The multifocal approach to sharing in shared decision making: a critical appraisal of the MAPPIN'SDM

Forner, D.; Noel, C.W.; Boland, L.; Pieterse, A.H.; Borkhoff, C.M.; Hong, P.

Review

The Multifocal Approach to Sharing in Shared Decision Making: A Critical Appraisal of the MAPPIN'SDM

David Forner⑩, Christopher W. Noel, Laura Boland⑩, Arwen H. Pieterse⑩, Cornelia M. Borkhoff, and Paul Hong

Objective. Shared decision making integrates health care provider expertise with patient values and preferences. The MAPPIN'SDM is a recently developed measurement instrument that incorporates physician, patient, and observer perspectives during medical consultations. This review sought to critically appraise the development, sensibility, reliability, and validity of the MAPPIN'SDM and to determine in which settings it has been used. **Methods**. This critical appraisal was performed through a targeted review of the literature. Articles outlining the development or measurement property assessment of the MAPPIN'SDM or that used the instrument for predictor or outcome purposes were identified. **Results**. Thirteen studies were included. The MAPPIN'SDM was developed by both adapting and building on previous shared decision making measurement instruments, as well as through creation of novel items. Content validity, face validity, and item quality of the MAPPIN'SDM are adequate. Internal consistency ranged from 0.91 to 0.94 and agreement statistics from 0.41 to 0.92. The MAPPIN'SDM has been evaluated in several populations and settings, ranging from chronic disease to acute oncological settings. Limitations include high reading levels required for self-administered patient questionnaires and the small number of studies that have employed the instrument to date. **Conclusion**. The MAPPIN'SDM generally shows adequate development, sensibility, reliability, and validity in preliminary testing and holds promise for shared decision making research integrating multiple perspectives. Further research is needed to develop its use in other patient populations and to assess patient understanding of complex item wording.

Shared decision making (SDM) is a collaborative process that allows patients and their health care providers to make medical decisions together, incorporating scientific expertise while considering the values and preferences of the patient.[1] SDM is particularly relevant for preference-sensitive decisions (when there is more than one reasonable treatment option).[1] While this unifying definition for SDM is pervasive, there is a lack of a unified conceptual model.[2] There is an increasing interest in context-specific conceptual models, as SDM is theorized to differ in unique scenarios.[3] Further, SDM may differ substantially depending on the various perspectives that may be

taken to assess its quality.[4] The importance of each individual actor within consultations cannot be understated, nor can the view by which their participation in SDM is assessed, such as self-perception or external observation.[5] Incongruency between SDM instruments and their outcomes therefore underpinned the development of a

Corresponding Author
David Forner, Division of Otolaryngology–Head & Neck Surgery, Victoria General Hospital, Dalhousie University, Queen Elizabeth II Health Sciences Centre, Dickson Building, Halifax, NS B3H 4R2, Canada (David.Forner@dal.ca).

cohesive SDM inventory integrating multiple perspectives with the goal of analyzing intersubjective perceptions of SDM.

The MAPPIN'SDM is a measurement instrument that was designed to assess SDM by examining interactions from 3 perspectives, those of the patient, physician, and observer.[6] The instrument measures both the behavior contributing to, and resulting from, participation in SDM according to the physician–patient dyad. Behaviors are defined as actions that attempt to involve both parties in the decision making process, and results are considered the perceptions of involvement in the process. The physician–patient dyad is defined as the integration of the patient and physician relationship. As it is important to consider which party contributes aspects of SDM to the interaction, as well as the extent to which each party contributes, all 3 (physician, patient, physician–patient dyad) units are assessed.[6]

Separate self-administered questionnaires for physicians (referred to as $Qdoc_{dyad(behavior)}$ and $Qdoc_{dyad(result)}$, which measure the physician's perception of the behaviors employed by the dyad and the SDM result of those behaviors, respectively) and patients (referred to as $Qpat_{dyad(behavior)}$ and $Qpat_{dyad(result)}$, which measure the patient's perception of the behaviors employed by the dyad and the SDM result of those behaviors, respectively) are employed within the MAPPIN'SDM. Observer-based components of the instrument allow external assessment of the physician, patient, and physician–patient dyad contributions to SDM by using 3 separate observer-administered questionnaires (Obs_{doctor}, $Obs_{patient}$, and Obs_{dyad}). The 3 observer-based questionnaires and 4 self-administered questionnaires comprise the 7 "foci" of interest to the instrument. The MAPPIN'SDM is currently the only SDM measurement instrument to integrate these multiple perspectives and

consider both the behavior related to, and the perception of, SDM.

The appropriateness of a measurement instrument for a specific use is a reflection of its development, sensibility, reliability, and validity.[7,8] Previous systematic reviews of SDM measurement instruments have not assessed these properties in depth, and thus this crucial set of information is lacking in the literature for specific measurement instruments.[2,4] Regardless of the specific conceptual model employed by clinicians and researchers, in order to accurately measure the construct of interest, sound measurement properties of the instrument of interest must be ensured. This review therefore sought to critically appraise the development and measurement properties of the MAPPIN'SDM and to provide a compendium of settings it has currently been used in.

