

# The multimorbidity interaction severity index (MISI)

## A proof of concept study

Dimitri Gassmann, MMed<sup>a,b</sup>, Marcus Cheetham, PhD<sup>a,b,c</sup>, Klarissa Siebenhuener, MD<sup>a,b</sup>, Barbara M. Holzer, PhD<sup>a,b</sup>, Claudine Meindl-Fridez, MD<sup>a,b</sup>, Florian F. Hildenbrand, MD<sup>a,b</sup>, Vanessa Virgini, MD<sup>d</sup>, Mike Martin, PhD<sup>b,c,e</sup>, Edouard Battegay, MD, FACP<sup>a,b,c,\*</sup>

### Abstract

Therapeutic decision-making for patients with multimorbidity (MM) is challenging. Clinical practice guidelines inadequately address harmful interactions and resulting therapeutic conflicts within and among diseases. A patient-specific measure of MM severity that takes account of this conflict is needed.

As a proof of concept, we evaluated whether the new Multimorbidity Interaction Severity Index (MISI) could be used to reliably differentiate patients in terms of lower versus higher potential for harmful interactions.

Two hypothetical patient cases were generated, each with 6 concurrent morbidities. One case had a low (i.e., low conflict case) and the other a high (i.e., high conflict case) potential for harmful interactions. All possible interactions between conditions and treatments were extracted from each case's record into a multimorbidity interaction matrix. Experienced general internists (N=18) judged each interaction in the matrix in terms of likely resource utilization needed to manage the interaction. Based on these judgements, a composite index of MM interaction severity, that is, the MISI, was generated for each physician and case.

The difference between each physician's MISI score for the 2 cases (MISI<sub>diff</sub>) was computed. Based on MISI<sub>diff</sub>, the high conflict case was judged to be of significantly greater MM severity than was the low conflict case. The positive values of the inter-quartile range, a measure of variation (or disagreement) between the 2 cases, indicated general consistency of individual physicians in judging MM severity.

The data indicate that the MISI can be used to reliably differentiate hypothetical multimorbid patients in terms of lesser versus greater severity of potentially harmful interactive effects. On this basis, the MISI will be further developed for use in patient-specific assessment and management of MM. The clinical relevance of the MISI as an alternative approach to defining MM severity is discussed.

**Abbreviations:** CSS = cascading style sheets, DDI = disease–disease, drug–disease, and drug–drug interactions, DG = Dimitri Gassmann, EB = Edouard Battegay, HTML = Hypertext Markup Language, MISI = multimorbidity interaction severity index, MM = multimorbidity.

**Keywords:** computerized decision support, drug–disease interactions, drug–drug interactions, MISI, multimorbidity, severity index, therapeutic conflict

Editor: Ahmet Emre Eskazan.

DG and MC contributed equally to this work.

Author Contributions: conceptual development: EB, DG, CM, KS, MC, VV, FH. Technical development: DG, CM, KS, BH. Conceived and designed the experiment: MM, EB, MC, DG. Performed the experiment: DG. Analyzed the data: MC. Contributed to the writing of the manuscript: MC, EB, DG.

This study was funded by the Center of Competence Multimorbidity, University of Zurich, and the University Research Priority Program "Dynamics of Healthy Aging," University of Zurich.

The authors have no conflicts of interest to disclose.

<sup>a</sup> Department of Internal Medicine, University Hospital Zurich, <sup>b</sup> Center of Competence Multimorbidity, <sup>c</sup> University Research Priority Program "Dynamics of Healthy Aging," University of Zurich, <sup>d</sup> Department of Dermatology, University Hospital Zurich, <sup>e</sup> Division of Gerontopsychology and Gerontology, Department of Psychology, University of Zurich, Zurich, Switzerland.

\* Correspondence: Edouard Battegay, Department of Internal Medicine, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland (e-mail: Edouard.Battegay@usz.ch).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2017) 96:8(e6144)

Received: 11 August 2016 / Received in final form: 21 December 2016 /

Accepted: 25 January 2017

<http://dx.doi.org/10.1097/MD.0000000000006144>

## 1. Introduction

Multimorbidity (MM) refers to the presence of multiple concurrent acute or chronic diseases within a person.<sup>[1,2]</sup> Therapeutic decision making for multimorbid patients is challenging.<sup>[3–6]</sup> This is because decisions rely on recommendations from clinical practice guidelines that were developed for the treatment of single diseases.<sup>[7–9]</sup> With some exceptions,<sup>[9–12]</sup> these guidelines are based on evidence from studies that excluded or under-represented multimorbid cases.<sup>[13–16]</sup> This monomorbidity approach to therapeutic decision-making in MM does not, therefore, adequately address the combined impact of potentially harmful disease–disease, drug–disease, and drug–drug interactions (DDIs) and multiple drug regimens<sup>[9,17–21]</sup> or adequately guide the clinical decision-maker through the therapeutic conflict.

Therapeutic decision making in MM requires consideration of potential DDIs<sup>[22,23]</sup> and their combined impact to determine a suitable clinical strategy, on a case-by-case basis. In some patients, co-occurring conditions and treatments can be managed without risk of harmful effects (e.g., physical exercise for hypertension, diabetes, and dyslipidemia) due to the treatment of one of the constituent conditions. But therapeutic conflict of various degrees of severity is typically encountered in that the treatment for 1 condition is contraindicated by the presence of

1 or more other conditions or treatments.<sup>[24]</sup> This applies, for instance, to concurrent gastrointestinal bleeding and anticoagulant treatment for heart disease, concurrent severe lung disease and benzodiazepine treatment for a sleeping disorder, or severe lung disease and opiate treatment for a pain disorder.<sup>[8,25]</sup> The overall complexity of such cases can place particularly high demands on the clinical decision-maker in reconciling the range of harmful interactions with a therapeutic strategy that is specifically tailored to the particular needs of the patient.<sup>[26,27]</sup>

