Classification of severity of health problems in family/general practice: an international field trial

George R Parkerson Jr*, Charles Bridges-Webb, Juan Gervas, Inge Hofmans-Okkes, Henk Lamberts, Jack Froom, Gisela Fischer, Betty Meyboom-de Jong, Bent Bentsen, Mike Klinkman and Jan de Maeseneer


Background. A methodology is needed for classification of health problems by severity.

Objectives. We aimed to test the Duke Severity of Illness Checklist (DUSOI) for feasibility and usefulness.

Method. The DUSOI was field tested internationally by 22 family/general practitioners in 9 countries.

Results. The DUSOI was found to be feasible for rating severity of illness of health problems in family/general practice. The measure was shown to be clinically useful in older patients and those with chronic and more severe health problems. Variability of severity ratings was less within the same rater than between different raters (i.e. higher intrarater than interrater reliability). Clinical face validity was supported by the finding that DUSOI ratings classified patients with the same diagnosis and those with different diagnoses according to the severity differences that would be expected clinically.

Conclusions. Although research is needed to improve reliability and to test validity further, the DUSOI was shown in the present study to be a methodology that is reasonable for consideration as an international classification of health problems by their severity in primary care patients.

Keywords. Severity of illness, classification of health problems, ICPC.

Introduction

As part of its ongoing effort to improve classification systems for primary care, the Classification Committee of the World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians (WONCA) has field tested a methodology for coding the severity of health problems. Once fully tested and refined, the severity classification system is expected to become part of the next revision of the WONCA International Classification of Primary Care (ICPC).

In recent years, a number of instruments have been developed to measure severity of illness because severity has become recognized as an important factor in health care research and allocation of health service resources. This report presents the results of a field trial that the Classification Committee has sponsored during the past 3 years using one of the severity measures, the Duke Severity of Illness Checklist (DUSOI). The DUSOI differs from the others in that it is based upon the clinical judgement of health care providers and was developed entirely in the primary care setting. The reliability and validity of the DUSOI has been established, and the instrument has been shown to be a useful research tool and a predictor of health-related clinical outcomes. Its feasibility and possible clinical usefulness have been reported. However, the present report is the first to describe the use of the DUSOI internationally.
This study addressed the following questions from an international perspective: (i) is the DUSOI feasible for use in family/general practice? (ii) is the DUSOI useful to clinicians in providing health care for primary care patients? (iii) does the DUSOI provide a methodology by which the family/general practitioner can rate the severity of each of a patient's health problems at the time of the patient encounter? and (iv) can this severity rating be used to classify severity for the same health problem in different patients, and among different health problems in the same patient? A system that is found to fulfill these requirements not only would have great research value, but also would enable the practising doctor to measure severity of illness for all health problems in a standardized way, to distinguish quantitatively the sickest patients from those who are less sick with the same health problem, and to classify health problems more accurately. None of these clinical uses has been demonstrated for the existing classification systems.

Methods

Participants in the study were volunteer family/general practitioners from multiple countries who were English literate. They were either members of the WONCA Classification Committee or recruited by those members. They were trained to perform the DUSOI by one of the authors (GRP), entirely by correspondence. In the training process, each participant received an instruction manual written in English and used the DUSOI to rate severity of illness on each health problem of two fictitious practice patients who were described in brief vignettes. The completed DUSOI form for the first practice case (a patient with three health problems) was mailed to the author, who critiqued it and returned written constructive feedback to the rater. The same process was followed for the second practice case (a patient with five health problems). Successful completion of the two practice cases with eight health problems was considered to be sufficient for the basic instruction of the raters.

The DUSOI is a measure of severity of illness in which the health care provider rates the severity of each of the patient's health problems at the time of an encounter. Severity is rated along four severity parameters: symptom status, complications, prognosis during the next 6 months without treatment, and treatability or expected response to treatment.

The study was conducted in two phases. In Phase 1 (1993–1994), each participating family/general practitioner completed the DUSOI on 25 patients in her or his practice. Selection of patients was decided by the rater, with the inclusion criteria that each patient had to be at least 17 years of age, and that there were at least five patients with each of five chronic health problems (i.e. chronic ischaemic heart disease, chronic obstructive pulmonary disease (COPD), osteoarthritis, diabetes mellitus type II, or depression). Although patients were selected according to whether or not they had certain specified health problems, the raters listed and rated severity of illness for all of the problems each patient had at the time of the encounter.

