probabilities domain is even more puzzling because they seem to question the importance of presenting probabilities in decision making as a way of increasing patient understanding.

In addition, while the references propose different theories for how individuals make decisions under uncertainty, they do not offer an explanation for how the proposed criteria will contribute to achieving the goal of helping patients understand the uncertainty associated with the potential outcomes of their therapeutic options or why these criteria are the best ones to use in pursuit of this goal. Therefore, they do not address the task of providing a theoretical basis to explain the mechanisms by which the criteria within the presenting probabilities domain support the goal of this domain.

Examples of key findings from our review of the empirical support provided by the IPDAS Collaboration for criteria within the presenting probabilities domain are provided below.

Presenting numbers
The IPDAS Collaboration recommend that event rates should be used in presenting probabilities (criterion 3.1). Two references are provided,[23,24] neither of which offers empirical support for this claim. For example, one reference found that use of '1 in X' scales (i.e. scales with the same numerator and varying denominators) is hard for patients to use and 'perform substantially worse than the other scales (evaluated)',[23] while criterion 3.2 of the IPDAS checklist recommends that patient DAs compare the probabilities of options using the same denominator for event rates.

Visual aids
The IPDAS Collaboration recommends that visual aids be used in presenting probability information to patients (criterion 3.5). Two references are offered in the Background Document in support of presenting visual aids, but neither offers direct empirical support for the criterion. One of the references cites a study by Feldman-Stewart *et al.*,[25] in which different formats for presenting information visually were compared but the study does not offer support for the more fundamental link suggested in criterion 3.5 that visual aids facilitate accurate understanding of probabilities. Of note, Feldman-Stewart *et al.* state that 'There are few systematic, comprehensive, empirical studies of quantitative information and there is virtually no information about what format is best for patients making medical treatment decisions'.[25] Therefore, even the reference provided by the IPDAS Collaboration seems to underscore the scarcity of data to support this criterion.

Framing probabilities
The IPDAS group recommends that 'the patient DA [should] present probabilities using both positive and negative frames' (criterion 3.13) because the way that information is presented can affect preferences and decision making. The reference offered for the existence of framing effects[26] provides a degree of empirical support for the criterion. However, there are inconsistencies between the criterion and the empirical support. For example, the study found no clear pattern of framing effects depending on whether the information was presented in a positive vs. negative format, while criterion 3.13 suggests that both positive and negative frames should be used.

Conveying uncertainty
Criterion 3.4 states that the DA should describe the uncertainty around the probabilities (e.g. by giving a range or by using phrases such as 'our best guess is'). Two references are provided to support this criterion.[27,28] However, these papers are prescriptive in nature and do not offer direct empirical support for the recommendation

from the IPDAS Collaboration that describing uncertainty around probabilities using specific phrases helps patients understand the possible benefits and harms of their choice and the chances that these will occur.[12]

Tailoring probabilities, probabilities in context and evidence for probabilities used

Criterion 3.9 indicates that the DA (should) allow patients to see the probabilities of what might happen based on their own individual situations (e.g. specific to their age or severity of their disease). Criterion 3.10 recommends that a DA place the chances of what might happen in the context of other situations (e.g. chances of developing other diseases, dying of other diseases or dying from any cause). Criterion 3.11 indicates that the patient DA should have a section that shows how the probabilities were calculated and 3.12 indicates that if the chance of disease is provided by the sub-groups, the patient DA (should) describe the tool that was used to estimate the risks. None of these criteria are supported with specific references in the Background Document.

Discussion

Patient DAs have emerged as one prominent mechanism to facilitate patient understanding and involvement in the treatment decision-making process, a practice that is increasingly under consideration at a policy level for the implementation as part of standard clinical practice. The IPDAS Collaboration was formed to develop quality criteria to assess the development, content and quality of DAs.[1] In this paper, we presented a framework for assessing the conceptual clarity and evidence base of quality criteria/standards developed for evaluating DAs. We then applied this framework to offer a critical review of the IPDAS quality criteria checklist in one domain: how to present probability information to patients on treatment benefits and side effects.

We found first that some of the central concepts underlying the inclusion of this domain in the quality standards, such as uncertainty in treatment outcomes, were not defined. Second we found gaps in empirical evidence, contradictory evidence and gaps in theoretical support for inferring that this domain, and particular criteria within this domain, would lead to increased patient understanding of information on treatment outcomes, the primary goal of this domain. We found similar gaps in the linkage of the presenting probabilities domain to the stated overarching goal for DAs. Given these gaps, it is unclear how the IPDAS Collaboration moved from description of available evidence, to interpretation of the evidence, to decisions about what specific criteria to include in the checklist that was sent out for voting. We do not know, for example, if and how 'thresholds' were defined for concluding whether the evidence (and/or theoretical support) was sufficient to justify inclusion of specific criteria in the checklist, and if any rules developed to evaluate the evidence were consistently applied across criteria.

