Development and initial validation of international severity scoring system for familial Mediterranean fever (ISSF)

Article in Annals of the rheumatic diseases · January 2016
DOI: 10.1136/annrheumdis-2015-208671

CITATIONS
10

READS
145

11 authors, including:

Erkan Demirkaya
Gulhane Military Medical Academy
171 PUBLICATIONS 1,382 CITATIONS
SEE PROFILE

Huri Ozdogan
Istanbul University
163 PUBLICATIONS 3,238 CITATIONS
SEE PROFILE

Omer Karadag
Hacettepe University
182 PUBLICATIONS 1,475 CITATIONS
SEE PROFILE

Seza Ozen
Hacettepe University
520 PUBLICATIONS 12,529 CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Standards and recommendations for transitional care management for adolescents and young people with rheumatic diseases View project

prospective study View project
EXTENDED REPORT

Development and initial validation of international severity scoring system for familial Mediterranean fever (ISSF)

Erkan Demirkaya,1,2 Cengizhan Acikel,2,3 Philip Hashkes,4 Marco Gattorno,5 Ahmet Gul,6 Huri Ozdogan,7 Turker Turker,8 Omer Karadag,9 Avi Livneh,10 Eldad Ben-Chetrit,11 Seza Ozen,12 FMF Arthritis Vasculitis and Orphan Disease Research in pediatric rheumatology (FAVOR)

ABSTRACT

Objective To develop widely accepted international severity score for children and adult patients with familial Mediterranean fever (FMF) that can be easily applied, in research and clinical practice.

Methods Candidate severity criteria were suggested by several FMF expert physicians. After three rounds of Delphi survey, the candidate criteria, defined by the survey, were discussed by experts in a consensus meeting. Each expert brought data of clinical manifestations, laboratory findings and physician’s global assessments (PGAs) of minimum 20 patients from their centres. We used the PGAs for disease severity as a gold standard. Logistic regression analysis was used to evaluate the predicting value of each item, and receiver operating characteristic curve analysis was performed to demonstrate the success of the criteria set.

Results A total of 281 patients consist of 162 children and 119 adults with FMF were enrolled and available for validity analysis: Nine domains were included in the final core set of variables for the evaluation of disease severity in FMF. The International Severity Score for FMF (ISSF) may reach a maximum of 10 if all items are maximally scored. The threshold values to determine: severe disease ≥6, intermediate disease 3–5, mild disease ≤2. Area under the curve was calculated as 0.825 for this set in the whole group.

Conclusions The initial validity of ISSF both in children and adult with FMF was demonstrated. We anticipate that it will provide a robust tool to objectively define disease severity for clinical trials, future research as well as for therapeutic decisions in managing patients with FMF.

INTRODUCTION

Familial Mediterranean fever (FMF) is an autoinflammatory hereditary disease characterised by recurrent attacks of fever and serositis.1 2 The disease may affect the peritoneum, pleura, joints and skin. The frequency of the attacks may vary from once a week to once per several months. The course and characteristics of the disease may be changed among patients with FMF and in the same patient over the years.

During the last two decades, the requirement to improve quality of life (QOL) in patients with chronic illness and create instruments to measure different aspects of diseases has become an important topic in health policy. A standardised assessment of disease severity will help physicians to compare patients with regard to amenability to treatment, cost effectiveness and safety of treatment. QOL, chronic sequel, burden imposed on the medical authorities and prediction of expected clinical course. Severity scoring systems have been developed to objectively quantify the disease severity for both therapeutic and prognostic purposes.3 A standardised assessment of disease severity will help physicians both to evaluate the response to therapy and to conduct clinical trials, especially in patients with colchicine-resistant FMF.