In Phase 2 (1994–1995), only those participating family/general practitioners who had completed Phase 1 were included. Their first task was to complete the DUSOI on a series of case vignettes that were different from the original practice cases but similar in format and content. These vignettes were intended to test the reliability of the DUSOI. There were three of these reliability test cases, each with three health problems. They were completed and mailed to the author at Time 1, and again at Time 2, at least 4 weeks later. Unlike the practice cases, no feedback was returned to the raters for the test cases.

The second task for Phase 2 was for each participant to perform the DUSOI on 30 patients in her or his practice. These patients could have any type of health problem except for problems entirely related to health maintenance, procedures, or process of medical care. Patients had to be at least 17 years of age, and they had to be selected from among the first consecutive eligible five patients who were seen during each of six different days over a period no longer than 6 weeks. Patients could be included only once during the study. As in Phase 1, all health problems of each patient were listed and rated for severity by the participating rater. The DUSOI form was modified for Phase 2 to include the capability for immediate scoring of the severity of each health problem by the practising physician, as shown below under the description of the DUSOI form. Except for this change, the form was identical for the two phases.

Additional data collected on each study patient in both phases were birthdate, gender, time to complete the DUSOI, difficulty completing the DUSOI, and potential clinical usefulness of the DUSOI severity rating in that particular patient. All data forms on patients were mailed to the author and entered into a computerized database. At the end of each phase, a complete summary of the raw data and descriptive analyses of the patients of each participant were sent to the respective participants. At the end of Phase 2, a final evaluation survey questionnaire was sent to all participants to determine their opinion as to the relative value of each of the DUSOI severity parameters, the type of patient in which the rating was most useful, whether or not the DUSOI was thought to be clinically accurate, whether or not immediate scoring of the DUSOI improved its clinical usefulness, how easy it was to complete, and in what way the raters planned personally to use the DUSOI in the future.
The DUSOI form used in Phase 2 of the study is shown in Figure 1, filled in as an example with fictitious data. To complete the form, each health problem for the given patient for the given visit is listed, and then a single digit severity rating on the scale of 0–4 is entered for each of the four severity parameters. The criteria for each level of severity for each severity parameter are displayed for easy reference in the large box at the bottom of the form. In Figure 1, patient number 10 had two health problems on 11 January 1995 (i.e. diabetes mellitus type II and acute appendicitis). The symptoms parameter was rated ‘1’ for diabetes (indicating questionable symptoms caused by diabetes in this patient at this visit), and ‘4’ for acute appendicitis (indicating major symptoms). The ‘0’ for complications for both diagnoses indicates the absence of complications at that time. The ‘2’ rating for prognosis for diabetes indicates that without any treatment for diabetes during the next 6 months this patient would be expected by her doctor to experience moderate disability. On the other hand, the ‘4’ for acute appendicitis indicates that without treatment this patient would be likely to die. In both diagnoses, the ‘2’ rating for treatability indicates that with treatment both problems would be expected to respond well in this particular patient.

For immediate scoring in Phase 2, each rater in the field trial summed the raw severity scores for each health problem, and entered the severity code as determined from the algorithm in the small box at the bottom of the form. Severity codes of 0, 1, 2, 3 and 4 indicate the following classes of severity, respectively: none, mild, intermediate, moderate and maximum. For example, in Figure 1, the total raw score for diabetes was 5, meaning that the severity code was 2 (i.e. since the total raw score was in the range 5–8), indicating that diabetes in this patient would be classified having "intermediate" severity. Likewise, the total raw score of 10 for acute appendicitis results in a severity classification of 3, indicating a “moderate” severity classification. If this patient had had a ruptured appendix, the complications raw score would have been 4, the total raw score would have been 14, and the severity classification code would have been 4 for "maximum" severity. In addition to the immediate scoring method, the DUSOI can be scored on a scale of 0–100 for each health problem. In the example, appendicitis with a raw score of 10 (scale of 0 to 16) would have a diagnosis DUSOI score of 62.5 (scale of 0–100), i.e. 10/16 × 100.