In a recent conceptual paper integrating the science behind decision making with the demands of designing complex health care interventions, Bekker[29] discusses whether using the IPDAS quality criteria checklist as a gold standard to judge a DA's quality may be premature. She notes that '… given that the evidence-base underpinning the IPDAS checklist is likely to change, it seems premature to encourage patient decision aid developers to adhere to it uncritically and/or to use it as a gold standard with which to judge intervention quality'.[29] She further suggests that 'encouraging people to adhere to the IPDAS checklist rather than reason critically about the content of the intervention could be detrimental to and/or stifle, the development of patient decision aids and their evaluations'.[29] One could argue that the danger of (premature) closure on the IPDAS quality criteria checklist is already foreshadowed by recent events. The IPDAS instrument, or IPDASi, based on the IPDAS quality criteria checklist is now being promoted as a way to 'assess whether a DA has undergone a comprehensive and rigorous development' and to 'provide assurance a decision aid has been considered against internationally agreed standards of quality'.[8]

The developers of the IPDAS quality criteria checklist were positively motivated to solve an important problem related to the increasing number of DAs and the ongoing debate about how these should be developed and evaluated. The Collaboration has also indicated publicly that they will be undertaking further work to update the Background Documents and test various versions of the IPDASi in different contexts.[9,30] Based on our results and with the benefit of hindsight, we offer the following suggestions for further development work.

1. We suggest that the key concepts underlying the goal of the presenting probabilities domain be clearly defined. Without a clear description of such concepts, it is difficult to evaluate whether the criteria included in the checklist to achieve the goal of this domain (increasing patient understanding of the concept of uncertainty and of the benefits and side effects of relevant treatment options) are achieved. A similar review should be undertaken of concepts underlying the remaining domains and criteria within these.

2. In updating the Background Documents, we suggest that the IPDAS Collaboration adopt an analytic framework (ours or another framework if a better one can be found) to move from: (i) description of the evidence to ii) evaluation of the evidence to (iii) decision making about what domains and criteria to include in the checklist. In particular, a more transparent process is needed to clarify what type of evidence and theoretical justification are needed to support the inclusion of (i) particular domains and criteria within the checklist and (ii) thresholds used for making these decisions. Where there is insufficient information to make a judgment, recommendations for further research could be made.

3. We think that identifying theoretical frameworks that inform how DAs should be developed is critically important. Durand *et al.* explored the extent to which theoretical frameworks informed the development and evaluation of decision support technologies by assessing literature cited in a 2003 Cochrane review of DAs.[13,31] They concluded that there is a need to 'give more attention to how the most important decision-making theories could be better used to guide the design of key decision support components and their modes of action'.[31] Our framework can be used to help assess the links between the goals of DAs and the theoretical support in the literature for design features hypothesized to achieve these goals.

4. We think it would be useful to apply our framework (or another one if a better one can be found) to assess the types and levels of theoretical and empirical support for other domains and criteria in the IPDAS checklist.

Conclusion

In this paper, we offer a framework for assessing the conceptual clarity and evidence base used to support the development of quality criteria/standards for evaluating DAs, in general. We then applied this framework to assess the IPDAS checklist criteria on how DAs can best present probability information to patients on treatment benefits and risks. Based on our results, we recommend that a similar assessment be undertaken of the other IPDAS checklist domains and criteria. If the empirical evidence and theoretical support for the criteria included in the checklist are found to be unclear, then clarifying these links should be an important focus of future research attention.

Appendix A

A note on methods

To undertake our analysis, we first identified and reviewed publicly available documents related to the IPDAS quality criteria checklist development process and posted on the IPDAS website: http://www.ipdas.ohri.ca/index.html. The following documents, which we label as primary documents, were reviewed:

1. IPDAS Background Document[12]
2. IPDAS First Round Voting Document[32]
3. IPDAS Second Round Voting Document[33]

4. IPDAS Checklist for Judging the Quality of Patient Decision Aids[34]
5. Publication describing the development of the IPDAS quality criteria checklist[7]

We also reviewed secondary documents cited in the IPDAS Background Document as offering either theoretical or empirical support for the inclusion of this domain and related criteria. The steps in our analysis were as follows. First, we read all the primary documents above prepared by the IPDAS Collaboration to determine whether there were clear statements regarding the goal(s) of the presenting probabilities domain and whether key constructs underlying the goal(s) of this domain (e.g. chance/probability), and the quality criteria within these were clearly defined and operationalized. Second, we looked for references cited in the Background Document (the first primary document listed above) that were offered as either theoretical or empirical support to justify including this domain and criteria within this domain in the checklist. Third, we located and read each secondary reference offered as providing support (as per above) for each criterion and summarized in tables whether there was an appropriate match (i.e. whether and extent to which the secondary reference cited in the primary document provided theoretical or empirical information to justify inclusion of each criterion in the quality checklist). Finally, we used the data in these tables to answer each of the questions in our analytic framework.

References

1 Llewellyn-Thomas HA. Studying patients' preferences in health care decision-making. *Canadian Medical Association Journal*, 1992; **147:** 859.

2 Llewellyn-Thomas HA. Patients' health-care decision making: a framework for descriptive and experimental investigations. *Medical Decision Making*, 1995; **15:** 101.

