Classification criteria for autoinflammatory recurrent fevers

MarcoGattorno,1 Michael Hofer,2,3 Silvia Federici,4 Federica Vanoni,5 Francesca Bovis,6 Ivona Aksentijevich,7 Jordi Anton,9 Juan Ignacio Arostegui,9 Karyl Barron,10 Eldad Ben-Cherit,11 Paul A Brogan,12 Luca Cantarini,13 Isabella Ceccherini,14 Fabrizio De Benedetti,15 Fatma Dedeoglu,16 Erkan Demirkaya,17 Joost Frenkel,18 Raphaella Goldbach-Mansky,19 Ahmet Gul,20 Veronique Hentgen,21 Hal Hoffman,22 Tilmann Kallinich,23 Isabelle Kone-Paut,24 Jasmin Kuemmerle-Deschner,25 Helen J Lachmann,26 Ronald M Laxer,27 Avi Livneh,28 Laura Obici,29 Seza Ozen,30 Dorota Rowczenio,26 Ricardo Russo,31 Yael Shinar,32 Anna Simon,33 Nataša Toplak,34 Isabelle Touitou,35 Yosef Uziel,36,37 Marielle van Gijn,38 Dirk Foell,39 Claudia Garassino,40 Dan Kastner,10 Alberto Martini,40 Maria Pia Sormani,6,41 Nicolino Ruperto,42 for the Eurofever Registry and the Paediatric Rheumatology International Trials Organisation (PRINTO)

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-215048).

For numbered affiliations see end of article.

Correspondence to
Dr Marco Gattorno, UOSD Centro Malattie Autoinflammatorie e Immunodeficienze, IRCCS Istituto GianninaGaslini, Genoa 16147, Italy; marcogattorno@gaslini.org

MG and MH contributed equally.

Received 9 January 2019
Revised 30 March 2019
Accepted 1 April 2019
Published Online First 24 April 2019

ABSTRACT

Background Different diagnostic and classification criteria are available for hereditary recurrent fevers (HRF)—familial Mediterranean fever (FMF), tumour necrosis factor receptor-associated periodic fever syndrome (TRAPS), mevalonate kinase deficiency (MKD) and cryopyrin-associated periodic syndromes (CAPS)—and for the non-hereditary, periodic fever, aphthosis, pharyngitis and adenitis (PFAPA). We aimed to develop and validate new evidence-based classification criteria for HRF/PFAPA.

Methods Step 1: selection of clinical, laboratory and genetic candidate variables; step 2: classification of 360 random patients from the Eurofever Registry by a panel of 25 clinicians and 8 geneticists blinded to patients’ diagnosis (consensus ≥80%); step 3: statistical analysis for the selection of the best candidate classification criteria; step 4: nominal group technique consensus conference with 33 panellists for the discussion and selection of the final classification criteria; step 5: cross-sectional validation of the novel criteria.

Results The panellists achieved consensus to classify 281 of 360 (78%) patients (32 CAPS, 36 FMF, 56 MKD, 37 PFAPA, 39 TRAPS, 81 undefined recurrent fever). Consensus was reached for two sets of criteria for each HRF, one including genetic and clinical variables, the other with clinical variables only, plus new criteria for PFAPA. The four HRF criteria demonstrated sensitivity of 0.94–1 and specificity of 0.95–1; for PFAPA, criteria sensitivity and specificity were 0.97 and 0.93, respectively. Validation of these criteria in an independent data set of 1018 patients shows a high accuracy (from 0.81 to 0.98).

Conclusion Eurofever proposes a novel set of validated classification criteria for HRF and PFAPA with high sensitivity and specificity.

