The multimorbidity interaction severity index (MISI)
A proof of concept study

Dimitri Gassmann, MMeda,b, Marcus Cheetham, PhDa,b,c, Klarissa Siebenhuener, MDa,b, Barbara M. Holzer, PhDa,b, Claudine Meindl-Fridez, MDa,b, Florian F. Hildenbrand, MDa,b, Vanessa Virginio, MDc, Mike Martin, PhDd,e, Edouard Battegay, MD, FACPa,b,c,*

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\text{\textsubscript{diff}}) was computed. Based on MISI\text{\textsubscript{diff}}, 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 multimorb 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.


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

1. Introduction
Multimorbidity (MM) refers to the presence of multiple concurrent acute or chronic diseases within a person. Therapeutic decision making for multimorb patients is challenging. This is because decisions rely on recommendations from clinical practice guidelines that were developed for the treatment of single diseases. With some exceptions, these guidelines are based on evidence from studies that excluded or under-represented multimorbid cases. This monomorbid 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 or adequately guide the clinical decision-maker through the therapeutic conflict.

Therapeutic decision making in MM requires consideration of potential DDIs 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...
or more other conditions or treatments.\textsuperscript{[24]} 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.\textsuperscript{[8,25]} 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.\textsuperscript{[26,27]}

Intensive work to support clinical decision making in MM is in progress (e.g.,\textsuperscript{[9,11,13,28]}). 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.,\textsuperscript{[29]}). 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.,\textsuperscript{[30–34]}). 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\textsuperscript{[35–39]} or prognosis.\textsuperscript{[40]} While this approach is considered useful for informing the planning and prioritization of treatment,\textsuperscript{[41]} 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 DDLs, 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 the design of the hypothetical cases. The experiment lasted

<table>
<thead>
<tr>
<th>Hypothetical patients</th>
<th>Conditions</th>
<th>Medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>High conflict case</td>
<td>Hypertensive emergency</td>
<td>Lisinopril</td>
</tr>
<tr>
<td>Renal failure</td>
<td>Diabetic Mellitus type II</td>
<td>Metformin</td>
</tr>
<tr>
<td>Depression</td>
<td>Rheumatoid arthritis</td>
<td>Morphine</td>
</tr>
<tr>
<td>Chronic alcohol abuse</td>
<td>Iron deficiency anemia</td>
<td>Ferrous sulfate</td>
</tr>
<tr>
<td>Diabetes mellitus type II</td>
<td>Lisinopril</td>
<td></td>
</tr>
<tr>
<td>Panic disorder</td>
<td>Arterial hypertension</td>
<td>Lisinopril</td>
</tr>
<tr>
<td>Gastroesophageal reflux</td>
<td>Esomeprazole</td>
<td></td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>Hypothyroidism</td>
<td>Levothyroxine</td>
</tr>
</tbody>
</table>

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.
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, $t_{177} = 4.53$, $P < 0.001$.

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.\(^4\) 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,\(^4\) 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 ($D = 0.177$, $P = 0.34$). 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 ($Min = −5$, $1^\text{st Quartile} = 16.75$, $Mdn = 32.50$, $M = 30.44$, $3^\text{rd Quartile} = 47.75$, $Max = 64.00$). 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\(^5\) (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 that they did not perceive a conflict in the timings of the cases.

\[\text{Table 2}\]

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

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.

\[\text{Table 3}\]

<table>
<thead>
<tr>
<th>MISI composition</th>
<th>Low conflict case</th>
<th>High conflict case</th>
</tr>
</thead>
<tbody>
<tr>
<td>General relevance</td>
<td>Number of conditions</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Number of medications</td>
<td>6</td>
</tr>
<tr>
<td>Current relevance</td>
<td>Active conditions</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Modified medications</td>
<td>0</td>
</tr>
<tr>
<td>Harmful DDIs*</td>
<td>Severity ratings of individual DDIs</td>
<td>10</td>
</tr>
<tr>
<td>MISI</td>
<td>Overall impact of DDIs</td>
<td>27</td>
</tr>
</tbody>
</table>

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).

and higher MM severity and therapeutic consider. Physicians distinguished between each case in terms of lower severity. The test showed that the difference between the MISI severity scores for cases 1 and 2 (i.e., high conflict case minus the low conflict case, or MISIdiff) on the Y-axis. The mean overall mean difference between cases 1 and 2 across all physicians is $M = 30.44$ (SD = 19.83).

![Figure 2.](image-url) 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 MISIdiff) on the Y-axis. 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 multimorbidity patient and this is likely to reveal differences between clinical decision makers[52,53] in terms of completion time and the absolute values of the MISI[54–56].

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.[57,58] To develop a common understanding (referred to in rater training as a frame of reference, e.g.,[60]) potential sources of disagreement between users of this tool need to be identified.[60] 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.[61]).

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
judging MM severity. Or that the severity ratings were biased by the physician's own judgements of severity, making the physicians judge the low conflict case (right side of main panel) and a high conflict case (left side of main panel) and a high conflict case (right side of main panel).

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 MASI 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.[40] 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.[63,64] The MASI could be easily adapted to generate a composite score based on 1 or more 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[42] by considering the immediate relevance of the variously harmful interactions for clinical prioritization, planning, and management.[65] This is important considering the impact of MM on health service utilization and costs.[66–68] These costs increase exponentially with the increasing number of chronic diseases.[69]

5. Conclusion

The present study demonstrated that the MASI can be used as a measure of subjective judgement of harmful interactions to reliably distinguish between 2 hypothetic 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 MASI 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 MASI focuses attention on the patient as a whole, rather than on any 1 multimorbid condition, by highlighting the interacting conditions and treatments (cf.[70]) 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 MASI 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