Methods

This critical appraisal was informed through a targeted search of the MEDLINE, Embase, and CINAHL databases for articles that developed or used any MAPPIN'SDM version. Search terms included "MAPPIN'SDM" or "Multifocal Approach to Sharing" (Suppl. Table S1). Eligible studies included any randomized controlled trials, case-control, case series ($N \geq 5$), cohort, cross-sectional, or mixed-methods studies that discussed the development or measurement property testing of the MAPPIN'SDM or that used the MAPPIN'SDM as a predictor or an outcome for any population of patients from April 2012 to April 2020. Articles were screened and abstracted by a single reviewer (DF).

This critical appraisal focuses on instrument development, sensibility, reliability, and validity of the MAPPIN'SDM and provides a synthesis of its measurement properties. Instrument development includes item generation, item reduction, and the creation of the rating scale for items in the instrument.

Division of Otolaryngology–Head & Neck Surgery, Dalhousie University, Halifax, NS, Canada (DF, PH); Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada (DF, CWN, CMB); Department of Otolaryngology–Head & Neck Surgery, University of Toronto, Toronto, ON, Canada (CWN); Knowledge Translation Laboratory, Health Sciences, Western University, London, ON, Canada (LB); Integrated Knowledge Translation Research Network, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada (LB); Medical Decision Making, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Zuid-Holland, The Netherlands (AHP); Pediatric Outcomes Research Team, Department of Pediatrics and Child Health Evaluative Sciences, Research Institute, Hospital for Sick Children, Toronto, ON, Canada (CMB). The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) received no financial support for the research, authorship, and/or publication of this article.

Search Results

The search strategy yielded 13 nonduplicate articles (Suppl. Figure S2). Reference review and snowballing techniques were used to further identify relevant articles. Ultimately, 13 articles were included in the study after screening (Table 1). Most studies were cross-sectional in design ($n = 7$), were performed predominantly in Germany and Norway, and were published from 2012 to 2019. Study sample sizes ranged from 27 to 197 participants. Seven studies reported on the development of the MAPPIN'SDM or directly assessed its measurement

Table 1 Study Characteristics

Author	Year	Journal	Country	Design	Participants	Development or Measurement Properties	Predictor or Outcome[a]
Kasper et al.[6,b]	2012	PLoS One	Germany	Cross-sectional	Dental medicine, MS clinic, GP; N = 40	✓	
Kasper et al.[5]	2012	ZEFQ	Germany	Cross-sectional	MS clinic, N = 76	✓	
Geiger et al.[33]	2012	ZEFQ	Germany	Cross-sectional	Dental medicine, MS clinic, GP; N = 40	✓	
Flierler[47]	2013	Anaesthesia	Germany	Cross-sectional	Anesthesia; N = 197		✓
Berger-Hoger et al.[15]	2017	BMC Bioinformatics and Decision Making	Germany	Pilot study, prospective cohort	DCIS; N = 27		✓
Geiger et al.[33]	2017	Patient Education and Counselling	Germany	Randomized controlled trial	Dental medicine, MS clinic, stroke unit, SCT unit, radiation oncology, surgery, gynecology; N = 152	✓	✓
Kasper et al.[34]	2017	Health Expectations	Germany	Prospective cohort	Dental medicine, MS clinic, GP; N = 40		✓
Kienlin et al.[41]	2017	Patient Education and Counselling	Norway	Cross-sectional	Surgery, internal medicine, pediatrics, obstetrics and gynecology, neurology; N = 32	✓	✓
Landmark[48]	2017	Patient Education and Counselling	Norway	Mixed methods, cross-sectional	Varying nonpsychiatry medical specialties; N = 27		✓
Geessink et al.[16]	2018	Patient Education and Counselling	Netherlands	Cross-sectional	Colorectal and pancreatic cancer; N = 80	✓	✓
Rahn et al.[19]	2018	International Journal of Nursing Studies	Germany	Mixed methods, randomized controlled trial	MS clinic, N = 73		✓
Rose et al.[17]	2019	Clinical Rehabilitation	United Kingdom	Mixed methods, cross-sectional	Intermediate care rehabilitation; N = 40		✓
Berger-Hoger et al.[15]	2019	International Journal of Nursing Studies	Germany	Randomized controlled trial	DCIS; N = 192	✓	✓

DCIS, ductal carcinoma in situ; GP, general practice; MS, multiple sclerosis; SCT, stem cell transplant.

[a]Refers to studies that used the MAPPIN'SDM as either a predictor (independent) variable or as an outcome (dependent) variable beyond sole use for development or testing of measurement properties.

[b]Original development paper.

properties, and 10 studies included the MAPPIN'SDM as a predictor or outcome.