Intensive work to support clinical decision making in MM is in progress (e.g.,<sup>[9,11,13,28]</sup>). But, at present, there is no instrument for measuring patient-specific burden of harmful DDIs that can be applied to any combination of medical conditions. A reliable measure of MM burden would facilitate the development of tools and guidance to support diagnostic and therapeutic decision making and provide a valuable frame of reference for comparing research findings (e.g.,<sup>[29]</sup>). The most commonly applied measures in the field of MM were originally constructed to support clinical decision making in specific patient groups in hospital settings or for research purposes (e.g.,<sup>[30–34]</sup>). These measures typically quantify the degree of MM in an individual by summing the number of concurrent diseases. Given the limitations of this approach, these measures have been superseded in part by the use of indices. Indices use weights to differentially assess each disease or condition, for instance, in terms of the physician's judgement of likely severity or resource utilization<sup>[35–39]</sup> or prognosis.<sup>[40]</sup> While this approach is considered useful for informing the planning and prioritization of treatment,<sup>[41]</sup> a complex case can render assessment particularly difficult.

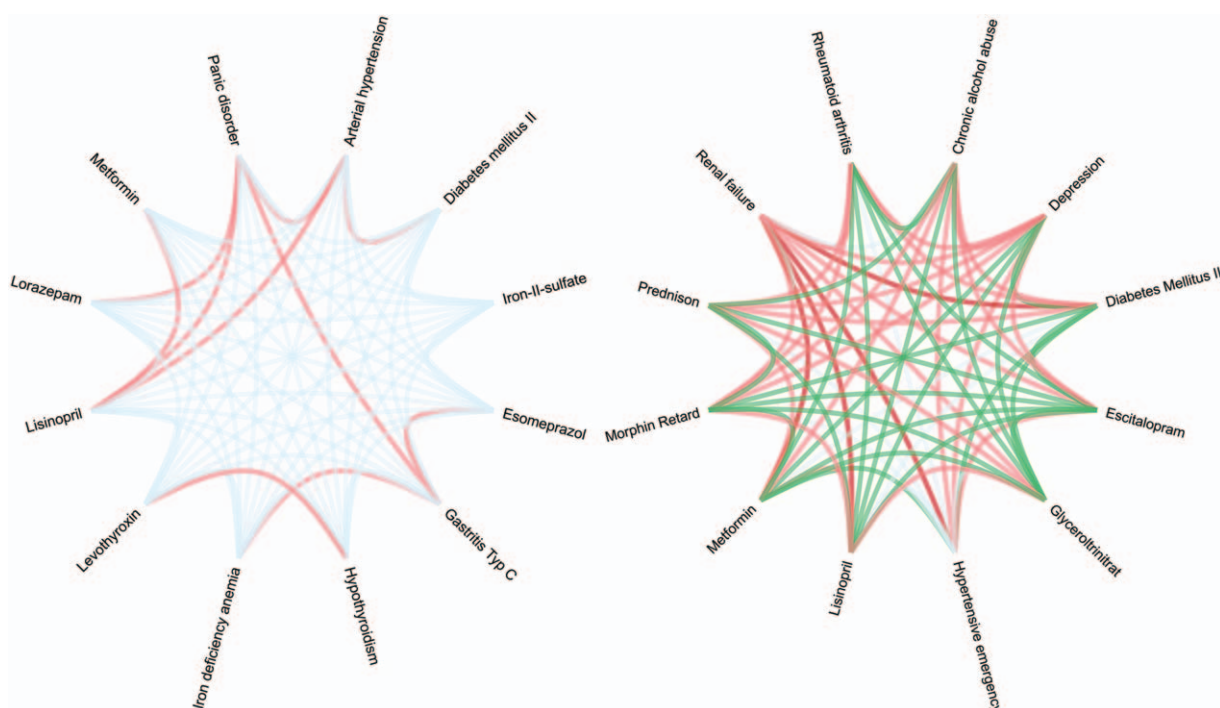
We developed therefore a web-based decision support tool for use at patient encounter to facilitate assessment of complex cases and therapeutic decision making. The tool has 3 main features. First, it generates a case-specific multimorbidity interaction matrix

of all potential DDIs from the patient record. The matrix is used to score each DDI in terms of likely resource utilization (i.e., expected intensity of effort needed to manage the interaction) and is intended to aid the physician's consideration of harmful DDIs and points of caution, uncertainty, and priority in the clinical management of the specific case. Second, the tool generates a network graph from these scores to help the physician visualize his or her assessment of the case (see Fig. 1). Third, the tool generates an overall composite score, or multimorbidity interaction severity index (MISI), of the case as a summary evaluation of the interaction severity of the case on a comparative scale. This preliminary study used the MISI as a basis for the initial evaluation of the tool. We tested the expectation that the MISI could be used to reliably differentiate 2 hypothetical multimorbid patient cases in terms of potential lower versus higher therapeutic conflict. Each case comprised a cluster of 6 similar concurrent conditions, in 1 case with a low and in the other a high potential for harmful interactions and therapeutic conflict.

## 2. Methods

### 2.1. Participants

Eighteen senior physicians (8 females, mean years of clinical experience as General Internists = 12.94, *SD* = 6) volunteered to participate. All participants were senior General Internists and staff members of the Department of Internal Medicine (General Internal Medicine), University Hospital of Zurich, Switzerland and native or fluent speakers of Swiss German (or Standard German) and fluent speakers of English. This pilot study was designed to explore the new index of MM severity on the basis of different patterns of responses of the participants in judging the severity of patients DDIs in terms of expected resource utilization. In the absence of comparable data for the application of this



**Figure 1.** Network graph generated from the multimorbidity interaction matrix on the basis of the participants' ratings of interaction severity, showing (A) the graph for the low conflict case with little risk of harmful interactions, and (B) the graph for the high conflict case with a high risk of harmful interactions.

construct to all combinations of potential DDIs, the sample size was selected on the basis of the experience gained developing the MISI. Sample size calculations for future development of the MISI can be based on estimates derived from this study. Local ethics committee approval was not required for the purpose of our sample of physicians evaluating hypothetical patients. Written informed consent was obtained according to the guidelines of the Declaration of Helsinki, data coded anonymously, and each physician debriefed at the end of the study. The data used and analyzed for the current study are available from the corresponding author on reasonable request.