INTRODUCTION

In the last 20 years the discovery of the inflammasome and the related genes of the now called systemic autoinflammatory diseases (SAIDs) has led to a completely new line of research. SAIDs are caused by exaggerated activation of the innate immune system, in the absence of high-titre auto-antibodies or antigen-specific T-cells.1 Recurrent (or periodic) fevers are characterised by inflammatory flares separated by intervals of general overall well-being. Some conditions are caused by a genetic defect and are collectively referred to as hereditary recurrent fever (HRF). Familial Mediterranean fever (FMF) is caused by mutations of MEFV2,3;
mevalonate kinase deficiency (MKD), by mutations of the mevalonate kinase gene (MVK)\(^8\); tumour necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS), by mutations of type I TNF receptor (TNFRSF1A)\(^9\); and cryopyrin-associated periodic syndromes (CAPS), by mutations of NLRP3.\(^7\) More common forms of recurrent fever syndromes include periodic fever, aphthosis, pharyngitis and adenitis (PFAPA) syndrome, which is a multifactorial disorder.\(^3\) So far, several clinical diagnostic and classification criteria have been proposed for HRF.\(^10–15\) and PFAPA.\(^9\)\(^\text{16}\) Overall, these criteria lack accuracy and do not consider the results of genetic analyses, now an essential tool for the accurate diagnosis and classification of HRF.

This distinction between classification and diagnostic criteria is not always clear in clinical practice, and the two terms are often (wrongly) used interchangeably.\(^7\)\(^\text{17}\) Classification criteria facilitate accurate identification of diseases for clinical or epidemiological studies, in this context reliably differentiating one autoinflammatory disease from another, but are not designed to diagnose that autoinflammatory disease; hence, classification criteria make the assumption that important disease mimics (eg, chronic infection or malignancy) have already been excluded. In contrast, diagnostic criteria are designed to positively rule in a specific diagnosis in an individual patient, while excluding all conditions with different overlapping disease manifestations based on derivation and validation in cohorts that include important disease mimics. As such, classification criteria cannot be used as diagnostic criteria.\(^\text{18}\)\(^\text{17}\) The purpose of this study was to develop and validate new evidence-based classification criteria for HRF and PFAPA, combining international expert consensus and statistical evaluation of real patients from a large data set of patients in the Eurofever Registry.

### METHODS
A multistep process using consensus formation techniques (Delphi and nominal group technique (NGT)) and statistical evaluations on real patients was used to develop and test the classification criteria (online supplementary figure 1 and supplementary material), based on a methodological framework used successfully in previous studies in rheumatology.\(^\text{20–25}\)

**Step 1: selection of clinical, laboratory and genetic candidate variables**

A panel of 162 international adult and paediatric experienced clinicians completed successive Delphi questionnaires in order to propose and then select and rank the variables (clinical manifestations, genetic analyses, laboratory examinations) from 1 (less important) to 10 (most important), for classification of each HRF and PFAPA.\(^\text{27}\)

**Step 2: classification of patients from the Eurofever Registry**

After selection (online supplementary figure 2), a random sample of 360 patients, 60 patients for each disease (FMF, TRAPS, MKD, CAPS, PFAPA and undefined recurrent fevers (uRF)) were selected from the Eurofever Registry. The inclusion criteria for the enrolment in the registry have been previously described\(^\text{28}\) (see online supplementary material).

Twenty-five international experienced clinicians/researchers and eight geneticists (total of 33 panelists) in the field of SAID blinded on patients’ original diagnosis were invited to participate in a multiround, secured web process to classify each of the 360 patients into one of six mutually exclusive diagnoses.\(^\text{29}\) Clinicians and geneticists worked separately in the first steps (clinicians blinded to genetic results and geneticists blinded to clinical data) and then together to reach consensus ≥80% on all classifiable patients.

**Step 3: statistical analysis for the selection of the best candidate classification criteria**

The statistical analysis plan (full details in the online supplementary material) foresaw the following steps:

- Evaluation through a univariate logistic regression of the relationship between each individual top variable identified in step 1 and each disease as derived from the panel’s classification.
- Computer generation of more than 30 000 new candidate sets of classification criteria through linear combinations of genetic and clinical variables with improper linear modeling. Additionally, 11 sets of criteria were derived from the literature\(^\text{6–16}\) or proposed by members of the panel based on their expertise.
- Identification of the top-performing criteria through ranking according to the Akaike information criterion (AIC), with best model having the lowest AIC.

**Step 4: NGT Consensus Conference for the selection of the final classification criteria**

The Consensus Conference was held in Genoa, Italy, on 18–21 March 2017. Clinicians and geneticists, who participated in the step 2 web consensus classification exercise, attended a meeting. The overall goal of the meeting was to decide on the final set of criteria, using a combination of statistical and consensus (≥80%) formation techniques with the consensus panel classification as reference standard.