Instrument Development

Item Generation

Items within the MAPPIN'SDM were generated primarily through expert opinion and amalgamated with the preexisting 12-item Observing Patient Involvement in Decision Making (OPTION) scale.[9] Experts involved in the development of the MAPPIN'SDM were experienced SDM researchers. Two items from the OPTION scale were merged due to lack of distinctiveness ("expectations" and "worries").[5] In order to stress the importance of chronological information exchange during the SDM process, items within the OPTION scale were rearranged. Four new items were developed in order to incorporate additional aspects of SDM deemed important by experts, including supporting decision making, evaluation of physician understanding of patient values, providing opportunities for questions from the physician to the patient, and indicating the source of information provided as evidence. Therefore, 15 items per observer foci were developed, yielding 45 observer-based items. All items were either obtained from the OPTION scale or created through the expert opinion of the authors.

Fifteen items were adapted from the observation instrument to create the self-administered instrument for the patient- and physician-related foci. As each item consisted of 2 parts each ("behavior" and "results"), 30 items were created. Development of the patient- and physician-reported items is not well defined in published literature or widely available in supporting documents.

The physician- and patient-based questionnaires were pilot tested with 10 participants each, leading to iterative, stepwise revisions of item wording; however, exact details beyond this are not included in published, peer-reviewed materials. Physicians in the pilot test included 4 clinicians from a multiple sclerosis clinic, 3 from dental medicine, and 3 general practitioners in private practice.[6] There are no details on the sample of patients included.

Item Reduction

No item reduction was performed for the first version of the MAPPIN'SDM. The total number of items for each perspective was reduced from 15 to 11 in the most recent version. In the most recently available training manual, the authors state that there was a "restructuring of the sub-criteria underpinning the MAPPIN'SDM concept" due to rater training experiences showing lack of distinctiveness between some items. How particular items were combined was a function of expert opinion and not specific psychometric or statistical methods.

Rating Scale Development

Each observer item is scored from 0 (poor performance) to 4 (excellent performance). Each physician- or patient-reported item is scored as agreement about a statement, with response options ranging from 0 (*statement is not at all true*) to 4 (*absolutely true*). Additional detailed response options and appropriate choices for the observation instruments are found within the dedicated training manual, promoting transparency and consistency in the scoring process.[10] For example, in items that focus on defining the problem, a score of 1 is consistent with the problem being defined, but there is no reference to the patient's condition or expectations.

There is no original scoring method specified for the original MAPPIN'SDM aside from comparisons of the raw scores of each of the 7 foci ($Qdoc_{dyad(behavior)}$, $Qdoc_{dyad(result)}$, $Qpat_{dyad(behavior)}$, $Qpat_{dyad(result)}$, Obs_{doctor}, $Obs_{patient}$, and Obs_{dyad}).

Sensibility

Sensibility is often defined as enlightened common sense that incorporates both ordinary common sense and clinical knowledge.[11] No formal sensibility assessments of MAPPIN'SDM have been undertaken. This critical appraisal uses Bombardier's framework,[12,13] adapted from Feinstein[11] and Rowe and Oxman,[14] as the framework for assessing sensibility. The Bombardier framework considered the purpose, population, and setting of an instrument, as well as its content validity, face validity, and feasibility.

Purpose, Population, and Setting

The intent of the MAPPIN'SDM is well defined—to integrate multiple relevant perspectives in the assessment of SDM. The intended population of patients and physicians is clear, and what specific disciplines and settings were used during the development phase are well presented. The instrument has also been successfully used in the oncology and rehabilitation settings.[15–17] While there is no specific purpose that can be mapped to various frameworks (e.g., the methodological framework for assessing health indices by Kirshner and Guyatt[18]), the MAPPIN'SDM is descriptive of SDM in general and has been successfully used as a discriminative instrument.[15,19] For example, scores on the MAPPIN'SDM

have been shown to be significantly different between patients using a decision aid to promote SDM compared to those that did not use a decision aid.[15] Notably, the patient–physician dyad must agree on what decision is being assessed before completing self-administered questionnaires.

Content Validity

Overall, content validity of the MAPPIN'SDM, as assessed by the authors, seems good and incorporates content that is reflective of its underlying constructs (i.e., SDM consists of the perceptions of physicians, patients, and external observers), although this has not been formally assessed by stakeholders external to the development group. The authors integrate 7 foci in the inventory, but some experts have called for a longitudinal assessment of SDM,[20] as well as the inclusion of patient supporters such as family members and caregivers.[21] However, while these considerations have not been included within the MAPPIN'SDM instrument, conceptual models and definitions of SDM are dynamic, and thus the exclusion of patient supporters is not particularly limiting overall.[3,22] The MAPPIN'SDM situates itself using a construct of SDM that values the physician, patient, and external observer perceptions of SDM at a singular time point but not necessarily caregivers, family members, or other individuals over time. Thus, researchers and clinicians must be wary of this when assessing whether the MAPPIN'SDM is appropriate for their required use. The categories within the MAPPIN'SDM are sensible and, with reduced overlap of items in the 11-item version, can be considered mutually exclusive.