Frequent and severe FMF attacks may result in serious complications such as development of secondary amyloidosis, growth and puberty delay, chronic arthritis, anaemia and compromise of QOL.4 Clinical spectrum of FMF is heterogeneous with a wide range, ranging between minimal activity of few affected sites and excellent response to colchicine and large number of frequent, intolerable, treatment-resistant attacks.5 A widely accepted measure, rationally ranking disease expression according to the degree of disease severity, would be useful in the management of this lifelong disease.

There are at least three severity scores in FMF. The most commonly used are those developed by Pras et al6 and Mor et al.7 There is no consensus on any of them. Thus, Kalkan et al7 showed that these two scoring systems were not statistically consistent with each other. International group of FMF experts realised the occurrence of this important concern.8 9 and established an international consortium assembled to develop a severity scoring tool that will be widely acceptable and guide in the assessment of disease severity in patients with FMF. The newly developed criteria were intended to be appropriate for use both in clinical practice and drug trials in children as well as in adults.

PATIENTS AND METHODS

The goal of the project was to reach a consensus on international severity criteria for children and adult patients with FMF that can be applied easily, mainly for research purposes, and that would be widely accepted. For this specific project it was

Handling editor Tore K Kvien

Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2015-208671).

For numbered affiliations see end of article.

Correspondence to Dr Erkan Demirkaya, Associate Professor of Paediatrics, Gulhane Military Medical Faculty, Pediatric Rheumatology Unit, FMF Arthritis Vasculitis and Orphan Disease Research in pediatric rheumatology (FAVOR), Etlik, Ankara 06018, Turkey; erkandemirkaya@yahoo.com

Received 28 September 2015
Revised 4 January 2016
Accepted 10 January 2016

decided to involve an international group of experts with extensive clinical experience of FMF.

The project was divided into three phases. In the first phase the Delphi technique was used to gather opinions from a large number of paediatric and adult rheumatologists. The second phase involved a consensus conference using nominal group technique (NGT). Validation of the new criteria was the third phase of the study.\(^\text{10–11}\)

**Step 1: Delphi technique (results of surveys through electronic mail)**

Delphi survey technique was used for the initial phase of the study. Three sequential questionnaire-based surveys were carried out to select and then rank the variables used in routine clinical practice to assess the severity of a patient with FMF. A body of experts including 24 FMF specialists from 16 countries participated in the survey. The first Delphi round aimed to identify all clinical and laboratory features considered to be associated with the severity of FMF. In the second round, 33 structured questions were developed to collect expert opinions about FMF severity and physicians were asked to indicate up to 10 variables that they judged as best reflecting severity. Mailing, email, fax or telephone reminders were used to ensure a response rate of at least 80% for both surveys. At the last rounds, the experts in the Delphi panel were asked to evaluate the answers given to the questions in the previous two rounds of Delphi survey.

**Step 2: consensus meetings and evaluation of preliminary results**

NGT is a structured face-to-face meeting designed to facilitate reaching consensus, through round discussion. Following the Delphi surveys the first consensus panel (NGT) was held in Washington DC, USA on 13 November 2012, with nine experts and one facilitator (CA) to define the necessary items for a set of criteria to be used for the evaluation of severity criteria in FMF. Participants (listed as coauthors) were international experts in management of FMF. The second panel was held in Lausanne, Switzerland on 23 May 2013, to discuss the results related to validation phase of the study and to have last commentaries from the international expert group. Both of the consensus meetings were organised by FMF Arthritis Vasculitis and Orphan disease Research in Paediatric Rheumatology (FAVOR) (http://www.favor.org.tr). Consensus formation methodology was designed so that each step was based on the results of the previous steps and items for the candidate criteria and their standard definitions were discussed. Prior to the meeting participants received a booklet containing relevant articles and previous results of the Delphi surveys.