## 2.2. Materials

**2.2.1. Hypothetical patients.** Authors EB and DG generated 2 hypothetical clinically plausible patient cases, a low and a high conflict case, each with 6 concurrent morbidities. In the low conflict case, the selection of concurrent conditions and treatments was deemed to have a low risk of harmful interactive effects and therapeutic conflict, whereas in the high conflict case these were considered to carry a high risk of harmful interactive effects and therapeutic conflict (see Table 1). The data of diagnosed conditions and medications of the hypothetical cases were imported into the web-based MISI.

**2.2.2. Technical description of the web-based MISI.** The MISI is integrated into a web-based decision support tool. The tool comprises a platform-independent, web browser-based system, with a graphical user interface and a dedicated server component. It is built on open source technology, including node.js and MongoDB for server-side application (express.js, passport.js, mongoose.js, socket.io) and HTML, CSS, and various JavaScript libraries (angular.js, d3.js, and Twitter Bootstrap) for the browser client.

## 2.3. Study procedure

All participants were tested individually in a small quiet room, located at the University Hospital of Zurich, and were blind to the design of the hypothetical cases. The experiment lasted approximately 35 min. The 2 cases were presented across participants in counter-balanced order. The experimental setup

included a 13-inch MacBook Pro with a Retina display and an optical mouse. The server of the web-based tool operated on the machine locally, with the webapp displayed in Firefox v42.0 in full-screen mode. Participants provided demographic information (e.g., age, clinical experience as General Internists). The web-based MISI was then presented to the participants. It was explained how to use the MISI and ensured that the explanation had been understood.

Each participant subsequently read the written instructions presented on the PC monitor as to how to proceed. First, and using the web-based MISI, a panel displayed information on the hypothetical case (e.g., the reason for the encounter, relevant findings in clinical examination, further relevant findings in diagnostic tests, medical history). Second, the hypothetical list of conditions and medications was displayed and the participants were required to rate each condition as either active or inactive and each medication as either modified or unmodified. The web-based MISI provides a definition of active and modified: Active is any medical condition that requires current and ongoing diagnostic or therapeutic (pharmacological or other) attention or measures, and modified is any medication or other therapeutic measure that needs to be altered in content or dose. These and all further ratings were conducted by mouse click on the checkboxes of the MISI's interface. Third, upon completing the last step, the MISI automatically generated a panel displaying a multimorbidity interaction matrix of all possible interaction pairs extracted from the list of active conditions and medications. Participants were required to rate each interaction pair on a qualitative 4-point severity scale of resource utilization (i.e., expected intensity of effort needed to manage the interaction). The scale ranges from "not harmful" (requiring therefore no action) to "life threatening" (requiring therefore severe action), as shown in Table 2. The severity scale also included a further option "don't know" to indicate that the participant is unable to judge whether a potential interaction might be harmful.

Finally, after completing the multimorbidity interaction matrix, the composite MM severity score was computed automatically and displayed above the matrix. The composite score of the MISI is computed by summing up the total number of conditions, medications, active conditions, modified medications, and the number of all interactions rated in terms of severity as minor, as major (multiplied by 3), and as life threatening (multiplied by 10) (see Table 3). Upon completion of the test, the participant was debriefed and asked to report any difficulties experienced using the web-based MISI.

To aid decision-making, the web-based decision support tool also generates a visually intuitive network graph to represent the internist's judgements of interaction severity, as rated in the multimorbidity interaction matrix (see Fig. 1). The graph provides a visual impression of the overall MM severity of the case. The aim of this visualization is to support the physician by highlighting nontrivial relationships that might otherwise be overseen in the complex patient by showing the relevant interactions between the conditions and treatments weighted according to the internist's judgements of severity. MM severity of each case can be visualized in network graph generated from the multimorbidity interaction matrix (see Fig. 1).

**Table 1**  
**Conditions and medications of 2 hypothetical patient cases.**

Hypothetical patients	Conditions	Medications
High conflict case	Hypertensive emergency	Lisinopril
	Renal failure	Nitroglycerin
	Diabetes Mellitus type II	Metformin
	Depression	Escitalopram Prednisone
	Rheumatoid arthritis	Morphine
	Chronic alcohol abuse	
Low conflict case	Iron deficiency anemia	Ferrous sulfate
	Diabetes mellitus type II	Metformin
	Panic disorder	Lorazepam
	Arterial hypertension	Lisinopril
	Gastritis Type C	Esomeprazole
	Hypothyroidism	Levothyroxine

The conditions in the left column are shown in relation to the corresponding medications to treat the conditions in the right column. The interactions in the low conflict case were designed to have a low risk of harmful interactive effects and those of the high conflict case to have a high risk of harmful interactive effects.

## 3. Results

All data analyses were performed using SPSS version 21.0 (SPSS Inc, Chicago, IL). There were no missing data.

**Table 2**

The table shows the 4-point severity rating scale, with a corresponding description and example (condition and medications) for each level of severity.

Severity rating scale	Description	Example: Oral anticoagulation drugs and git-bleeding
Don't know	Cannot be judged due to a lack of knowledge	Gastritis 5 years ago and ongoing oral anti- coagulations drugs
Not harmful	Requires no action	Gastroenteritis with stomach ache and oral anti- coagulation drugs
Minor	Requires moderate action	Duodenal angiodysplasia and ongoing oral anti- coagulation drugs
Major	Requires intense action	Gastric ulcer with a relevant anaemia and ongoing oral anticoagulation drugs
Life threatening	Requires full action	Bleeding from varices and ongoing oral anti-coagulation drugs

The scale ranges from "not harmful" to "life threatening" interaction. The option "don't know" is also given to indicate lack of appropriate knowledge to judge the potential interaction.