Face Validity

The items of MAPPIN'SDM make intuitive sense and thus have good evidence of face validity. For example, it is straightforward what "the clinician elicits the patient's preferred level of involvement in decision making" means. That said, double-barreled questions in both the self-administered questionnaire and the observer-based instrument are found, although these are mostly resolved in the 11-item version. Examples include "The way I exchanged information with the doctor during the consultation suited both parties and contributed towards a mutual understanding." As well, there has been no published assessment of the face validity of the MAPPIN'SDM by groups of end users. Despite being developed in German, published versions of the MAPPIN'SDM are shown in English, and the language structure is satisfactory. That said, another important caveat

is the high grade level and advanced language used in patient-facing questions. For example, item 2a reads, "It was stated within the consultation that from a medical point of view there is not only one correct way to deal with my problem and the doctor cannot decide on his own which option is the right one in my case." This statement represents a Flesch-Kincaid grade level of 16.5, and other items are similarly advanced.[23] There have been no formal assessments of interpretability of items by patients. It is known that education and reading levels have an impact on both the perception of SDM involvement and the interpretation and response to instrument items. As such, this could create potential issues with the instrument.[24] However, the phrasing of the physician- and observer-based instruments seems appropriate, and possible issues with patient reading levels could be mitigated in the research setting with administrative assistance.

Feasibility

Feasibility assesses whether the measurement instrument could be used within a research or clinical environment or if aspects of its administration would be prohibitive to its use. Within a research setting, the administration of the MAPPIN'SDM is feasible, although possibly more time and resource intensive to use than other less comprehensive measures such as the SDM-Q-9 or OPTION5.[25,26] Video or audio recordings of the consultation are necessary for the observer-based instrument. In a research setting, the MAPPIN'SDM seems feasible, although there are clear logistical obstacles to overcome. Within a clinical setting, the authors believe current iterations of the MAPPIN'SDM would not be feasible, as portions of the instrument require external review of the videorecorded consultation, and busy clinic environments may not support routine use of the self-administered questionnaires.

Reliability

Reliability is the degree to which measurements of individuals on different occasions, or by different observers, produce the same or similar results.[27] Reliability is therefore a comparison of variance in the true scores to total variance, including systematic and random error.[28,29] Furthermore, reliability is a prerequisite for validity and responsiveness, as it must be shown that instrument results are consistent across administrations ceteris paribus. Across various settings and populations, 7 studies have assessed different forms of reliability for the MAPPIN'SDM, including internal consistency and

agreeability (Table 2). As reliability is not a fixed property of the instrument but is instead a characteristic of the test scores (and therefore is unique to the setting and population under consideration), assessment of reliability in specific populations is essential.[30,31] Further adding to the complexity of assessing the reliability of SDM instruments is the perception of the actors involved—patients and providers may change their minds on their levels of involvement over time. The time in which questionnaires were administered was not readily available in the majority of studies included in this critical appraisal.

Internal Consistency

Measures of internal consistency demonstrate how well items within the instrument relate to each other.[32] Three studies[6,33,34] supplied measures of internal consistency (Table 2). Cronbach's α was shown to range between 0.91 and 0.94 and was given for each of the 4 self-administered questionnaires (Qpat$_{dyad(behavior);dyad(result)}$ [i.e., the patient's perception of the behaviors and results of those behaviors] and Qdoc$_{dyad(behavior);dyad(result)}$ [i.e., the physician's perception of the behaviors and results of those behaviors]). These results are acceptable, as defined by Nunnally,[35] where 0.7 to 0.8 is a minimum for early scale development and research purposes. The internal consistency of the MAPPIN'SDM approaches values upon which redundancy of the items may be of concern (i.e., higher than 0.90 to 0.95[27]). On the other hand, application of Cronbach's α may not be appropriate overall. For nearly a decade,[36] some authors have considered SDM to be a formative[37] model,[22] and thus the Cronbach's α (typically applied for reflective models and instruments) would not be suitable.[38] As Cronbach's α is provided in the original development study,[6] it could be assumed the model is reflective. However, the developers do not make explicit the type of model underlying the MAPPIN'SDM, and thus this assumption may be questioned.

Agreement and Reproducibility

Of those studies that documented interrater reliability, the majority reported correlation statistics, including the Pearson correlation coefficient and Spearman ρ (Table 2). Only 1 study used intraclass correlation (ICC) as a measure of interrater reliability, which is the preferred practice given it measures both correlation and agreement, as opposed to correlation alone.[16,39] The ICC was sufficient for both physician and dyad (0.81–0.85; Table 2) but was insufficient for interrater agreement of patient involvement in the SDM process (0.61; Table 2).

Validity

Validity is the degree to which an instrument measures what it is intending to measure,[32] of which 2 main types are forefront: criterion and construct validity. Several studies (below) have shown the MAPPIN'SDM to be valid for the measurement of interperspective SDM in different populations of patients, including in the settings of multiple sclerosis, pancreatic cancer, dental medicine, colon cancer, and others.