Statistical methods used for the analyses in this study included the chi-square for categorical data and Student’s t-test for comparison of continuous data. Reliability (i.e. reproducibility or agreement) was tested.
by using intraclass correlation coefficients (ICC) derived from a two-way mixed effects analysis of variance model assuming fixed raters' effects.11,12 The ICC for interrater reliability is the proportion of the total variance in the DUSOI scores that is explained by the variance between the raters' scores across the different health problems, when the total variance includes not only the variance between raters' scores across the different health problems, but also the variance due to differences between the raters' scores for each health problem, and the unexplained variance due to random error. The ICC for intrarater reliability is the proportion of the total variance in the DUSOI scores that is explained by the variance between the raters' scores for each health problem, when the total variance includes not only the variance between the raters' scores for each health problem, but also the variance between the scores at Time 1 and Time 2, and the unexplained variance due to random error. ICCs can range from 0 to 1.0, from lowest to highest, with excellent reliability indicated by ICCs > 0.75, fair to good reliability by ICCs of 0.40–0.75, and poor reliability by ICCs < 0.40.12

Results

Initially, 47 family/general practitioners from 16 different countries volunteered to participate in the field trial and completed at least part of Phase 1. Of these, 30 (63.8%) from 12 countries completed Phase 1, and 22 (46.8%) from 9 countries completed both Phase 1 and Phase 2. The reasons for dropping out of the study included lack of time, lack of sufficient clinical practice to supply the range of patients required, unexplained cessation of communication, and, in one instance, sudden death.

The data reported here are those collected by the 22 participants who completed the entire study. Their distribution by countries was: The Netherlands, 5; Spain, 4; Belgium, 3; UK, 3; Hong Kong, 2; USA, 2; Germany, 1; Israel, 1; and Japan, 1. Their gender distribution was 86.4% male, with a mean age of 43.9 ± 9.2 SD years, ranging from 30 to 64 years. They had been in practice for a mean of 14.2 ± 9.2 SD years, with 13.6% practising for <5 years and 27.3% for >20 years. Practice location was 71.4% urban, 23.8% rural, and 4.8% other type.

The 22 participating physicians performed DUSOI ratings on 548 of their patients during Phase 1 and on 643 patients in Phase 2, for a total of 1191 in the study. In both phases, there were 59.6% females. However, the age distribution differed because of the selection criterion for Phase 1 that required the presence of at least one specified chronic illness. The mean age of patients in Phase 1 was 65.6 ± 14.2 SD years (range 17–98) compared with 53.7 ± 18.0 SD years (range 17–96) in Phase 2. In Phase 1, 3.3% of patients were in the age range 17–33 years, and 58.7% in the >65 year range, compared with 17.8% and 32.3%, respectively, for Phase 2. For all study patients, the mean age was 59.2 ± 17.4 SD years, with 11.1% in the 17–33 range, and 44.4% in the >65 range.

The 1191 study patients were found to have 2488 health problems. Because patients in Phase 1 were selected only if they had certain health problems, the distribution of problems did not reflect the practice profile of the participating physicians. A much more representative problem profile was shown in Phase 2, where 1154 health problems were identified for the 643 patients, with the following prevalences for the 15 most frequent problems: hypertension, 14%; diabetes mellitus, 8.6%; depression, 7.6%; acute upper respiratory infection, 6.2%; back pain, 5.3%; chronic ischaemic heart disease, 3.6%; asthma, 3.3%; acute bronchitis, 3.0%; osteoarthritis, 3.0%; obesity, 2.8%; COPD, 2.8%; anxiety, 2.6%; ill-defined conditions, 2.6%; irritable bowel, 2.2%; and urinary tract infection, 2.2%. There were insufficient numbers of practitioners in each country to allow meaningful comparisons of physician or patient characteristics among the nine countries.

For interrater reliability, as tested on the three fictitious patient vignettes, the ICC was 0.45 for the DUSOI scores of the nine health problems. Intrarater ICCs for each of the nine health problems were as follows: obesity, 0.78; cystitis, 0.68; anxiety, 0.65; ischaemic heart disease, 0.62; diabetes type 1, 0.61; acute bronchitis, 0.46; low back pain, 0.42; depression, 0.40; and partner being ill, 0.39.