**Step 3: validation of the defined criteria set**

**Data collection**

Physicians who were running a dedicated outpatient clinic for FMF were invited to register their patients with FMF for the study. Registered patients were required to be diagnosed as FMF according to Tel-Hashomer\(^\text{12}\) or paediatric FMF\(^\text{13}\) criteria without any other accompanying inflammatory disease. All patients were receiving colchicine at stable doses for at least 6 months. The severity of patient was evaluated by senior FMF expert in each centre based on the clinical course of the disease. In addition, for each patient, a five-page case report form was completed. This included demographic data, and information regarding clinical variables that might represent risk factors for the severe disease course was collected at the study visit based on history and medical records. Demographic information included date of diagnosis of FMF, date of first observation at the tertiary care centre and disease duration (from the diagnosis to the time of the study visit). Additional data were collected on disease complications, chronic sequelae (growth retardation, anaemia, diarrhoea, amyloidosis, proteinuria, etc) and information about the attacks: (i) characteristics of fever episodes (duration, frequency, triggers); (ii) presence and frequency (always or often/sometime) of the clinical manifestations; (iii) type and site of the attacks and treatment modalities. The physician’s global assessment (PGA) of disease severity was recorded on a 10 cm (21 circle numbered) visual analogue scale (VAS). Physician’s final decision about patient’s severity status was also recorded dichotomously.

Detailed laboratory findings were also collected including C-reactive protein, erythrocyte sedimentation rate, fibrinogen, serum amyloid A and MEFV mutations, previously determined by acceptable methods of molecular analysis (according to the Infever database, http://fmf.igh.cnrs.fr/ISSAID/infevers/). Each centre enrolled at least 10 severe and 10 non-severe patients for the study. The electronic forms contained predefined rules to avoid errors and missing data, and were reviewed for consistency by a dedicated FAVOR research assistant. Nine rheumatologists, expert in FMF and autoinflammatory diseases, from three countries enrolled their patients. The study was approved by local ethics committees and informed consent was obtained from parents/guardians if the participant was younger than 18.

**Validation**

Validation of the core set measures was conducted with the use of the Outcome Measurement Sets for Clinical Trials (OMERACT) filter for outcome measures in rheumatology.\(^\text{14–15}\) The applicability and practicality of the measures were determined by addressing the topics of briefness, simplicity, ease of scoring and percentage of missing values.

We used the PGA for disease severity as a gold standard. Expert physicians assigned their patients to the severe or non-severe groups. Additionally, they also evaluated their patient with the 21 circle numbered (10 cm VAS) according to degree of disease severity. To increase standardisation between different raters, joint working group studies and discussions session were performed before the data collection step. Records were then re-evaluated in a blinded manner (patient demographic data blinded) by independent experts (SO, ED and HO) to confirm severity status. The disease experts had the mandate to control the consistency and the quality of data. In case of inconsistency or other doubts, specific queries were resubmitted to the participating centres for resolution.

Student’s t test was used for comparison of continuous variables between the severe and non-severe groups. Construct validity was determined to analyse the relations of other measures with the priori prediction. PGA of the patient’s overall disease severity was adopted as a representative measurement.

Performance of existing scores: there are two currently used severity scoring sets in FMF, the first one was established by Pras et al.,\(^\text{5}\) based on their experience and the appreciation, and the other set was developed by Mor et al.,\(^\text{6}\) based on statistical methods. The scoring system of Mor et al has six elements, including age of onset, dose of colchicine, number of involved sites in one attack and during the course of the disease, and the presence of pleuritic and erysipelas-like attacks during the course of the disease. The scoring system of Pras et al also has six elements, including age of onset, dose of colchicine, number of attacks per month, presence of arthritis, erysipelas-like erythema and amyloidosis.

---

The Mor and Pras severity scores were automatically calculated, at the time of enrolment, and validated independently by a statistician (CA). The ability of the Pras and Mor severity scores and the new International Severity Score for FMF (ISSF) to discriminate patients with FMF with and without severe course, according to the attending physician, was evaluated using the receiver operating characteristic (ROC) curve analyses.

