Of blind men and elephants: suggesting SDM-MASS as a compound measure for shared decision making integrating patient, physician and observer views

Von Blinden und Elefanten: Einführung von SDM-MASS als Verbundmaß der geteilten Entscheidungsfindung zur Integration der Sichtweisen von Patient, Arzt und Beobachter

Friedemann Geiger¹,²,*, ⁵, Jürgen Kasper³,⁴,⁵

¹ Tumor Center, University Medical Center Schleswig-Holstein, Kiel, Germany
² Department of Pediatrics, University Medical Center Schleswig-Holstein, Kiel, Germany
³ Institute of Neuroimmunology and Clinical MS Research (INiMS), University Medical Center Hamburg-Eppendorf, Hamburg, Germany
⁴ Department of Primary Medical Care, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Summary
Objective: Shared decision making (SDM) between patient and physician is an interpersonal process. Most SDM measures use the view of one party (patient, physician or observer) as a proxy to capture this process although these views typically diverge. This study suggests the compound measure SDMMASS (SDM Meeting its concept’s ASSumptions) integrating these three perspectives in one single index.

Methods: SDMMASS was derived theoretically and compared empirically to unilateral perspectives of patients, physicians and observers by application to a data set of 10 physicians (40 consultations) receiving an SDM training.

Results: The constituting parts of SDMMASS were highly reliable (Cronbach’s alpha .94; interrater reliability .74-.87). Unilateral appraisal of training effects was divergent. SDMMASS revealed no effect.

Conclusion: SDMMASS combines noteworthy information about SDM processes from different viewpoints and thereby delivers plausible assessments. It could overcome immanent shortcomings of unilateral approaches. However, it is a complex measure needing further validation.

* Correspondence to: Dr. Friedemann Geiger, Klinik für Allgemeine Pädiatrie, Universitätsklinikum Schleswig-Holstein, Schwanenweg 20, 24105 Kiel
E-Mail: f.geiger@uksh.de (F. Geiger).
⁵ Both authors contributed equally to this work.

1865-9217/$ – see front matter
http://dx.doi.org/10.1016/j.zefq.2012.03.020
Zusammenfassung


Ergebnisse: Die konstituierenden Teile von SDM_MASS waren hoch reliabel (Cronbach’s alpha .94; Interrater-Reliabilität .74 bis .87). Die unilaterale Bewertung von Trainingseffekten war divergent. SDM_MASS offenbarte keinen Effekt.


Introduction

"It was six men of Hindustan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind."

This is the beginning of a famous fairy tale [1]. Each Hindu then studies one part of the elephant — side, tusk, trunk, knee, ear, and tail. They compare an elephant to a wall, a spear, a snake, a tree, a fan, or a rope. Their conclusions about the true nature of an elephant hence are quite heterogeneous.

„And so these men of Hindustan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and wrong,
Though each was partly in the right
And all were in the wrong."

Measurement of shared decision making (SDM) may lead to similar disputes. For decades, SDM has been studied by use of unilateral instruments being administered either by patients, physicians or observers [2]. These approaches are still predominant [3]. However, assessing one and the same consultation by use of instruments from different perspectives typically leads to marginal or at best moderate correlations between different viewpoints [4—11]. It has been discussed that if not serious methodological problems inherent in the instruments cause these incongruencies, they might be due to inequality of the SDM sub-constructs addressed by these instruments assessing different points of view [9, 12]. In that case, the actual situation among instruments claiming to assess SDM is similar to that of the Hindus in the tale. A lack of idea of the subject’s entire shape prevents consensus among the Hindus.

In view of such discrepancies, it could be argued, that the patients’ perspective should be prioritized, as they are the main addressees of the communication. It seems arbitrary to say SDM had “happened” if patients do not perceive it “happening”, i.e. perceive themselves as active self-determined and respected partners — a judgment that can only be made by the patients themselves. On the other hand, they may be emotionally affected and thereby not a reliable source of information regarding other aspects of SDM. Furthermore, patients are typically not aware of SDM definitions and underlying concepts of evidence-based risk communication [13, 14].

Others may argue that the physician is to be prioritized, as he or she also participates directly in the decision making process supposedly without facing emotional burden. Moreover, the physician is supposed to guide the communicative process. Thus he or she is the only participant knowing about the intentions behind observable communicative actions. However, physicians may have role conflicts influencing judgment of their own communicative performance. And they may also not be aware of the full gamut of SDM characteristics [14—16].