Criterion Validity

Criterion validity is the degree to which an instrument produces a similar result to a gold standard, or criterion measure. Shared decision making criterion measures have not been established.[40] Only a single study has attempted to demonstrate criterion validity of the MAPPIN'SDM.[41] Kienlin et al.,[41] in a study of 35 decision sequences across 17 medical specialties, contrasted the ability of the MAPPIN'SDM observer subinstrument to measure the presence of SDM compared to a reference standard, defined as a senior SDM researcher categorizing sequences as having SDM or not. The MAPPIN'SDM showed excellent sensitivity (78%–93%) and specificity (80%–91%) across each of the patient, doctor, and dyad levels. However, it should be noted that the senior researcher acting as the criterion measure was the senior investigator in the MAPPIN'SDM development study, and bias is therefore inherent.

Construct Validity

Additional studies have investigated the construct validity of the MAPPIN'SDM.[6,16,41] Construct validity is the degree to which an instrument measures what it intends to measure and is tested against a series of a priori defined hypotheses.[42] In its initial development, the creators hypothesized that there would be limited congruence between measurement perspectives, a positive correlation between physician and patient SDM activity, interperspective agreement in those with high SDM skill, reduced correlation between physician and dyad with higher patient activity, high within-party correlations for behavior and result, and high within-party correlations for the items *equipoise*, *communication of risks*, and *agreeing on a decision*.[6] While these hypotheses were supplied a priori, no magnitudes were given. Ultimately, 4 of these 6 hypotheses were confirmed across the study population consisting of multiple sclerosis, dental medicine, and general practice environments.

Table 2 Reliability of the MAPPIN'SDM

Study	Internal Consistency	Instrument Aspect	Result	Agreement or Reproducibility	Instrument Aspect	Result
Kasper et al. 2012[6]	Cronbach's α	$Qdoc_{dyad(b)}$	0.91	Pearson	Obs_{doctor}	0.87
		$Qdoc_{dyad(r)}$	0.94		$Obs_{patient}$	0.81
		$Qpat_{dyad(b)}$	0.92		Obs_{dyad}	0.74
		$Qpat_{dyad(r)}$	0.94			
Kasper et al. 2012[5]					Obs_{doctor}	0.92
					$Obs_{patient}$	0.41
					Obs_{dyad}	0.85
Berger-Hoger et al.[15]				Modified Cohen's κ[a,b] (Spearman ρ)	Obs_{dyad}	0.77 (0.68)
Geessink et al.[16]				ICC (95% CI)	Obs_{doctor}	0.85 (0.76–0.90)
					$Obs_{patient}$	0.61 (0.36–0.76)
					Obs_{dyad}	0.81 (0.67–0.88)
Kasper et al.[34]	Cronbach's $α$[c]	Qdoc	0.94	Spearman ρ	Obs_{doctor}	0.87
		Qpat	0.94		$Obs_{patient}$	0.81
					Obs_{dyad}	0.74
Kienlin et al.[41]				Modified Cohen's κ[a,b] (percent agreement)	Obs_{doctor}	0.77 (0.79)
					$Obs_{patient}$	0.82 (0.86)
					Obs_{dyad}	0.77 (0.82)
Geiger et al.[33]	Cronbach's $α$[c]	Qdoc	0.94	Spearman ρ (range)[d]	Obs_{doctor}	0.74–0.87
		Qpat	0.94		$Obs_{patient}$	
					Obs_{dyad}	

CI, confidence interval; ICC, intraclass correlation.

[a]Mean κ of all items within the scale reported.

[b]Uses theoretical instead of expected values for calculation of the κ statistic—no detail directly provided for these assumptions.

[c]Uncertain if behavior, result, or combined.