**Step 5: cross-sectional validation of the final classification criteria**

The performance of the final set of classification criteria to discriminate patients with the different HRF and PFAPA was tested, using the original treating physician patients’ diagnosis as reference standard for the cross-sectional validation, postconsensus, in a separate set of 1018 patients selected after random computer generation from the Eurofever Registry, which contains all variables included in the final set of classification criteria.

### RESULTS

The demographic, clinical, genetic and laboratory features of the 360 patients randomly selected from the Eurofever Registry are provided in table 1 and online supplementary table 1.

A total of 100 different genotypes were reported in the 360 patients included in the classification process as reported in online supplementary table 2.

Nine patients with CAPS and two with TRAPS had no mutations detected using Sanger sequencing; thus, at the time of enrolment, somatic mosaicism could not be formally excluded in them. Low penetrant or incidental (non-confirmatory) genetic findings were also reported in 7 patients with PFAPA and 14 with uRF (online supplementary table 3).

### Classification of patients from the Eurofever Registry

In the first two rounds, evaluation of clinical data by clinicians (blinded to genetic results) resulted in consensus of ≥80% for a total of 216 of 360 (60%) patients (figure 1); consensus was reached in 51 patients with MKD, 43 with TRAPS, 29 with FMF, 29 with CAPS, 26 with PFAPA and 38 with uRF. Similarly evaluation of demographic and genetic data by geneticists (blinded...
to clinical data) in two separate rounds reached consensus on 319 of 360 (89%) with 278 (77%) patients with 80% consensus after the first round. At the end of the two initial rounds, 128 (36%) patients were concordant between the independent evaluation of both the clinicians and the geneticists. At the end of the fourth round, consensus was achieved in 281 of 360 (78%) as follows: 56 (95%) MKD, 39 (76%) TRAPS, 37 (70%) PFAPA, 36 (71%) FMF, 32 (63%) CAPS and 81 (85%) uRF (figure 1, online supplementary table 4). K ( concordance coefficient) agreement between the panel reference standard classification and the original patient diagnosis by the treating physician was 0.99 for MKD, 0.87 for TRAPS, 0.86 for CAPS, 0.84 for FMF and 0.68 for PFAPA.

### Table 1  Demographic features of the 360 patients included in the study

<table>
<thead>
<tr>
<th>Criteria</th>
<th>FMF n=60</th>
<th>CAPS n=60</th>
<th>MKD n=60</th>
<th>TRAPS n=60</th>
<th>PFAPA n=60</th>
<th>uRF n=60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>30 (50%)</td>
<td>32 (53%)</td>
<td>26 (43%)</td>
<td>35 (58%)</td>
<td>28 (47%)</td>
<td>28 (47%)</td>
</tr>
<tr>
<td>Paediatric/Adults</td>
<td>54/6</td>
<td>33/27</td>
<td>45/15</td>
<td>29/31</td>
<td>59/1</td>
<td>39/21</td>
</tr>
<tr>
<td>Age, years, median (range)</td>
<td>10.5 (7.0–15.5)</td>
<td>16.0 (8.9–31.6)</td>
<td>16.2 (9.1–23.0)</td>
<td>21.9 (10.5–41.1)</td>
<td>6.6 (3.8–9.5)</td>
<td>13.5 (8.2–26.4)</td>
</tr>
<tr>
<td>Age at disease onset, median (range)</td>
<td>3.4 (1.2–6.4)</td>
<td>3.0 (0.5–11.2)</td>
<td>0.4 (0.2–0.9)</td>
<td>3.4 (0.8–10.6)</td>
<td>1.5 (0.7–3.0)</td>
<td>5.9 (2.0–19.1)</td>
</tr>
<tr>
<td>Disease duration, median (range)</td>
<td>5.6 (2.7–10.2)</td>
<td>9.0 (4.6–19.1)</td>
<td>14.2 (7.9–20.8)</td>
<td>13.3 (6.8–23.2)</td>
<td>3.9 (2.3–6.8)</td>
<td>4.8 (3.0–8.2)</td>
</tr>
<tr>
<td>Episode duration, median (range)</td>
<td>3.0 (2.0–4.0)</td>
<td>2.0 (0.8–5.0)</td>
<td>5.0 (4.0–7.0)</td>
<td>8.0 (5.0–18.0)</td>
<td>4.0 (3.0–5.0)</td>
<td>4.0 (3.0–7.0)</td>
</tr>
<tr>
<td>Number episodes/year, median (range)</td>
<td>12.0 (10.0–20.0)</td>
<td>12.0 (6.0–25.0)</td>
<td>12.0 (10.0–16.0)</td>
<td>6.0 (4.0–12.0)</td>
<td>12.0 (12.0–18.0)</td>
<td>12.0 (5.0–13.0)</td>
</tr>
</tbody>
</table>