One might anticipate that the high conflict case demands more careful consideration in making judgements about potential harmful interactions and resource allocation than does the low conflict case and that greater consideration is reflected in the time taken to judge each case. In fact, the mean time taken by the internists to score the interactions in the MISI's multimorbidity interaction matrix was  $M=11.58$  min ( $SD=4.21$ ; range = 3.25–21.48) for the low conflict case and  $M=15.76$  min ( $SD=4.99$ ; range = 5.27–24.37) for the high conflict case. A paired  $t$  test showed that the high conflict case took significantly longer to score than the low conflict case,  $t_{(17)}=4.53$ ,  $P<0.001$ .

Each patient case was designed to have 6 generally relevant medical conditions and medications. Before using the multimorbidity interaction matrix to rate each potential DDI in terms of likely resource allocation, the internists were first required to judge whether each condition was currently active and whether the medication used needed to be modified. Before testing for differences between the 2 hypothetical cases, the Kolmogorov–Smirnov test for normality was conducted.<sup>[42]</sup> This indicated that the distribution of active condition judgements deviated significantly from a normal distribution in the low ( $D=.298$ ,  $P<0.01$ ) and high conflict case ( $D=.274$ ,  $P<0.01$ ). A Wilcoxon Signed-ranks test was therefore applied,<sup>[43]</sup> this showing no significant difference between the high ( $Mdn=5$ ) and low conflict case ( $Mdn=6$ ) in terms of active judgements,  $Z=-0.78$ ,  $p=0.44$ . The Kolmogorov–Smirnov test for normality indicated that the distribution of modified medication judgements also deviated significantly from a normal distribution for the low ( $D=0.226$ ,  $P=.02$ ) but not for the high conflict case ( $D=0.177$ ,  $P=0.14$ ). A Wilcoxon Signed-ranks test was therefore performed. This showed that the high conflict case ( $Mdn=3.5$ ) was judged as requiring modification significantly more often than the low conflict case ( $Mdn=1$ ),  $Z=-2.37$ ,  $P=0.018$ .

In other words, the preceding analyses show that the internists judged no difference between the 2 cases in terms of the number

of active concurrent conditions. But they did judge a difference in terms of the number of required medication modifications. Given the same number of conditions in each case, the results suggest that the cumulative number of concurrent conditions or conditions (i.e., the oft used definition of MM) does not necessarily reflect the clinical resource effort needed to manage (i.e., in this case, to modify medications) in a multimorbid patient.

We then evaluated whether the MISI could be used to reliably differentiate between the 2 patient cases in terms of lower versus higher therapeutic conflict. To do this, the difference between each physician's composite MM severity score,  $MISI_{diff}$ , was computed and used for further analyses.

The descriptive data for  $MISI_{diff}$  largely show positive values ( $Min=-5$ ,  $1^{st}Quartile=16.75$ ,  $Mdn=32.50$ ,  $M=30.44$ ,  $3^{rd}Quartile=47.75$ ,  $Max=64.00$ ). This means that most raters gave the low conflict case a lower score than the high conflict case. In fact, the mean 50% of internists produced a  $MISI_{diff}$  value between 16.75 and 47.75 between the 2 cases. The positive values of the inter-quartile-range ( $IQR=31$ ) and the standard deviation ( $SD=19.84$ ) for  $MISI_{diff}$ , used as measures of variation (or disagreement) between the 2 cases, suggest overall consistency of the individual physicians in their judgments. However, the minimum value of  $-5$  indicates that at least 1 physician rated the low conflict case as having slightly greater MM severity than the high conflict case.

For illustrative purposes,  $MISI_{diff}$  can be shown in a scatter plot based on the Bland–Altman (B&A) approach<sup>[43]</sup> (see Fig. 2). The scatter plot allows the pattern of MISI data to be inspected: each point in the plot shows the mean of each physician's MM severity scores for cases 1 and 2 on the X-axis and each physician's  $MISI_{diff}$  for cases 1 and 2 on the Y-axis. The overall mean of  $MISI_{diff}$  across all physicians is  $M=30.44$  ( $SD=19.83$ ). Closer inspection of Fig. 2 reveals that 2 of the 18 physicians show differences in their MM severity scores of close to zero, meaning

**Table 3**

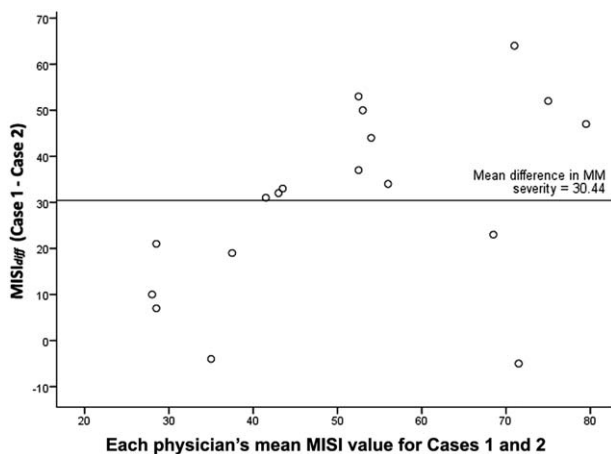
The table shows an example of the composition of the MISI on the basis of one of the participant's ratings of a low and a high conflict case.

	MISI composition	Low conflict case	High conflict case
General relevance	Number of conditions	6	6
	Number of medications	6	6
Current relevance	Active conditions	5	4
	Modified medications	0	5
Harmful DDIs*	Severity ratings of individual DDIs	10	38
MISI	Overall impact of DDIs	27	59

The MISI is computed by summing up the total number of conditions, medications, active conditions, modified medications, and the number of all interactions. The interactions were rated in terms of severity as minor, as major (multiplied by 3), and as life threatening (multiplied by 10). The result in this participant's example is  $MISI=27$  for the low conflict case (with little risk of harmful interactions) and  $MISI=59$  for the high conflict case (with a high risk of harmful interactions).