A possible way to overcome the mentioned shortcomings of patients’ and physicians’ views could be use of trained independent unbiased observers. Video or audio documents allow to thoroughly check if e.g. all available treatment options have been listed. However, while observations may yield reliable assessments of SDM behavior, they might in turn miss essential parts of the communication. Watching a physician listing all possible treatment options does not necessarily mean that the patient has actually understood which options he or she personally has. Additionally, technical restrictions (angle of view, low resolution, noise) could limit the amount of information observers are able to get, especially regarding para- and non-verbal emotional aspects of communication [17].

Trying to determine which party is ’right’ — and whose perspective could be ignored — is an ill-posed question comparable to the Hindu example. Instead, we plea to appreciate and make use of each perspective’s particular advantages. None of the abovementioned exclusive perceptive pathways is dispensable or redundant when trying to fully capture the phenomenon of SDM [11]. Attempts to use only one as a proxy for the others turned out to be misleading in view of the large number of studies reporting discrepancies [4—11]. A recently upcoming tendency to use dyadic approaches in SDM measurement addressing both patient and physician [12, 18—20] already responds to this evidence...
by considering the interdependency of subjective viewpoints as important to judge decision making quality. In this article, we argue for combining and integrating different perspectives as the best approximation to SDM.

Aim of the Study

This study aims to suggest an integrative compound measure for SDM called SDMMASS (SDM Meeting its concept’s ASSumptions) and to evaluate its assumed advantages compared to unilateral measures. In particular, differences between assessing effects of an SDM training using SDMMASS and several unilateral measures are analyzed. The study has an exploratory character.

Methods

Operationalization of SDM and construction of SDMMASS

Construction of a compound measure requires to identify important ingredients of SDM and to define how these are to be combined mathematically in one single SDM index. As outlined above, the views of patients, physicians and observers with their respective unique advantages are noteworthy to entirely grasp the phenomenon of SDM.

To avoid combining apples and oranges, a compound measure should be composed of identical instruments for each perspective. MAPPIN’SDM [see 10 for details] is a measurement system providing such a basis. It comprises seven coherent views (’foci’) on SDM: I) SDM related behavior shown by 1) patient, 2) doctor, 3) doctor-patient-dyad, judged by an independent observer on the basis of video documents; II) the doctor’s perception of 4) the dyad’s SDM related behavior and 5) its perceived result; III) the patient’s perception of 6) the dyad’s SDM related behavior and 7) its perceived result. Each focus is assessed by an identical set of 15 criteria with slightly adapted wording.

We hypothesize that SDM can be sufficiently captured combining three essential aspects out of these seven foci. The patient is the best — and the only — one to judge if and to which extent SDM has been realized in a way that he or she actually understood the need for a decision, the available treatment options, their pros and cons, if personal preferences were respected, etc. This patient-based judgment about the perceived result of the dyad’s SDM behavior is operationalized by MAPPIN’SDM focus 7 (mean scale score).

There are some issues of SDM that can better (or only) be judged by independent observers, e.g. if an available but maybe less profit-making option has been skipped. Even more importantly, typically only trained observers are familiar with optimal communicative ways that provide a frame of reference while ranking the presented SDM performance. This observable SDM-related communicative behavior is operationalized by MAPPIN’SDM focus 3 (mean scale score).

Finally, the dyadic process of SDM requires that patient and physician share the impression that essential aspects of SDM have been realized. This concordance between SDM judgment of patient and physician is operationalized by non-parametric correlation (Kendall’s tau) between MAPPIN’SDM foci 5 and 7.

The formula below shows how these 3 parts are adjusted and combined to the compound measure SDMMASS. To yield equal emphasis, all components are partially rescaled to a range between 0 and 1. Dividing the numerator by 3 leads to SDMMASS scores theoretically ranging from 0 to 1, where 0 indicates no SDM and 1 indicates perfect SDM.

SDMMASS hence ensures that physicians’ observable SDM behavior that has not been noticed and valued successful by the patient is as insufficient as patients’ (deluded) perceptions of involvement if SDM actually had not occurred at all.

$$SDM_{MASS} = \frac{M(MAPPIN - Q_{patient})}{4} + \frac{M(MAPPIN - Q_{dyad})}{4} + \tau_{1-15}(MAPPIN - Q_{patient}; MAPPIN - Q_{doctor})$$

$$M: \text{Mean of items of the particular domain (range 0-4)}$$

$$MAPPIN - Q_{patient}: \text{Patient questionnaire addressing his/her perceived result of SDM}$$

$$MAPPIN - Q_{dyad}: \text{Observer rating addressing the realized amount of SDM on the dyadic level}$$

$$MAPPIN - Q_{doctor}: \text{Physician questionnaire addressing his/her perceived result of SDM}$$

$$\tau_{1-15}: \text{Kendall’s tau coefficient indicating consensus (concordance) on items 1-15; negative values are fixed at 0.}$$

Sample and design

The complete MAPPIN’SDM inventory was applied to a sample of 40 consultations from different medical fields. All consultations comprised a medical decision (or its deferral). Ten physicians documented a sequence of four consultations alternating with SDM training modules (see below for details): baseline (t0), after receiving a training manual (t1), a video tutorial (t2), and a face to face feedback (t3).