We performed univariate and multivariate logistic regression analyses to evaluate importance of each criteria and tried to define the best set that can predict disease severity. To evaluate success of criteria set and to determine the best cut-off points for each item, several ROC analyses were performed. These results were presented during the second expert panel to guide the panel members’ decisions about candidate items.

Sensitivity, specificity, negative predictive value and positive predictive value were calculated for different cut-off points. The items that remained significant were used for the final criteria set (total items: 9; total point: 10). Statistical analysis was performed by using SPSS for Windows V.21.0.

RESULTS
A total of 281 patients with FMF were enrolled and available for analysis: 162 children and 119 adult with FMF. Demographic data were available for 281 patients and summarised in table 1. Clinical and laboratory characterisation of the patients were summarised in online supplementary table S1. Patients who were enrolled in the study originated from Turkey, Israel, Ukraine, Tunisia, USA, Egypt, Netherlands, Italy, Morocco and Iraq.

Step 1: Delphi surveys
The questionnaires of first round of Delphi were sent to 42 experts from the Eurofever database and 24 (57.1%) responded. The experts agreed on the necessity of a new severity assessment score for FMF. Twenty-nine clinical features, 16 laboratory parameters and 11 sociodemographic (such as gender, ethnicity, etc) variables were suggested by the experts at the end of first Delphi. After two rounds of Delphi, 21 variables were selected to be discussed in the nominal group meeting. All Delphi steps were completed by 21 of 24 (87.5%) experts who had joined the first Delphi round.

Step 2: classification by consensus panel and validation
At the first consensus meeting, the panellists discussed preliminary data from the Delphi surveys and voted candidate variables. Ten candidate variables were selected for the disease severity assessment and each variable was defined.

After preliminary data analysis the results were discussed in the second consensus meeting. Organ dysfunction was eliminated from the list after this meeting. The nine domains included in the final core set of variables for the evaluation of disease severity in FMF and the related suggested variables to measure each domain are shown in table 2. All items were scored as simplified no (0)/yes (1) except for the frequency of attacks. The ISSF may reach a maximum of 10 if there are all items recorded and scored. If the patient had an average of between one and two attacks per month this was scored as 1 and if there were > 2 attacks per month a score of 2 was given. If the subject experienced less than one attack per month this item was scored as 0. Validity measures including sensitivity and specificity for different threshold points are given in table 3.

In the final analytical step, in order to properly calculate sensitivity, specificity and ROC cut-off, the 281 patients were dichotomised as having severe or not severe disease. According to the ROC curve and related coordinate points (figure 1), the ISSF ≥ 6 points identified severe patients (positive predictive value 100%), while a total score between three and five points identified patients with intermediate severity, and ≤ 2 was...
defined as mild disease (negative predictive value 88%). The area under the curve (AUC) of ISSF was 84.8% (95% CI 80.4% to 89.3%). The new severity score had a greater AUC score in ROC curves than the Mor and Pras scores, both in adult and children groups (figure 1).

**DISCUSSION**

This is the first effort in establishing international severity criteria for FMF. The criteria set have been developed with a consensus-driven methodology by paediatricians and internists, with expertise in this disease, and validated in a large database comprising both children and adults with FMF. These measures enable to quantify and compare patient populations expressing extremely wide range and heterogeneous disease manifestations.