\* DDI = disease–disease, drug–disease, and drug–drug interactions. MISI = multimorbidity interaction severity index.





**Figure 2.** The scatter plot (based on the Bland–Altman approach) depicts the mean of a particular physician's 2 MM severity scores for cases 1 and 2 on the X-axis and the difference between that physician's 2 MM severity scores for cases 1 and 2 (i.e., high conflict case minus the low conflict case, or  $MISI_{diff}$ ) on the Y-axis. The mean overall mean difference between cases 1 and 2 across all physicians is  $M=30.44$  ( $SD=19.83$ ).

that these 2 judged the MM severity of the 2 cases as being highly similar.

The Kolmogorov–Smirnov test for normality indicated that the  $MISI_{diff}$  distribution did not deviate significantly from a normal distribution ( $D=0.122$ ,  $P=0.200$ ). A one sample  $t$  test was conducted to test the null-hypothesis that  $MISI_{diff}$  is zero, that is, the physicians judged the 2 cases similarly in MM severity. This test showed that the difference between the  $MISI_{diff}$  values ( $M=30.44$ ,  $SD=19.83$ ) and zero was highly significant,  $t_{(17)}=6.51$ ,  $P<0.001$ , 95% CI 20.58, 40.31. This indicates that the physicians distinguished between each case in terms of lower and higher MM severity and therapeutic conflict.

#### 4. Discussion

As a proof of concept, the present study evaluated whether the MISI could be used to reliably distinguish between 2 clinically relevant but hypothetical patient cases in terms of the intensity of effort (i.e., resource utilization) required to clinically manage harmful interactions. The 2 patients were designed to have either a low or high risk of harmful interactive effects with a corresponding degree of therapeutic conflict (see Table 1). In the high conflict case, the design included hypertensive emergency, renal failure, and diabetes mellitus type II as a typical cluster of conditions in MM.<sup>[44]</sup> Depression, often associated with noncompliance, may cooccur with these conditions, and noncompliance may be associated with hypertensive emergencies.<sup>[44–46]</sup> Diseases such as rheumatoid arthritis can cooccur with the other conditions, though not necessarily in relation to the typical cluster.<sup>[47]</sup> In the low conflict case, these prevalent diseases (iron deficiency anemia, diabetes mellitus type II, panic disorder, arterial hypertension, gastritis type C) often cooccur, with the exception of hypothyroidism, but not as a typical MM cluster.<sup>[48–50]</sup> Consistent with their design, the analyses showed that the low conflict case was judged to be significantly less MM severity than the high conflict case. The data thus support the use of the MISI as a means to distinguishing between different patient cases in terms of the severity of harmful interactions and therapeutic conflict.

This web-based tool is novel both as an instrument and for physicians who have never used a multimorbidity severity matrix to judge harmful interactions, as in our case. In this sense, the physicians were also under investigation. As a precursor to a conventional inter-rater agreement analysis, in which we would have a larger range of hypothetical or real cases (e.g.,<sup>[51]</sup>), this preliminary study was designed to give insight into the pattern of MISI data generated by the physicians while applying careful experimental control to the 2 cases. The general difference in the pattern of MISI data between the cases suggests that the physicians similarly understood and used the multimorbidity severity matrix to judge expected resource utilization. While this suggestion is supported by the physicians' feedback at debriefing, we are mindful of potential sources of error in our new approach. We consider therefore the present study in terms of a number of factors that require more attention before conducting a conventional study of inter-rater agreement.

This tool is intended for use as a brief instrument to facilitate assessment of complex cases and illness severity. The time taken for our experienced general internists (who received only a brief introduction to the tool) to complete their task averaged 12 minutes for the low conflict case and 16 minutes for the high conflict case. This difference likely reflects a greater level of difficulty in assessing the high conflict case, but there was large variability between internists. This ranged from 3 to 24 minutes. This variability might suggest potential for improving the time to administer the instrument. Multiple factors, including uncertainties in MM and differences in internists' knowledge and experience, influence the process of evaluating the complex multimorbid patient and this is likely to reveal differences between clinical decision makers<sup>[52,53]</sup> in terms of completion time and the absolute values of the MISI.<sup>[54–56]</sup>

While the internists were all highly experienced in MM, they were not trained in or otherwise accustomed to using a multimorbidity interaction matrix to score harmful interactions. Training would help to ensure a common understanding of the matrix and its scoring procedure and enhance agreement between the physicians.<sup>[57,58]</sup> To develop a common understanding (referred to in rater training as a frame of reference, e.g.,<sup>[59]</sup>) potential sources of disagreement between users of this tool need to be identified.<sup>[60]</sup> A primary consideration in the present study is the extent to which the descriptors used to characterize the severity scale (see Table 2) are open to discrepant interpretation by different internists. Severity was operationally defined in terms of internists' subjective judgements of projected resource utilization (i.e., intensity of effort) to manage harmful interactions. At debriefing, this definition was not reported as presenting any difficulties, and the inter-quartile-range, used as a measure of variation (or disagreement) between the 2 cases, suggests general consistency (albeit with the exception of 3 internists) in the use of the severity scales across the 2 cases (cf.<sup>[61]</sup>).

It is possible that good consistency (or, in a study with more than 2 patient cases, good interrater agreement) within a single-center sample reflects a good degree of shared knowledge of and routine in applying available center-specific resource and care management procedures. Any differences across centers in resources and procedures for MM in general or for specific clusters of MM might have a different influence on physicians' judgements of resource utilization for the treatment of MM patients. This might result therefore in some degree of systematic variation in inter-rater agreement across centers. As a broadly applicable measure, the MISI might provide a means to

evaluating this possibility. But the present study revealed that 2 of the physicians judged the low conflict case as slightly more severe than the high conflict case. Besides training, knowledge, and experience, testing is needed to establish what other (e.g., motivational) factors might influence the use of the tool for judging MM severity.