Feasibility was tested in terms of the amount of time required to perform the DUSOI rating on real patients and the amount of difficulty involved, as perceived by the raters themselves. For the 22 raters, the mean time required for each of the 1191 patients was 1.9 ± 1.3 SD minutes, with a minimum of <1 minute and a maximum of 10 minutes. No difficulty was found by the physicians in performing the ratings for 71.1% of the patients, some difficulty for 27.1%, and a lot of difficulty for 1.8%.

Potential clinical usefulness of the DUSOI severity ratings at the time of the rating process was judged to be none for 31.7% of the patients, some usefulness for 53.6%, and a lot of usefulness for 14.7%. Clinical usefulness was considered to be greater for the patients with the more severe health problems. This was shown by the higher mean overall DUSOI score (i.e. the combination of DUSOI scores for all health problems in a given patient) for the patients judged as having a lot of usefulness than the mean score for those with none or some usefulness (54.0 ± 18.4 SD versus 45.2 ± 19.7 SD on a scale of 0–100, P = 0.0001).

The severity of the 2488 health problems reported by the 22 practitioners covered the full range of the
Severity classification

Table 1: Inter-diagnosis and intra-diagnosis severity of illness distribution for diagnoses with a frequency of 20 or higher (n = 2488 health problems in 1191 patients of 22 providers in nine countries)

<table>
<thead>
<tr>
<th>Health problems</th>
<th>n</th>
<th>Severity scorea (scale 0–100)</th>
<th>Severity classification codeb (scale 0–4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>142</td>
<td>53.2 ± 19.0</td>
<td>0 (0%) 8.4 (1%) 44.4 (2%) 36.6 (3%) 10.6 (4%)</td>
</tr>
<tr>
<td>Acute bronchitis</td>
<td>27</td>
<td>45.8 ± 10.8</td>
<td>0 (0%) 3.7 (1%) 70.4 (2%) 25.9 (3%) 0 (4%)</td>
</tr>
<tr>
<td>Asthma</td>
<td>32</td>
<td>43.8 ± 17.1</td>
<td>0 (0%) 15.6 (1%) 59.4 (2%) 18.8 (3%) 6.2 (4%)</td>
</tr>
<tr>
<td>Acute upper respiratory infection</td>
<td>54</td>
<td>26.4 ± 12.9</td>
<td>0 (0%) 61.1 (1%) 33.2 (2%) 3.7 (3%) 0 (4%)</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other arthritisc</td>
<td>23</td>
<td>50.0 ± 17.5</td>
<td>0 (0%) 4.3 (1%) 47.8 (2%) 43.5 (3%) 4.3 (4%)</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>162</td>
<td>48.8 ± 17.6</td>
<td>0.6 (0%) 13.0 (1%) 45.1 (2%) 37.0 (3%) 4.3 (4%)</td>
</tr>
<tr>
<td>Back pain without radiating symptoms</td>
<td>42</td>
<td>40.9 ± 16.2</td>
<td>2.4 (0%) 19.0 (1%) 47.6 (2%) 31.0 (3%) 0 (4%)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart failure</td>
<td>24</td>
<td>48.7 ± 13.8</td>
<td>0 (0%) 0 (1%) 75.0 (2%) 25.0 (3%) 0 (4%)</td>
</tr>
<tr>
<td>Ischaemic heart disease</td>
<td>138</td>
<td>41.5 ± 17.8</td>
<td>0 (0%) 20.3 (1%) 59.4 (2%) 14.5 (3%) 5.8 (4%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>209</td>
<td>26.1 ± 13.1</td>
<td>0.9 (0%) 64.1 (1%) 31.6 (2%) 2.9 (3%) 0.5 (4%)</td>
</tr>
<tr>
<td>Mental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>181</td>
<td>46.5 ± 16.3</td>
<td>0 (0%) 14.4 (1%) 48.6 (2%) 34.8 (3%) 2.2 (4%)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>27</td>
<td>44.9 ± 16.6</td>
<td>0 (0%) 22.2 (1%) 55.6 (2%) 18.5 (3%) 3.7 (4%)</td>
</tr>
<tr>
<td>Metabolic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity</td>
<td>48</td>
<td>43.6 ± 17.2</td>
<td>0 (0%) 14.6 (1%) 62.5 (2%) 18.7 (3%) 4.2 (4%)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>183</td>
<td>35.5 ± 17.3</td>
<td>0.6 (0%) 37.2 (1%) 47.5 (2%) 12.0 (3%) 2.7 (4%)</td>
</tr>
<tr>
<td>Lipid disorders</td>
<td>26</td>
<td>17.1 ± 8.6</td>
<td>3.8 (0%) 88.5 (1%) 7.7 (2%) 0 (3%) 0 (4%)</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other stomach and duodenal disordersd</td>
<td>22</td>
<td>42.9 ± 21.5</td>
<td>0 (0%) 18.2 (1%) 63.6 (2%) 9.1 (3%) 9.1 (4%)</td>
</tr>
<tr>
<td>Ill-defined conditions</td>
<td>21</td>
<td>33.3 ± 21.1</td>
<td>14.3 (0%) 28.6 (1%) 33.3 (2%) 19.0 (3%) 4.8 (4%)</td>
</tr>
<tr>
<td>Cystitis</td>
<td>22</td>
<td>33.2 ± 15.7</td>
<td>4.5 (0%) 40.8 (1%) 50.0 (2%) 4.5 (3%) 0 (4%)</td>
</tr>
<tr>
<td>All health problems</td>
<td>2488</td>
<td>39.1 ± 19.4</td>
<td>1.6 (0%) 29.9 (1%) 45.9 (2%) 19.3 (3%) 3.3 (4%)</td>
</tr>
</tbody>
</table>