CAPS, cryopyrin-associated periodic syndromes; FMF, familial Mediterranean fever; MKD, mevalonate kinase deficiency; PFAPA, periodic fever, aphthosis, pharyngitis and adenitis; TRAPS, tumour necrosis factor receptor-associated periodic fever syndrome; uRF, undefined recurrent fevers.

### Statistical analysis for the selection of the best classification criteria

The top variables arising from step 1 (see the Methods section) were included in a univariate logistic regression analysis using the 281 patients for which consensus was achieved by the panel as outcome. Clinical variables positively and negatively associated with each disease are reported in online supplementary table 5 with the related OR and 95% CI. The strategy for the classification of the genotypes is described in online supplementary table 6. A total of 198 over >30 000 possible new sets of classification criteria (available on request; 50 for CAPS, 45 for FMF, 44 for TRAPS, 32 for MKD and 22 for PFAPA) were retained, based on

---

Figure 1  Flow chart of the consensus nominal group technique for classification of patients from the Eurofever Registry. CAPS, cryopyrin-associated periodic syndromes; FMF, familial Mediterranean fever; MKD, mevalonate kinase deficiency; PFAPA, periodic fever, aphthosis, pharyngitis and adenitis; pts, patients; TRAPS, tumour necrosis factor receptor-associated periodic fever syndrome.
The performances of all the criteria chosen by the consensus together with 11 criteria from the literature (online supplementary table 7) for glossary. See table 2 for prerequisite criteria.

their AIC, for further evaluation at the Consensus Conference, together with 11 criteria from the literature (online supplementary figure 4).

NGT Consensus Conference for the selection of the final classification criteria

The first disease discussed was FMF. After multiple voting sessions, all three tables of experts, which worked independently from each other, showed a complete convergent validity selecting the same top definition number 38 (online supplementary figure 4, session A), including genetic and clinical variables with a positive association (table 2). After general discussion, a second set of criteria based solely on clinical criteria was selected to be used as a possible tool for the indication for molecular analysis or as classification criteria in case genetic testing is not locally available.

The performances of all the criteria chosen by the consensus in the 281 patients who reached a consensus are reported in tables 2 and 3 (see also glossary in online supplementary table 7).

**Table 2** New Eurofever/PRINTO classification criteria for hereditary recurrent fevers and their performance in the 281 patients with consensus

<table>
<thead>
<tr>
<th>Criteria</th>
<th>CAPS</th>
<th>FMF</th>
<th>TRAPS</th>
<th>MKD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of a confirmatory NLRP3 genotype* and at least one among the following:</td>
<td>Presence of confirmatory MEFV genotype* and at least one among the following:</td>
<td>Presence of confirmatory TNFRSF1A genotype* and at least one among the following:</td>
<td>Presence of a confirmatory MVK genotype* and at least one among the following:</td>
<td></td>
</tr>
<tr>
<td>Urticarial rash.</td>
<td>Duration of episodes 1–3 days.</td>
<td>Duration of episodes ≥7 days.</td>
<td>Gastrointestinal symptoms.</td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>Abdominal pain.</td>
<td>Periorbital oedema.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of not confirmatory NLRP3 genotype† and at least two among the following:</td>
<td>OR</td>
<td>Presence of a not confirmatory TNFRSF1A genotype† and at least two among the following:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urticarial rash.</td>
<td>Duration of episodes 1–3 days.</td>
<td>Duration of episodes ≥7 days.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>Abdominal pain.</td>
<td>Periorbital oedema.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity: 1
Specificity: 1
Accuracy: 1

A patient with (1) evidence of elevation of acute phase reactants (ESR or CRP or SAA) in correspondence to the clinical flares and (2) careful consideration of possible confounding diseases (neoplasms, infections, autoimmune conditions, other inborn errors of immunity) and a reasonable period of recurrent disease activity (at least 6 months) is classified as having hereditary recurrent fever if the criteria are met.