3 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). *Social Science and Medicine*, 1997; **44:** 681–692.

4 Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Social Science and Medicine*, 1999; **49:** 651–661.

5 Charles C, Gafni A, Whelan T, O'Brien M. Cultural influences on the physician-patient encounter: the case of shared treatment decision-making. *Patient Education and Counseling*, 2006; **63:** 262–267.

6 Wennberg J. Practice variation, shared decision-making and health care policy. In: Edwards A, Elwyn G (eds) *Shared Decision-Making in Health Care: Achieving Evidence-Based Patient Choice*. Oxford: Oxford University Press, 2009: 1–3.

7 Elwyn G, O'Connor A, Stacey D *et al*. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *British Medical Journal*, 2006; **333:** 417.

8 IPDASi Assessment Brochure. http://www.ipdasi. org/IPDASi%20Information.pdf, accessed 15 June 2010.

9 IPDAS Collaboration website: What's new section. http://www.ipdas.ohri.ca/news.html, accessed 15 June 2010.

10 IPDAS Collaboration website. What are Patient Decision Aids? http://www.ipdas.ohri.ca.what.html, accessed 15 June 2010.

11 IPDAS Collboration website. Home page. http://www.ipdas.ohri.ca/index.html, accessed 15 June 2010.

12 O'Connor A, Llewellyn-Thomas H, Stacey D (eds). *IPDAS Collaboration Background Document*. 2003; Available at: http://ipdas.ohri.ca/IPDAS_Background.pdf, accessed 31 January 2009.

13 O'Connor AM, Stacey D, Entwistle V *et al*. Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews* 2003; Art. No.: CD001431. DOI: 10.1002/14651858.CD001431.

14 Charles C, Redko C, Whelan T, Gafni W, Renyo L. Doing nothing is no choice: lay constructions of treatment decision-making among women with early-stage breast cancer. *Sociology of Health and Illness*, 1998; **20:** 71–95.

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

16 Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*, 1981; **211:** 453–458.

17 Loewenstein G, Weber EU, Hsee CK, Welch N. Risk as feelings. *Psychological Bulletin*, 2001; **127:** 267–286.

18 Slovic P, Finucane M, Peters E, MacGregor DG. The affect heuristic. In: Gilovich T, Griffin D, Kahneman D (eds) *Heuristics and Biases*. New York: Cambridge University Press, 2002: 397–420.

19 Von Neumann K, Morganstern O. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1947.

20 Bell DE, Farquhar PE. Perspective on utility theory. *Operations Research*, 1986; **34:** 179–183.

21 Michina MJ. Choice under uncertainty: problems solved and unsolved. *Economic Perspectives*, 1987; **1:** 121–154.

22 Fishburn PC. Expected utility: an anniversary and a new era. *Journal of Risk and Uncertainty*, 1988; **1:** 267–288.

23 Woloshin W, Schwartz LM, Byram S, Fischhoff B, Welch HG. A new scale for assessing perceptions of chance. *Medical Decision Making*, 2000; **20:** 298–307.

24 Woloshin W, Schwartz LM, Welch HG. Risk charts: putting cancer in context. *Journal of the National Cancer Institute*, 2002; **94:** 799–804.

25 Feldman-Stewart D, Kocovski N, McConnell BA, Brundage MD, Mackillop WJ. Perception of quantitative information for treatment decisions. *Medical Decision Making*, 2000; **20:** s228–s238.

26 Edwards A, Elwyn G, Covey J, Matthews E, Pill R. Presenting risk information – a review of the effects of "Framing" and other manipulations on patient outcomes. *Journal of Health Communication*, 2001; **6:** 61–82.

27 Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *British Medical Journal*, 2002; **324:** 827–830.

28 Skolbekken J. Communicating the risk reduction achieved by cholesterol reducing drugs. *British Medical Journal*, 1998; **316:** 1956–1958.

29 Bekker HL. The loss of reason in patient decision aid research: do checklists damage the quality of informed choice interventions? *Patient Education and Counseling*, 2010; **78:** 357–364.

30 Elwyn G, O'Connor AM, Bennett C *et al*. Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASi). *PLoS ONE* 2009; **4:** e4705. Epub 2009 Mar 4.

31 Durand M-A, Stiel M, Boivin J, Elwyn G. Where is the theory? Evaluating the theoretical frameworks described in decision support technologies. *Patient Education and Counseling*, 2008; **71:** 125–135.

32 O'Connor AG, Elwyn G, Stacey D (eds). *IPDAS Voting Document*. 2005; Available at: http://ipdas.ohri.ca/IPDAS_First_Round.pdf, accessed 31 January 2009.

33 O'Connor AG, Elwyn G, Stacey D (eds). *IPDAS Voting Document– 2nd Round*. 2005; Available at:http://www.ipdas.ohri.ca/IPDAS_Second_Round.pdf, accessed 31 January 2009.

34 IPDAS Collaboration. 2005. *Criteria for Judging the Quality of Patient Decision Aids*. Available at: http://www.ipdas.ohri.ca/IPDAS_checklist.pdf, accessed 31 January 2009.