a Duke Severity of Illness Checklist (DUSOI) scores (scale = 0–100 from lowest to highest severity).
b DUSOI classification codes (scale = 0–4, from lowest to highest severity).
c Arthritis other than rheumatoid arthritis, osteoarthritis, or traumatic arthritis.
d Stomach and duodenal disorders other than peptic ulcer and neoplasms.

DUSOI measures (i.e. from 0 to 4 for the severity classification codes, and from 0 to 100 for the severity scores), with a mean of 1.9 ± 0.8 SD and 39.1 ± 19.4 SD, respectively. The distribution of health problems among the five severity code categories followed a normal distribution, with 1.6% for the no severity category (i.e. code = 0), 29.9% for mild severity, 45.9% for intermediate severity, 19.3% for moderate severity, and 3.3% for maximum severity (i.e. code = 4).

The variation in magnitude of the mean DUSOI scores among the 18 most prevalent health problems in this dataset, as well as the distribution of DUSOI classification codes within each of these health problems, are shown in Table 1. For example, interdiagnosis variation in severity for respiratory health problems was shown by the expected higher severity for COPD (53.2 ± 19.0 SD) than for acute upper respiratory infection (URI) (26.4 ± SD). Using the same example from Table 1, intra-diagnosis variation in severity was shown by the difference in distribution of severity codes, where 10.6% of patients with COPD had the DUSOI code 4 severity classification (i.e. maximum severity) and 8.4% had code 1 classification (i.e. mild severity), contrasted with patients having URI, where 0% were classified code 4, and 61.1% were classified code 1.

Of the 18 family/general practitioners who responded to the evaluation survey at the end of the field trial, 77.8% believed that the DUSOI was easy enough to use in practice; 11.1%, possibly easy enough; and 11.1%, not easy enough to use. The DUSOI was thought to be more useful in older patients by 77.8% and in patients with chronic health problems by 72.2%. It was not considered to be useful in any aged patient by 22.2%, or any type of health problem by 27.8%.

The immediate scoring method used in Phase 2 was thought definitely to improve the usefulness by 41.2% of the participants; somewhat to improve usefulness by 23.5%; and not to improve usefulness by 35.3%. The DUSOI was considered to be a very accurate indicator of the severity of the patient's health problems by 22.2%, somewhat accurate by 66.7%, and not accurate
Discussion

This field trial demonstrated that the DUSOI in a manual format was feasible for family/general practitioners to use in everyday clinical practice and was considered by most of the practitioners to have potential clinical usefulness, especially for older patients and those with chronic health problems. Similar findings were reported in a study of US family physicians using the computerized version of the DUSOI. This verification of results on an international scale adds strong support for the potential clinical applicability of the DUSOI and emphasizes the importance of collaborative studies performed in multiple countries. A measurement system such as the DUSOI that is developed initially in one country can be tested internationally to assess its potential usefulness on a much broader scale. Also, international testing will become increasingly important as more instruments that have been used primarily for research are tested for their clinical applicability.