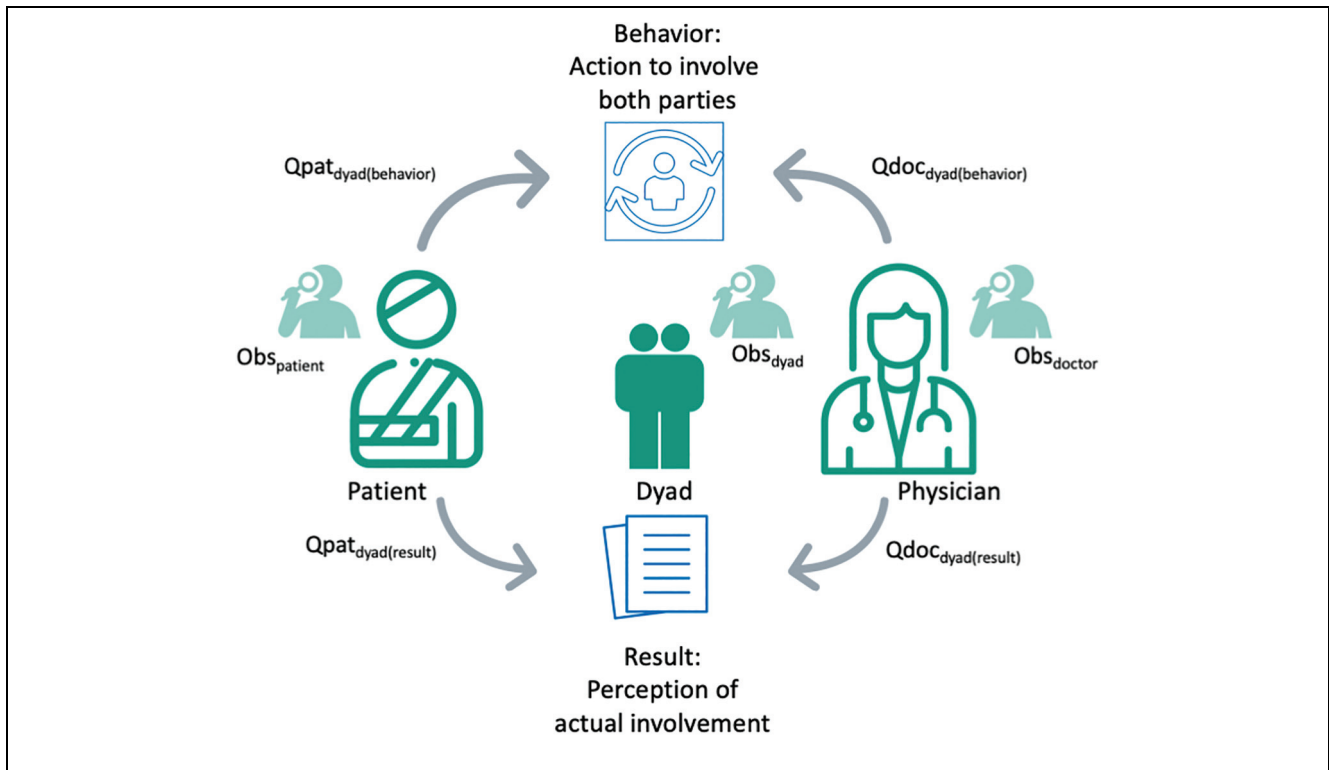[d]Does not report individual instrument interrater reliability, only a range.

Figure 1 Visualization of the 7 foci assessed by the MAPPIN'SDM, including the patient's perception of the behavior ($Qpat_{dyad(behavior)}$) and result ($Qpat_{dyad(result)}$) of the interaction, the physician's perception ($Qdoc_{dyad(behavior)}$ and $Qdoc_{dyad(result)}$), and an external reviewer's perception of involvement of the patient ($Obs_{patient}$), physician (Obs_{doctor}), and the combined dyad (Obs_{dyad}).

In a study of 80 pancreatic and colon cancer patients, Geessink et al.[16] investigated the convergent validity of the MAPPIN'SDM,[16] a subtype of construct validity. There was very strong correlation between Obs_{doctor} and scores on the OPTION scales ($OPTION_5$ and $OPTION_{12}$; Spearman ρ 0.845 and 0.841, respectively) and strong correlation between Obs_{dyad} and the OPTION scales ($OPTION_5$ and $OPTION_{12}$; 0.819 and 0.770, respectively). As the observer-based MAP-PIN'SDM was developed using the OPTION scales, this finding is not surprising and is a limitation of the study. There was a lower, albeit moderate, correlation between the OPTION scales and $Obs_{patient}$ ($OPTION_5$ and $OPTION_{12}$; 0.526 and 0.525, respectively). This is expected given the focus of the OPTION scales on the physician only. The same pattern was seen by Kienlin et al.,[41] who found the Obs_{doctor} and the Obs_{dyad} to correlate more strongly with the $OPTION_5$ than the $Obs_{patient}$, although each was only low to moderately correlated.[41] However, in both studies, a priori hypotheses had not been defined.[16,41]

Synthesis and Discussion

This review describes the development process and psychometric testing literature available for the MAP-PIN'SDM. This is the only available SDM measurement instrument that integrates the perspectives of the patient, physician, and observer with assessments of each party as well as the patient–physician dyad. As such, it is uniquely positioned within SDM research. Through this critical appraisal, strengths and limitations have been identified that investigators should consider before using the MAPPIN'SDM instrument.

The MAPPIN'SDM was developed using expert opinion and incorporation of items from a preestablished observer-based instrument. While the content validity of the instrument appears satisfactory, it should not be assumed that all aspects of SDM are captured as formal testing has not been performed.[43,44] Content validity may also be assessed through a computational assessment of the agreement of content experts, which could be employed at an institutional level before adoption of the MAPPIN'SDM.[43] Furthermore, full understanding of

the development, and therefore utilization of, the MAP-PIN'SDM is limited by a lack of description of how the novel items were created.

Assessments of the reliability and validity of the MAP-PIN'SDM are promising but remain at a paucity, and their individual study qualities are generally low.[45] The majority of reliability studies employed correlation statistics only and therefore did not assess the agreement between raters. Furthermore, validation studies often did not fully specify their a priori hypotheses, including an expected magnitude of correlation. Recognizing that the MAPPIN'SDM is still in its infancy, further studies are needed to gather more evidence on this aspect of the instrument.