These factors might include experimental considerations. Every effort was made during recruitment and at instruction to ensure that participants were not aware of any potential difference in the severity of the hypothetical cases. Once physicians began using the multimorbidity severity matrix to rate the expected resource utilization of potentially harmful interactions, the physician may have tried to infer from those ratings the general severity of the case. The motivation behind using a multimorbidity severity matrix is that it might make it easier to build up a picture of a more complex case. This might be particularly helpful for less experienced physicians. But from an experimental perspective, we cannot exclude the possibility that a physician's ratings influenced subsequent ratings in the matrix, or that the severity ratings were biased by the physician's anticipation of the general severity of the case before completing the matrix. While we ensured that conditions and medications were always presented in the same sequence across all participants, future experiments could randomize the order of entries in the matrix to test for such bias. On the other hand, the network graph and the final index were only accessible to the physician after the completion of the matrix and did not influence the physician's ratings.

The strengths and weaknesses of the web-based tool may be considered in relation to its main features. The multimorbidity interaction matrix may help the physician to build up a cohesive and systematic assessment of potentially harmful DDIs. This can be important because, unlike the well-structured reporting of diagnoses and conditions, interactions are often described, sometimes lengthily, in free text. Given the potential complexity of many cases, the matrix might be especially helpful for less experienced physicians. To ensure the relevance of the judgements of resource allocation as a basis for using the matrix and determining severity of DDIs, these judgements need to be evaluated in relation to outcome measures such as hospital stay, morbidity, mortality, and resource use. A major weakness of the present matrix is that it requires manual work. Future development aims to simplify its use and usability by, for instance, integrating a database query of DDIs and data from an expert panel to automatically highlight likely and relevant DDIs for a given MM cluster. A limitation at present is that all diagnoses in a database query of DDIs must be encoded or structured in a way that can be further processed by the web-based tool. This is not yet the case.

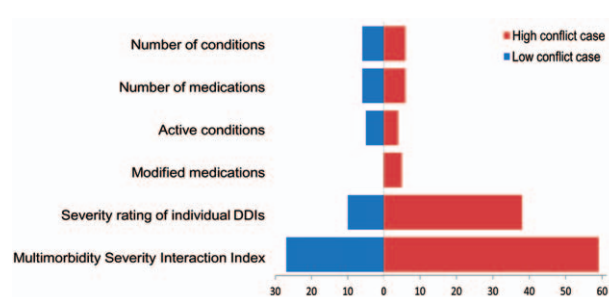
The main strength of the network graph is that it allows the user to visualize severity of interactions at one glance. Especially in a complex case, a simple visual overview of the pattern of DDIs might reduce the cognitive load of maintaining a mental representation of DDIs (see Fig. 1B).<sup>[62]</sup> A further advantage is that the visual representation highlights the interactions according to the physician's own judgements of severity, making the network graph immediately relevant for the physician and the case. Potentially, the graph could augment the reporting of complex cases as a means to gaining a quick overview of the DDIs in a case and by highlighting potentially critical points of convergence between conditions and medications that require might particular attention. The graph cannot replace clinical contact to the patient. Empirical evaluation of the usability and

effectiveness of the network graph for less and more experienced internists as a simple aid to representing, communicating and considering a complex case awaits investigation.

As a summary measure, the MISI may be used to gauge and communicate the general severity of a case. Importantly, the MISI provides a summary evaluation of the harmful interactions at a given point in time. As a momentary assessment, the value of the MISI will change over the course of treatment. With further development, the use of the MISI as a basis for measuring change in interaction severity over time will be tested. One priority for further development is to first ensure the reliability of this index. The index might then be used as a reference for typical MM clusters (and combinations of harmful DDIs), as established for example by an expert panel. The advantage of this is that a physician could judge his or her assessment of case severity against the reference. While the MISI is conceptualized as a comparative scale, empirically based thresholds of higher, medium, and lower risk are needed. These thresholds are likely to be time-consuming and costly to develop, but they would foster for purpose of research and clinical assessment consistent and meaningful interpretation within particular and across different MM clusters.

To consider potential strengths and limitations from the perspective of the user, we collected feedback from the physicians at debriefing. This supported the view that the matrix evoked greater reflection about and awareness of potential DDIs and therapeutic conflict. The network graph was not investigated, but the feedback supported the idea that the visualization could help to make the user more aware of these interactions and conflicts. On the other hand, the use of the matrix was reported by some as requiring a high level of attention. The task of explicitly judging the severity of DDIs generated a sense of uncertainty in some raters, in part because it revealed to the rater imperfect knowledge: internet links to an interaction checker were suggested. Clearly, the task was made more difficult in this study by the fact that anamnestic, clinical, and laboratory data were not provided.

The data of this study show that quantifying the degree of MM simply on the basis of the cumulative number of concurrent conditions does not necessarily reflect the clinical effort needed to prioritize, plan, and manage harmful interactions. As indicated in Table 3 and Fig. 3, the total number of a patient's conditions and medications can deviate from the number of conditions and medications that the physician considered to be relevant for current treatment. While each patient case was designed to have 6 concurrent conditions, the analyses showed a significant



**Figure 3.** The figure illustrates the composition of the MISI, based on one physician's ratings of a low conflict case (right side of main panel) and a high conflict case (left side of main panel).

difference between the low conflict and high conflict cases in terms of the judged need to modify the patient's medications; the high conflict case requiring a significantly greater number of modifications. The physician must also prioritize the clinical effort needed to manage and monitor more harmful interactions. The computation of the MISI considers, therefore, the cumulative number of concurrent conditions and medications, while weighing the index to take account of more severe DDIs that the physician considers to require more urgent attention. Further work is needed to ensure that the judgments of anticipated resource utilization and our weighting procedure reasonably reflect real resource utilization.