*Pathogenic or likely pathogenic variants (heterozygous in AD diseases, homozygous or in trans (or biallelic) compound heterozygous in AR diseases).
†Variant of uncertain significance (VUS). Benign and likely benign variants should be excluded.

**Table 3** Eurofever/PRINTO clinical classification criteria for PFAPA and hereditary recurrent fevers and their performance in the 281 for whom consensus was achieved

<table>
<thead>
<tr>
<th>Criteria</th>
<th>PFAPA</th>
<th>CAPS</th>
<th>FMF</th>
<th>TRAPS</th>
<th>MKD</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least seven out of eight:</td>
<td>Presence of at least two of five*:</td>
<td>At least six out of nine:</td>
<td>Score ≥5 points:</td>
<td>Presence of at least three of six:</td>
<td></td>
</tr>
<tr>
<td>Presence</td>
<td>Urticarial rash.</td>
<td>Presence</td>
<td>Eastern Mediterranean ethnicity.</td>
<td>Age at onset &lt;1 years.</td>
<td></td>
</tr>
<tr>
<td>Duration of episodes, 3–6 days.</td>
<td>Sensorineural hearing loss.</td>
<td>Duration of episodes, 1–3 days.</td>
<td>Painful lymph nodes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absence</td>
<td>Arthritis.</td>
<td>Absence</td>
<td>Maculopapular rash.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diaphragna.</td>
<td>Absence</td>
<td>Aphthous stomatitis.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chest pain.</td>
<td>Urticarial rash.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin rash.</td>
<td>Maculopapular rash.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthritis.</td>
<td>Painful lymph nodes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity: 0.97
Specificity: 0.93
Accuracy: 0.99

*Modified by Kuenemmer-Deschner et al.14 See online supplementary table 6 for glossary. See table 2 for prerequisite criteria.

CAPS, cryopyrin-associated periodic syndromes; FMF, familial Mediterranean fever; MKD, mevalonate kinase deficiency; PFAPA, periodic fever, aphthosis, pharyngitis and adenitis; PRINTO, pediatric rheumatology international trial organization; TRAPS, tumour necrosis factor receptor-associated periodic fever syndrome.
The same approach was followed for the other HRFs (CAPS, TRAPS, MKD), leading to the selection of criteria with genetic and clinical variables (number 32 for CAPS, number 46 for TRAPS, number 37 for MKD) (table 2, online supplementary figures 6-8). As per the process to establish FMF criteria, a set of purely clinical criteria (ie, without genetic results) was also selected for each HRF, namely definitions number 20 and number 1 for MKD and TRAPS, respectively (table 3). For CAPS classification, the experts reached consensus on a modified version of recently published criteria.\textsuperscript{14} The performance of the original Kummerle criteria (using two out of six criteria) in the context of the present study population displayed a good sensitivity (0.91), but a low specificity (0.80).\textsuperscript{14} In contrast, when the variable ‘musculoskeletal pain’ was excluded, a higher specificity (0.94, with a sensitivity of 0.80) was achieved, if two out of five criteria are present (table 3).

The most severe form of CAPS, chronic infantile neurological cutaneous articular (CINCA)/neonatal onset multisystemic inflammatory disorder (NOMID), displays a chronic rather than a recurrent disease course. Patients with CINCA were not included in the validation process described above. However, when the new genetic and clinical CAPS criteria were tested in a separate set of 70 patients with CAPS with chronic disease course enrolled in the Eurofever Registry, the sensitivity was 100% for the genetic and clinical criteria and 80% for the clinical criteria (not shown).

Clinical classification criteria for PFAPA were discussed between the 25 clinical panelists distributed in two tables (no geneticists). After discussion (online supplementary figure 8), definition number 13 (clinical variables with both positive and negative association) was chosen (table 3). During the Consensus Conference, the panel agreed on a few suggested mandatory criteria that should be fulfilled in all the patients before the application of the new classification criteria (table 3) with a consensus of 100% for point 1 and 96% for point 2.