Reliability (i.e. reproducibility or agreement) of the DUSOI severity ratings was shown to be better within the same raters (intrarater reliability) than between different raters (interrater reliability). Intrarater ICCs were >0.60 (i.e. "good", according to Fleiss) for five of the nine health problems, while the interrater ICC for the nine problems in combination was 0.45 (i.e. "fair", according to Fleiss).

Considerable variability among clinician raters might be expected for the DUSOI ratings because the instrument relies entirely upon the clinical judgement of the individual raters, which in itself may vary widely. However, the limits of clinically acceptable variability have not been established for a measure like the DUSOI. Additional research is needed to determine, for example, whether ICCs in the range of 0.30–0.50, which are considered by Fleiss to be of "poor", or at best "fair", reliability, may be actually "good" for measures that depend upon clinical judgement.

Although it is possible in the present study that native language differences among the raters may have resulted in decreased reliability because of different interpretations of the vignettes (all of which were written in English), it is doubtful that this was a determining factor since all of the raters were English literate. However, it is possible in an international field trial like the present one, where all of the instructions for how to perform the DUSOI ratings were given by correspondence written in English, and where successful completion of only two brief practice vignettes was considered adequate DUSOI training, that the minimum training resulted in higher variability. Eccles et al. in the UK found that intrarater reliability improved after the rater received additional DUSOI training.

Although this study was not originally designed to measure the validity of the DUSOI, the resulting data provided support for its clinical face validity. The clinically expected intra-diagnosis and inter-diagnosis variations in severity were shown by the correspondingly appropriate variations in DUSOI scores and the DUSOI severity classification codes. These findings indicate to the WONCA Classification Committee that the DUSOI may offer a valid severity classification component for the ICPC international classification system. Such an addition will be very important in strengthening the accuracy of classification for primary care health problems, because neither the present ICPC nor other systems include codes that adequately account for intra- and inter-diagnosis variation in severity of illness for all health problems. Severity classification is available only for selected diagnoses and usually for coding only complications, such as the supplementary severity codes for complications of diabetes mellitus that are offered by the International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10).

Although the raters were not asked their opinion about the suitability of the DUSOI for severity classification, they were asked about feasibility, clinical usefulness, accuracy, and their plans for personal use of the DUSOI. While most of the raters found the DUSOI easy to use and possibly of use in the clinical care of older and chronically ill patients, 58.8% responded that they did not intend to use the DUSOI in patient care. More raters planned to use the measure in research and teaching than in patient care and practice management. Further study is needed to assess the clinical usefulness when only the sickest patients or those at highest risk for ill health are rated for severity, rather than including patients whose severity is obviously mild, and who are at low risk.

Other studies are needed to refine the DUSOI methodology to improve its reliability. Among the
important factors that need testing are the effect of native language translations, the effect of more intense training of raters, and the effect of concurrent severity ratings of a series of real patients by two different raters, face to face with the same patient. This additional research is justified now that the DUSOI has been shown to be feasible in family/general practice, clinically useful in selected patients, and valid in quantifying variations in severity within the same diagnosis and among different diagnoses.

Acknowledgements

The 22 family/general practitioners who collected the complete dataset in time for the analyses in this paper are: Betty Meyboom-de Jong, Niek Jon de Wit, Klaas Reenders, Frans Rutten and Nico van Duijn of The Netherlands; Juan Gervas, Antonio Lopez-Miras, Raimundo Pastor-Sanchez and Mercedes Perez-Fernandez of Spain; Gerald Deneckere, Marc Jamoulle and Michel Roland of Belgium; Stephen Blades, Colin Bradshaw and Martin Eccles of the UK; Cindy Lam and Tai-Pong Tam of Hong Kong; Duane Lawrence and James Mold of the USA; Udo Schirmer of Germany; Philip Sive of Israel; and Manabu Yoshimura of Japan.

Data were managed by Alverta Sigmon and statistical analyses were performed by Jessica C-K Tse in the Department of Community and Family Medicine, Duke University Medical Center.

The DUSOI form was reproduced with permission of Duke University Medical Center.

References