Intervention

The intervention is a training curriculum addressing physicians [21,22] aiming at stimulating efforts to involve patients better in the decision making process. The training includes 3 educational components: a) a manual comprehensively explaining SDM skills by use of examples and their assessment by MAPPIN’SDM b) a video tutorial demonstrating examples of excellent performance of each skill, and c) a face to face training session (15 min)
Figure 1. SDM level rated by patient, physician and observer at the 4 measurement points. Physicians and patients rated the dyad’s SDM performance. Observers rated the SDM performance of physician, patient and dyad separately. In this diagram, SDMmass is rescaled to a mean of 2 to match the order of magnitude.

providing structured feedback based on the analysis of a video-recorded consultation from the particular physician referring to the MAPPIN’SDM framework. In this study, feedback was given by one of the authors (JK).

Measurement
Assessment was identical at every measurement point (t0-t3). Each consultation was video-taped. After having agreed on a reference decision (if there was more than one) doctor and patient filled out MAPPIN’SDM questionnaires (MAPPIN-Qdoctor, MAPPIN-Qpatient) directly after the consultation addressing the perceived level of realized SDM. Each video was rated by two independent trained raters by use of MAPPIN’SDM observer scales. The raters were blinded towards the training level of the physician and the questionnaire data. The ratings indicate individual and shared observable attempts to realize SDM by separately rating patient, doctor and dyadic level (MAPPIN-Opatient, MAPPIN-Odoctor, MAPPIN-Odyad). All MAPPIN’SDM scales refer to the identical set of 15 SDM indicators by use of 5-point Likert scales ranging from 0 (poor performance) to 4 (excellent performance) [see 10 for details].

On this basis, SDMmass was calculated according to the abovementioned formula.

Analyses
In this study, intervention effects are operationalised by differences on mean MAPPIN’SDM scale scores between baseline (t0) and after the full training (t3). These differences are analysed using one-tailed non-parametric Wilcoxon tests. Analysis focuses on SDMmass and unilateral MAPPIN’SDM foci. Inter-rater reliability was assessed by Pearson correlation coefficients. Internal consistency of the questionnaire data was assessed by Cronbach’s alpha.

Results
Sample characteristics
Among the 40 participating patients 18 were female. The 10 physicians (3 female) were specialists in neurology, internal, dental and general medicine. Decisions were made regarding treatment of e.g. multiple sclerosis, diabetes, tooth loss, or hypertension. Consultation lengths ranged from 3 to 51 minutes (M = 20; SD = 11). Lengths of decision sequences lay between 3 and 39 minutes (M = 15; SD = 8).

Reliabilities and scale properties
Inter-rater reliability of the observer data was high to excellent (.74 < r < .87). Internal consistency of doctor and patient questionnaire was excellent (alpha = .94 each).

Training effects
After the training physicians showed more attempts to involve their patients (MAPPIN-Odoctor p = .05). However, these invitations to participate did not result in patients actually showing higher involvement (MAPPIN-Opatient p = .11). Hence, the increase of observed SDM within the dyad was slightly above the significance level (MAPPIN-Odyad p = .07).

Physicians’ ratings of subjective views on the communication showed a significant increase from pre to post training.
(MAPPIN-\text{Qdoctor} p = .02), while patients’ did not (MAPPIN-Qpatient p = .87).

According to SDM\text{mass}, the training had no significant effect (p = .17). Figure 1 shows the courses of these six measures.

Discussion

This study aimed to derive the compound measure SDM\text{mass} from theoretical considerations and to compare it to unilateral measures empirically.

The comparison reveals that effects of an intervention being perceived by physicians’ subjectively and assessed by observers objectively disappear if additionally patients’ view and concordance between physician and patient are taken into account by SDM\text{mass}. After the training, observers noticed physicians showing more attempts to involve patients, but they did not detect patients responding or acting as an involved partner. In line with this, physicians reported an increase of sharing in their consultations while patients did not. But acting as if a physician involves patients into decision making processes does not necessarily mean that patients actually are involved or feel involved. Apparently, self-estimation of physicians is strongly oriented towards their own behavior rather than to whether this behavior turns out effective regarding patient involvement. This might also reflect an incomplete stage of implementation of adopted training contents. For example, a consultation where physicians experiment with newly learned skills may appear artificial, somewhat unusual or just mechanistic to the patient which, in turn, perhaps lowers patient ratings of SDM [see also 23, 24].