Globally, convergent validity among the three tables of experts was obtained for the genetic and clinical definitions of FMF and CAPS, whereas for all the other definitions a partial convergent validity (agreement in two out of three tables) was reached, with the need for a final plenary voting session (online supplementary figures 4-8 and online supplementary table 8).

Cross-validation of the final classification criteria

The ability of the new classification criteria to discriminate among the different recurrent fevers and uFR was further tested in the validation data set of 1018 patients extracted from the Eurofever Registry (online supplementary table 9) using as reference standard for each patient the diagnosis given by the treating physician. In the last column of table 4, the genotype (score 0=negative/not done; score 1=not confirmatory; score 2=confirmatory) of patients not identified by the clinical criteria for HRF is reported. Notably, almost all the patients not classified by the clinical and genetic criteria displayed a negative or not confirmatory genotype (table 4). The performance of the new classification criteria (either clinical and genetic or clinical only) was generally superior (accuracy ranging from 0.81 to 0.98; table 4) to those already available in the literature (accuracy 0.56–0.94) (online supplementary table 10).

**DISCUSSION**

The present study provides new evidence-based classification criteria for the four ‘classical’ HRF (FMF, MKD, TRAPS, CAPS) and PFAPA, incorporating combined consensus expertise of clinicians and geneticists with statistical analyses in real patients from the Eurofever Registry. At variance with past work\textsuperscript{15,16} these new classification criteria combine genetic and clinical variables to overcome the paradox of the absence of a role of the molecular analysis for the proper identification of patients affected by these (mainly) genetic conditions. As defined by the American College of Rheumatology, the proposed classification criteria have selected clinical and genetic findings able to identify the defined diseases and separate from other confounding autoinflammatory conditions.\textsuperscript{18} 19 Although these criteria may at times be helpful in clinical practice, they are explicitly not meant to be employed as diagnostic criteria. The advent of the so-called next-generation sequencing era resulted on one side to an increased availability of the molecular analysis at reduced costs but might often lead to difficulties in the proper interpretation of this large set of bioinformatic data. In fact, besides the identification of clearly pathogenic variants, in many circumstances (ie, low penetrance variants or variants of unknown significance, monolacile variants in autosomal recessive diseases, presence of variants in more than one gene) the genetic results are not unequivocal and should be placed in the context of a pertinent clinical setting. In these latter cases, the classification of the patient is usually problematic, as clearly shown in the process of patients’ validation in this study. For these reasons, the panel decided to introduce a distinction between a confirmatory (namely, surely or likely pathogenic variants) and not confirmatory (variants of unknown significance) genetic test. For the daily use of the new criteria, a parallel consensus classification effort by the genetic subcommittee of the INSAD project has established the pathogenicity of each currently known variant associated to HRF.\textsuperscript{40} A differential approach for the interpretation of the biallelic variants was chosen for the two autosomal recessive diseases, namely MKD and FMF. MKD is caused by loss-of-function mutations in the MVK gene. The panel excluded the possibility of classifying a patient as an MKD in the absence of biallelic mutations of the MVK gene. Conversely, recent evidence has clarified that FMF is secondary to gain-of-function mutations of the MEFV gene, with a clear dose effect.