There have been previous attempts to develop severity criteria in FMF. The severity scoring systems for FMF were built and intended for use in adults by two different groups. The most commonly used criteria were those developed by Pras and Mor criteria also performed well. However, both specificity and sensitivities were lower than the current criteria. It seems that one of the major issues in disease assessment is going to be solved by the ISSF. It may validated and both were developed for adult patients. Moreover, low agreement between the two sets of criteria has been previously shown.7, 16 There has been a paediatric attempt to develop a modified severity criteria; however, this has not been widely used.17 Modifications of the adult scoring instruments to fit paediatric FMF have been proposed and used in a few studies. These modified scores were proven to be very effective in distinguishing between disease severity of Turkish children living in Germany with mild phenotype to Turkish children living in Turkey with a more severe expression.17 However, the performance of the two modified sets did not correlate with each other giving the impression that a well-performing paediatric score tool should be established. In a recent study, Kalkan et al8 have evaluated the consistency of these modifications in a series of paediatric patients and they found that the results of these two scoring systems were not statistically consistent with each other. Inconsistency of two scoring systems and lack of correlation between the scoring systems raise concerns in the reliability of these scoring systems. It is not surprising that in a recent commentary the need for a new set of criteria has been highlighted.9 An international expert panel of physicians and researchers taking care of patients with FMF came together in order to establish ISSF for FMF because progress in the management of autoinflammatory diseases requires parallel improvements in the ability to measure and record disease severity in a uniform manner.

Due in part to lack of standard outcome measures very few randomised controlled trials have been conducted in FMF. In these studies authors used different primary outcome measures such as attack frequency,18–20 ‘abortion’ of the attacks,21 severity of the attacks as defined by the patient,22–24 QOL scales and the serum levels of acute-phase reactants.20, 24 The same problem can also be seen in observational clinical and epidemiological studies.16, 25–27 Because of the absence of validated uniform instrument to use in these studies it is too difficult to perform meta-analysis or systematic reviews.28, 29 We have assessed the ROCs for all three criteria (Mor, Pras and ISSF) and have shown that this new criteria had the highest sensitivity and specificity. The criteria published by Pras and Mor also performed well. However, both specificities and sensitivities were lower than the current criteria. It seems that one of the major issues in disease assessment is going to be solved by the ISSF. It may

| Table 3 Validity measures for different threshold points for International Severity Score for familial Mediterranean fever (ISSF) |
|---|---|---|---|---|---|
| Score | TP/FP | TN/FN | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
| 0 | 149/0 | 11/121 | 100 | 8 | 55 | 100 |
| 1 or more | 147/2 | 37/95 | 99 | 28 | 61 | 95 |
| 2 or more | 138/11 | 78/54 | 93 | 59 | 72 | 88 |
| 3 or more | 112/37 | 103/29 | 75 | 78 | 79 | 74 |
| 4 or more | 78/71 | 120/12 | 52 | 91 | 87 | 63 |
| 5 or more | 39/110 | 129/3 | 26 | 98 | 93 | 54 |
| 6 or more | 18/131 | 132/0 | 12 | 100 | 100 | 50 |
| 7 or more | 5/144 | 132/0 | 3 | 100 | 100 | 48 |
| 8 or more | 1/148 | 132/0 | 1 | 100 | 100 | 47 |
| 9 or more | 0/149 | 132/0 | 0 | 100 | 0 | 47 |
| 10 | 0/149 | 132/0 | 0 | 100 | 0 | 47 |

FN, false negative; FP, false positive; NPV, probability of true non-severe patients rate according to the ISSF; PPV, probability of true severe patients rate according to the ISSF; sensitivity, detected proportion of true severe patients; specificity, detected proportion of true non-severe patients; TN, true negative; TP, true positive.

Figure 1 ROC of disease severity according to different severity scoring systems and also age groups (children, adult and all patients) showing an AUC of 0.84 (95% CI 0.804 to 0.893) in 281 patients with FMF. AUC, area under the curve; FMF, familial Mediterranean fever; ISSF, International Severity Score for FMF; ROC, receiver operating characteristics curve.
provide the scientists and physicians clear measures for follow-up of their patients from childhood till adulthood in a uniform manner.

Perhaps another major role of severity scoring tool, and in this respect it differs from disease activity tool and disease outcome measure, is to define phenotypically different patient populations. This aspect of the tool has been used extensively to show ethnic variability in the severity of FMF

Ethics approval The study was approved by the local ethics committee of Gulhane Military Medical Academy.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
Clinical and epidemiological research