Other indicators of MM severity have been used.<sup>[40]</sup> These range from composite measures based on multiple dimensions (e.g., prognostic threat to life, number of organs affected, disability, complications, and seriousness of treatment) to more simple uni-dimensional measures of severity.<sup>[63,64]</sup> The MISI could be easily adapted to generate a composite score based on 1 or more other such dimensions. But the present definition of severity, in terms of resource utilization, serves to focus attention on the relationship between the use of health care resources and DDIs<sup>[42]</sup> by considering the immediate relevance of the variously harmful interactions for clinical prioritization, planning, and management.<sup>[65]</sup> This is important considering the impact of MM on health service utilization and costs.<sup>[66–68]</sup> These costs increase exponentially with the increasing number of chronic diseases.<sup>[69]</sup>

## 5. Conclusion

The present study demonstrated that the MISI can be used as a measure of subjective judgement of harmful interactions to reliably distinguish between 2 hypothetical patients of potentially low and high therapeutic conflict. The judgement of harmful interaction is based on the internist's subjective assessment of the intensity of effort (i.e., resource utilization) required to clinically manage the harmful interactions. In conducting this proof of concept study, we made the assumption that the main source of systematic variation in the subjective judgements of resource allocation would be attributable to the differences in DDIs between the low and high conflict cases. But the impact of the internists' expertise in treating MM and the validity of the subjective construct "resource utilization" itself needs to be subject to testing. Importantly, the MISI needs to be evaluated for a larger number of patients and for a greater range of variously severe MM. In supporting clinical decision making, the patient-specific MISI focuses attention on the patient as a whole, rather than on any 1 multimorbid condition, by highlighting the interacting conditions and treatments (cf.<sup>[70]</sup>). The visually intuitive network graph is designed to help the clinical decision-maker visualize his or her assessment of the case. The potential usability and effectiveness of the graph as an aid to representing, communicating, and considering a complex case awaits evaluation. The digital format of the MISI means that this instrument is practical to administer and easy to adapt for application in clinical practice and research.

## Acknowledgments

The authors thank Professor Beate Sick, Zurich University of Applied Science, Winterthur, Switzerland, for consultation regarding the statistical approach applied in this study.

## References

- van den Akker M, Buntinx F, Knottnerus J. Comorbidity or multimorbidity: what's in a name. A review of literature. *Eur J Gen Pract* 1996;2: 65–70.
- Bayliss EA, Edwards AE, Steiner JF, et al. Processes of care desired by elderly patients with multimorbidities. *Fam Pract* 2008;25:287–93.
- Marengoni A, Angleman S, Melis R, et al. Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev* 2011;10:430–9.
- WHO: The European health report 2009: health and health systems. Copenhagen: World Health Organisation; 2009.
- Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med* 2007;22(suppl 3):391–5.
- Glynn LG, Valderas JM, Healy P, et al. The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Fam Pract* 2011;28:516–23.
- Markun S, Holzer BM, Rodak R, et al. Therapeutic conflicts in emergency department patients with multimorbidity: a cross-sectional study. *PLoS One* 2014;9:e110309.
- Gutermann IK, Niggemeier V, Zimmerli LU, et al. Gastrointestinal bleeding and anticoagulant or antiplatelet drugs: systematic search for clinical practice guidelines. *Medicine* 2015;94:e377.
- Guthrie B, Payne K, Alderson P, et al. Adapting clinical guidelines to take account of multimorbidity. *BMJ* 2012;345:e6341.
- Blazik E, van den Busche H, Gurtner F, et al. Epidemiological strategies for adapting clinical practice guidelines to the needs of multimorbid patients. *BMC Health Serv Res* 2013;13:352.
- Fabbri LM, Boyd C, Boschetto P, et al. How to integrate multiple comorbidities in guideline development: article 10 in integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proc Am Thorac Soc* 2012;9: 274–281.
- Uhlig K, Leff B, Kent D, et al. A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med* 2014;29: 670–679.
- Gross C, Mallory R, Heiat A, et al. Reporting the recruitment process in clinical trials: who are these patients and how did they get there? *Ann Intern Med* 2002;137:10–6.
- Boyd CM, Vollenweider D, Puhon MA. Informing evidence-based decision-making for patients with comorbidity: availability of necessary information in clinical trials for chronic diseases. *PLoS One* 2012;7:e41601.
- Rothwell P. External validity of randomised controlled trials: "to whom do the results of this trial apply? *Lancet* 2005;365:82–93.
- Van Spall HG, Toren A, Kiss A, et al. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007;297:1233–40.
- Barnett K, Mercer S, Norbury M, et al. Epidemiology of multimorbidity and implications for healthcare, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37–43.
- Guthrie B, McCowan C, Davey P, et al. High risk prescribing in primary care patients particularly vulnerable to adverse drug events: cross sectional population database analysis in Scottish general practice. *BMJ* 2012;342: d3514.
- Bourgeois FT, Shannon MW, Valim C, et al. Adverse drug events in the outpatient setting: an 11-year national analysis. *Pharmacoepidemiol Drug Safety* 2010;19:901–10.
- Boyd C, Darer J, Boult C, et al. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases implications for pay for performance. *JAMA* 2005;294:716–24.
- Tinetti ME, Bogardus ST Jr, Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *N Engl J Med* 2004;351:2870–4.
- Muth C, van den Akker M, Blom JW, et al. The Ariadne principles: how to handle multimorbidity in primary care consultations. *BMC Med* 2014;12:223.
- Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. *N Engl J Med* 1998;338:1516–20.
- Markun S, Holzer BM, Rodak R, et al. Therapeutic conflicts in emergency department patients with multimorbidity: a cross-sectional study. *PLoS One* 2014;9: e110309.
- Abernethy AP, Currow DC. Safety of benzodiazepines and opioids in very severe respiratory disease: national prospective study. *BMJ* 2014;348:g445.
- Starfield B. Threads and yarns: weaving the tapestry of comorbidity. *Ann Fam Med* 2006;4:101–3.
- Luijckx HD, Loeffen MJ, Lagro-Janssen AL, et al. GPs' considerations in multimorbidity management: a qualitative study. *Br J Gen Pract* 2012;62:e503–10.