The reading level required for the self-administered patient questionnaires is high. It is possible that the formally translated English versions could contain more appropriate language compared to currently available translations that do not appear to have been performed through standard cross-culture validation approaches.[46] As well, reading-level issues may be partially mitigated with standardized guidance from a research assistant in a research setting where all patients are "walked through" the questionnaire. However, this was not the intended use of the instrument and creates additional burden in an already complex measurement tool, potentially limiting its use in large-scale research studies. As well, the measurement properties of the instrument may differ according to the route of administration used.

Last, the MAPPIN'SDM has been used in a variety of settings and thus may be suitable for use in similar populations. However, further demonstration of the measurement properties of the MAPPIN'SDM in unrelated settings would be necessary before widespread usage of the instrument is possible.

Whether the perception of involvement in the SDM process determined by the patient or physician should necessarily converge with the involvement in the SDM process as assessed by external reviewers remains to be determined. For clinicians and researchers with aspirations to integrate and assess all 3 of these viewpoints (patient, physician, and observer), this critical appraisal offers an assessment of the various measurement properties of the only currently available measurement instrument that may satisfy such goals. This critical appraisal was not designed to assess whether the underlying conceptual model employed by the MAPPIN'SDM is necessarily correct. Indeed, it may be assumed that observers and participants may reasonably be convergent or divergent in their perceptions of SDM.[16]

Conclusion

The MAPPIN'SDM is a comprehensive inventory integrating patient, physician, and observer perspectives across both SDM behaviors and resulting perceptions. Administrative burdens are expected to be high, and the self-administered items have potentially prohibitively high reading-level requirements for patient participants. However, given that the MAPPIN'SDM is the only instrument of its kind, seems to have otherwise reasonable sensibility, and has demonstrated preliminarily adequate measurement properties, it is a valuable contribution to SDM research.

ORCID iDs

David Forner https://orcid.org/0000-0002-1014-3545
Laura Boland https://orcid.org/0000-0001-8680-9844
Arwen H. Pieterse https://orcid.org/0000-0001-6395-0052

Supplementary Material

Supplementary material for this article is available on the *Medical Decision Making* website at http://journals.sagepub.com/home/mdm.

References

1. Elwyn G, Frosch D, Rollnick S. Dual equipoise shared decision making: definitions for decision and behaviour support interventions. *Implementation Sci*. 2009;4(1):75.
2. Bouniols N, Leclère B, Moret L. Evaluating the quality of shared decision making during the patient-carer encounter: a systematic review of tools. *BMC Res Notes*. 2016;9(1):382.
3. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. *BMJ Open*. 2019;9(12):1–11.
4. Gärtner FR, Bomhof-Roordink H, Smith IP, Scholl I, Stiggelbout AM, Pieterse AH. The quality of instruments to assess the process of shared decision making: a systematic review. *PLoS One*. 2018;13(2):e0191747.
5. Kasper J, Hoffmann F, Heesen C, Köpke S, Geiger F. Completing the third person's perspective on patients' involvement in medical decision-making: approaching the full picture. *Z Evid Fortbild Qual Gesundheitswes*. 2012;106(4):275–83.
6. Kasper J, Hoffmann F, Heesen C, Köpke S, Geiger F. MAPPIN'SDM–the multifocal approach to sharing in shared decision making. *PLoS One*. 2012;7(4):e34849.
7. Jung JJ, Borkhoff CM, Jüni P, Grantcharov TP. Non-Technical Skills for Surgeons (NOTSS): critical appraisal of its measurement properties. *Am J Surg*. 2018;216(5):990–7.
8. Prasad K. The Glasgow Coma Scale: a critical appraisal of its clinimetric properties. *J Clin Epidemiol*. 1996;49(7):755–63.

9. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. *BMJ Qual Saf.* 2003;12(2):93–9.

10. Kasper J, Liethmann K. Manual for training and coding MAPPIN'SDM revised-multifocal approach to the 'sharing' in SDM. Hamburg. 2014.

11. Feinstein AR. *Clinimetrics.* New Haven, CT: Yale University Press; 1987.

12. Bombardier C, Tugwell P. Methodological considerations in functional assessment. *J Rheumatol Suppl.* 1987;14:6–10.

13. Buchbinder R, Goel V, Bombardier C, Hogg-Johnson S. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? *J Clin Epidemiol.* 1996;49(2):141–9.

14. Rowe BH, Oxman AD. An assessment of the sensibility of a quality-of-life instrument. *Am J Emerg Med.* 1993;11(4): 374–80.

15. Berger-Hoeger B, Liethmann K, Muehlhauser I, Haastert B, Steckelberg A. Nurse-led coaching of shared decision-making for women with ductal carcinoma in situ in breast care centers: a cluster randomized controlled trial. *Int J Nurs Stud.* 2019;93:141–52.

16. Geessink NH, Ofstad EH, Rikkert MGO, van Goor H, Kasper J, Schoon Y. Shared decision-making in older patients with colorectal or pancreatic cancer: Determinants of patients' and observers' perceptions. *Patient Educ Counsel.* 2018;101(10):1767–74.

17. Rose A, Soundy A, Rosewilliam S. Shared decision-making within goal-setting in rehabilitation: a mixed-methods study. *Clin Rehabil.* 2019;33(3):564–74.

18. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis.* 1985;38(1):27–36.

19. Rahn A, Köpke S, Backhus I, et al. Nurse-led immuno-treatment DEcision Coaching In people with Multiple Sclerosis (DECIMS)—feasibility testing, pilot randomised controlled trial and mixed methods process evaluation. *Int J Nurs Stud.* 2018;78:26–36.

20. Clayman ML, Gulbrandsen P, Morris MA. A patient in the clinic; a person in the world. Why shared decision making needs to center on the person rather than the medical encounter. *Patient Educ Counsel.* 2017;100(3):600–4.

21. Ho A. Relational autonomy or undue pressure? Family's role in medical decision-making. *Scand J Caring Sci.* 2008; 22(1):128–35.

22. Dowie J. Shared decision making is a preference-sensitive formative construct: the implications. *Eur J Person Centered Healthcare.* 2019;7(3):506–17.

23. Flesch R. *Let's Start with the Formula: How to Write Plain English.* New York, NY: Harper & Row; 1979.

24. Ankuda CK, Block SD, Cooper Z, et al. Measuring critical deficits in shared decision making before elective surgery. *Patient Educ Counsel.* 2014;94(3):328–33.

25. Barr PJ, O'Malley AJ, Tsulukidze M, Gionfriddo MR, Montori V, Elwyn G. The psychometric properties of Observer OPTION5, an observer measure of shared decision making. *Patient Educ Counsel.* 2015;98(8):970–6.

26. Elwyn G, Grande SW, Barr P. *Observer OPTION 5 Manual.* Dartmouth: The Dartmouth Institute for Health Policy and Clinical Practice; 2016.

27. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess.* 2003;80(1):99–103.

28. Hoyle RH, Harris MJ, Judd CM. *Research Methods in Social Relations.* Thomson Learning; 2002.

29. Judd CM, Smith ER, Kidder LH. *Research Methods in Social Relations.* Fort Worth, TX: Holt, Rinehart and Winston; 1991.

30. Caruso JC. Reliability generalization of the NEO personality scales. *Educ Psychol Meas.* 2000;60(2):236–54.

31. Wilkinson L; Task Force on Statistical Inference, American Psychological Association, Science Directorate. Statistical methods in psychology journals: Guidelines and explanations. *Am Psychologist.* 1999;54(8):594–604.

32. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use.* Oxford, UK: Oxford University Press; 2015.

33. Geiger F, Liethmann K, Reitz D, Galalae R, Kasper J. Efficacy of the doktormitSDM training module in supporting shared decision making: results from a multicenter double-blind randomized controlled trial. *Patient Educ Counsel.* 2017;100(12):2331–8.

34. Kasper J, Liethmann K, Heesen C, Reissmann DR, Geiger F. Training doctors briefly and in situ to involve their patients in making medical decisions—preliminary testing of a newly developed module. *Health Expectations.* 2017; 20(6):1254–63.

35. Nunnally JC. *Psychometric Theory.* 2nd ed. New York: McGraw-Hill; 1978.

36. Wollschläger D. Where is SDM at home? Putting theoretical constraints on the way shared decision making is measured. *Z Evid Fortbild Qual Gesundheitswes.* 2012;106(4): 272–4.

37. De Vet H, Terwee C, Mokkink L, Knol D. Concepts, theories and models, and types of measurements. In: *Measurement in Medicine: A Practical Guide.* Cambridge, UK: Cambridge University Press; 2011. p 7–29.

38. Neuendorf K. Internal consistency reliability: can Cronbach's alpha be too high? *J Pers.* 2003;80(3):217–22.

39. Jinyuan L, Wan T, Guanqin C, Yin L, Changyong F. Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Arch Psychiatry.* 2016;28(2):115.

40. Elwyn G, Edwards A, Mowle S, et al. Measuring the involvement of patients in shared decision-making: a systematic review of instruments. *Patient Educ Counsel.* 2001;43(1): 5–22.

41. Kienlin S, Kristiansen M, Ofstad E, et al. Validation of the Norwegian version of MAPPIN'SDM, an observation-based instrument to measure shared decision-making in

clinical encounters. *Patient Educ Counsel*. 2017;100(3): 534–41.

42. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297–334.

43. Lawshe CH. A quantitative approach to content validity 1. *Personnel Psychol*. 1975;28(4):563–75.

44. Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for Lawshe's content validity ratio. *Meas Eval Counsel Dev*. 2012;45(3):197–210.

45. Mokkink LB, De Vet HC, Prinsen CA, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171–9.

46. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417–32.

47. Flierler WJ, Nübling M, Kasper J, Heidegger T. Implementation of shared decision making in anaesthesia and its influence on patient satisfaction. *Anaesthesia*. 2013;68(7): 713–22.

48. Landmark AMD, Ofstad EH, Svennevig J. Eliciting patient preferences in shared decision-making (SDM): Comparing conversation analysis and SDM measurements. *Patient Educ Couns*. 2017;100(11):2081–87.