Either way, as insufficiency of measurement has been ruled out as possible explanations [9,11], incongruities between the three perspectives demonstrate their non-redundancy [10] as they capture different aspects of SDM. This, in turn, confirms once again the shortcomings of traditional unilateral approaches. If only one perspective had been applied, the training would have either been regarded as being effective (physicians’/observers’ view on physicians) or rather counterproductive or at least useless (patients’ view). None of these perspectives can be seen as more important or “true” than the others. But considering several of them simultaneously, as SDM\text{mass} does, yields condensed information about the complete subject. Thereby, assessment is closer to the elephant than any single precise description of tusks, ears and trunks can ever be. However, just as the blind men have to figure out a useful procedure how to integrate their partial knowledge to come closest to the nature of the elephant, SDM\text{mass} should be seen as one possible operationalization of a trans-focal compound measure. Other combinations of constituting foci, multiplicative instead of additive compositions, or other than equal loadings between the components should be discussed by means of theoretical and empirical issues to be able to generate the most veridical capture of SDM.

Apart from the actual composition, a compound measure like SDM\text{mass} is not the best suiting measure for every study about SDM. Some may only focus physicians’ SDM skills, others may be primarily interested in perceptions of the patient. In these cases, assessment of SDM\text{mass} would be too extensive. However, instruments claiming to assess SDM could be validated and normed with SDM\text{mass} as a reference measure.

As another advantage, SDM\text{mass} combines markers of SDM that we call first order and second order indicators. First order indicators are direct measures calculated within one of the possible foci, e.g. mean observable dyadic SDM behavior. Second order indicators result from relations between different foci like degree of congruence or differences between judgments. While the former are easier to generate and to interpret, the latter may come closer to the idea of “sharing” between participants [25]. Furthermore, second order indicators can hardly be biased by one party as a result of social desirability — among physicians pretending to act professionally as well as among patients being afraid of irritating someone they are dependent on. The latter influence may be responsible for the high absolute values in the patient ratings in this study [see also e.g. 26]. In our view, any compound measure should therefore include first and second order indicators because of their respective advantages.

However, due to the small sample size this study in particular and SDM\text{mass} in general have some limitations. Assessing the components of SDM\text{mass} is complex. Furthermore, the results presented here to introduce construction and application of SDM\text{mass} exemplary are drawn from a study piloting the training and the MAPPIN’SDM system [10]. They were not intended to validate SDM\text{mass} — nor the training — in a strict sense. Both is currently done in a multi-center RCT [21].

Another limitation results from the fact that compound measures (like SDM\text{mass}) can currently only be calculated on basis of MAPPIN’SDM as this is at present the only measurement approach assessing SDM from all possible perspectives [11]. Although MAPPIN’SDM has been thoroughly developed and validated and is currently transferred to several countries and languages, any limitation of MAPPIN’SDM also applies to SDM\text{mass}. The next best alternative to MAPPIN’SDM, the expanded Dyadic OPTION system [19], does not comprise an observer scale assessing the SDM performance of the doctor-patient dyad. Moreover, it carries on of the theoretical shortcomings of its unilateral predecessor [9,27]. Therefore, it would not have been reasonable to build SDM\text{mass} based on components drawn from the OPTION system.

Conclusion

Together with other articles in this special issue, this study contributes to the overdue debate on the principals of SDM measurement by suggesting a new approach [28,29]. The current trend towards multi-perspective approaches in SDM measurement [3] is a welcome development. However, it remains unfinished as long as no theoretically derived and empirically proven rationale exists, how different — typically diverging — scale scores can be interpreted to indicate SDM [10]. This inevitably requires integrating compound measures as an addition to analysis of the constituting unilateral perspectives in order to reduce information complexity while holding on to the goal of getting as close as possible to the real nature of the elephant.
Acknowledgements

This work was undertaken as part of the project "Profile of perceived uncertainty as outcome of shared decision making" which is funded by the Tumor Center Kiel. JK has received travel fees to scientific conferences from Merck Serono, FG has nothing to declare. We thank Frauke Hoffmann and Mirela Neukirchner for their assiduous work as raters in the project and Sascha Köpke for carefully reviewing our manuscript.

Ethical approval: Ethics committee of the University Medical Center Kiel (D424/11).

References