- [28] Blozik E, van den Busche H, Gurtner F, et al. Epidemiological strategies for adapting clinical practice guidelines to the needs of multimorbid patients. *BMC Health Serv Res* 2013;13:352.
- [29] Boyd CM, Weiss CO, Halter J, et al. Framework for evaluating disease severity measures in older adults with comorbidity. *J Gerontol A Biol Sci Med Sci* 2007;62:286–95.
- [30] de Groot V, Beckerman H, Lankhorst GJ, et al. How to measure comorbidity: a critical review of available methods. *Clin Epidemiol* 2003;56:221–9.
- [31] Charlson ME, Sax FL, MacKenzie CR, et al. Assessing illness severity: does clinical judgement work? *J Chron Dis* 1986;39:439–52.
- [32] Fortin M, Bravo G, Hudon C, et al. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med* 2005;3:223–8.
- [33] van den Akker M, Buntinz F, Metsemakers JF, et al. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol* 1998;51:367–75.
- [34] Salisbury C, Johnson L, Purdy S, et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;61:e12–21.
- [35] Huntley A, Johnson R, Purdy S, et al. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med* 2012;10:134–41.
- [36] Valderas JM, Starfield B, Sibbald B, et al. Defining comorbidity: implications for understanding health and health services. *Ann Fam Med* 2009;7:357–63.
- [37] Harboun M, Ankri J. Comorbidity indexes: review of the literature and application to studies of elderly population. *Revue Epidemiol Sante Publique* 2009;49:287–98.
- [38] de Groot V, Beckerman H, Lankhorst GJ, et al. How to measure comorbidity: a critical review of available methods. *J Clin Epidemiol* 2003;56:221–9.
- [39] Groll DL, To T, Bombardier C, et al. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol* 2005;58:595–602.
- [40] Charlson ME, Pompe P, Ale KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- [41] Fortin M, Contant E, Savard C, et al. Canadian guidelines for clinical practice: an analysis of their quality and relevance to the care of adults with comorbidity. *BMC Family Practice* 2011;12:74.
- [42] Conover WJ. *Practical Nonparametric Statistics*. 3rd ed. New York: John Wiley & Sons; 1999:428–433.
- [43] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;327:307–310.
- [44] Schäfer I, Kaduszkiewicz H, Wagner HO, et al. Reducing complexity: a visualisation of multimorbidity by combining disease clusters and triads. *BMC Public Health* 2014;14:1285.
- [45] DiMatteo MR, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. *Arch Intern Med* 2000;160:2101–7.
- [46] Saguner AM, Dür S, Perrig M, et al. Risk factors promoting hypertensive crises: evidence from a longitudinal study. *Am J Hypertens* 2010;23:775–80.
- [47] Dickens C, McGowan L, Clark-Carter D, et al. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. *Psychosom Med* 2002;64:52–60.
- [48] Schäfer I, von Leitner E-C, Schön G, et al. Multimorbidity patterns in the elderly: a new approach of disease clustering identifies complex interrelations between chronic conditions. *PLoS One* 2010;5:e15941.
- [49] Vanderpump MP. The epidemiology of thyroid disease. *Br Med Bull* 2011;99:39–51.
- [50] Kessler RC, Chiu WT, Demler O, et al. Arch gen psychiatry. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication 2005;62:617–27.
- [51] Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quantit Methods Psychol* 2012;8:23–34.
- [52] Smith M, Higgs J, Ellis E. *Clinical Reasoning in the Health Professions*. 3rd ed. Amsterdam: Butterworth; 2008.
- [53] Hajjaj F, Salek M, Basra M, et al. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. *J R Soc Med* 2010;103:178–87.
- [54] Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective in medical expertise: theory and implications. *Acad Med* 1990;65:611–21.
- [55] Patel VL, Groen GJ. Developmental accounts in the transition from medical student to doctor: some problems and suggestions. *Med Educ* 1991;2:526–35.
- [56] Wainwright SF, Shepard KF, Harman LB, et al. Novice and experienced physical therapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. *Phys Ther* 2010;90:75–88.
- [57] Gorman CA, Rentsch JR. Evaluating frame-of-reference rater training effectiveness using performance schema accuracy. *J Appl Psychol* 2009;94:1336–44.
- [58] Davies M, Fleiss JL. Measuring agreement for multinomial data. *Biometrics* 1982;38:1047–51.
- [59] Woehr DJ, Huffcutt AI. Rater training for performance appraisal: a quantitative review. *J Occup Organizational Psychol* 1004;67:189–205.
- [60] Hoyt WT, Kerns M. Magnitude and moderators of bias in observer ratings: a meta-analysis. *Psychol Methods* 1999;4:403–24.
- [61] Davies M, Fleiss JL. Measuring agreement for multinomial data. *Biometrics*. 1982;38:1047–1051.
- [62] Anderson EW, Potter KC, Matzen LE, et al. A user study of visualization effectiveness using EEG and cognitive load. *Comput Graph Forum*. 2011;30:791–800.
- [63] Barsky AJ, Wyshak G, Klerman GL. Medical and psychiatric determinants of outpatient medical utilization. *Med Care* 1986;24:548–60.
- [64] Perkins AJ, Kroenke K, Unützer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol* 2004;57:1040–8.
- [65] Feinstein AR. Pre-therapeutic classification of co-morbidity in chronic disease. *J Chronic Dis* 1970;23:455–68.
- [66] Hoffman C, Rice D, Sung HY. Persons with chronic conditions: their prevalence and costs. *JAMA* 1996;276:1473–9.
- [67] Fortin M, Soubhi H, Hudon C, et al. Multimorbidity's many challenges. *BMJ* 2007;334:1016–7.
- [68] Ritchie C. Health care quality and multimorbidity: the jury is still out. *Med Care* 2007;45:477–9.
- [69] Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;162:2269–76.
- [70] Espallargues M, Philp I, Seymour DG, et al. ACMEplus PROJECT TEAM Measuring case-mix and outcome for older people in acute hospital care across Europe: the development and potential of the ACMEplus instrument. *QJM* 2008;101:99–109.