\textsuperscript{31,32} and therefore FMF could be classified with identification of either one or two pathogenic variants in exon 10 of MEFV in the presence of a consistent clinical phenotype. The same possibility was also considered for the two autosomal dominant diseases, CAPS and TRAPS, in the absence of confirmatory phenotype. In patients carrying variants of unknown pathogenic significance (such as R92Q and P46L for TNFRSF1A, or V198M for NLRP3),\textsuperscript{33–36} only the presence of some very specific clinical variables would support the proper disease classification. In parallel with the elaboration of the definitive criteria that include genetic/clinical variables, the panel agreed on additional clinical criteria that should be used to (1) identify patients with recurrent fevers that would need to undergo genetic testing for molecular confirmation; (2) search for possible somatic mosaicism using NG in patients with a clear phenotype, but negative Sanger sequencing results; and (3) classify patients (eg, for epidemiological studies) even in those countries where routine genetic testing is not possible. For PFAPA, the contemporary evaluation of either positive (presence) and negative (absence) clinical variables yielded a much higher accuracy when compared with the classical modified Marshall’s criteria.\textsuperscript{46} Following the consensus meeting, the new sets of criteria were valid.
### Table 4  Performance of the new classification criteria to discriminate different recurrent fevers in the validation data set of patients extracted from the Eurofever Registry (N=1018)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Positive/Negative</th>
<th>OR</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>AUC</th>
<th>Patients not fulfilling the criteria</th>
<th>Number of patients not fulfilling the clinical criteria but satisfying the clinical/genetic criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAPS clinical + genetics</strong></td>
<td>TP: 98/1013</td>
<td>4490</td>
<td>0.72</td>
<td>1</td>
<td>0.96</td>
<td>0.86</td>
<td>27/38 pts, score 0 (71.05%)</td>
<td>10/38 pts, score 1 (26.32%) 1/38 pts, score 2 (2.63%)</td>
</tr>
<tr>
<td><strong>CAPS clinical</strong></td>
<td>TP: 82/925</td>
<td>220.3</td>
<td>0.77</td>
<td>0.99</td>
<td>0.96</td>
<td>0.88</td>
<td>3/25 pts, score 0 (13.04%)</td>
<td>8/25 pts, score 1 (34.78%) 12/25 pts, score 2 (52.17%)</td>
</tr>
<tr>
<td><strong>FMF clinical + genetics</strong></td>
<td>TP: 304/1010</td>
<td>1725.3</td>
<td>0.89</td>
<td>1</td>
<td>0.96</td>
<td>0.94</td>
<td>12/39 pts, score 0 (30.77%)</td>
<td>Score 0: 0/3  Score 1: 1/8  Score 2: 2/112</td>
</tr>
<tr>
<td><strong>FMF clinical</strong></td>
<td>TP: 283/940</td>
<td>82.4</td>
<td>0.85</td>
<td>0.94</td>
<td>0.91</td>
<td>0.89</td>
<td>3/50 pts, score 0 (6.12%)</td>
<td>Score 0: 0/3  Score 1: 1/8  Score 2: 2/1920</td>
</tr>
<tr>
<td><strong>MKD clinical + genetics</strong></td>
<td>TP: 45/1000</td>
<td>5209.1</td>
<td>0.74</td>
<td>1</td>
<td>0.98</td>
<td>0.87</td>
<td>2/16 pts, score 0 (12.5%)</td>
<td>Score 0: 0/3  Score 1: 1/8  Score 2: 2/116</td>
</tr>
<tr>
<td><strong>MKD clinical</strong></td>
<td>TP: 43/818</td>
<td>13.2</td>
<td>0.75</td>
<td>0.81</td>
<td>0.81</td>
<td>0.78</td>
<td>2/14 pts, score 1 (15.38%)</td>
<td>Score 1: 0/2  Score 2: 1/114</td>
</tr>
<tr>
<td><strong>TRAPS clinical + genetics</strong></td>
<td>TP: 73/1000</td>
<td>2526.9</td>
<td>0.74</td>
<td>1</td>
<td>0.97</td>
<td>0.87</td>
<td>6/26 pts, score 0 (23.08%)</td>
<td>Score 1: 0/2  Score 2: 2/116</td>
</tr>
<tr>
<td><strong>TRAPS clinical</strong></td>
<td>TP: 52/940</td>
<td>30.4</td>
<td>0.55</td>
<td>0.96</td>
<td>0.92</td>
<td>0.76</td>
<td>27/42 pts, score 1 (71.05%)</td>
<td>Score 1: 8/27  Score 2: 2/111</td>
</tr>
<tr>
<td><strong>PFAPA clinical</strong></td>
<td>TP: 149/1000</td>
<td>61.4</td>
<td>0.66</td>
<td>0.97</td>
<td>0.9</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CAPS, cryopyrin-associated periodic syndromes; FMF, familial Mediterranean fever; FN, false negative; FP, false positive; MKD, mevalonate kinase deficiency; PFAPA, periodic fever, aphthosis, pharyngitis and adenitis; pts, patients; TN, true negative; TP, true positive; TRAPS, tumour necrosis factor receptor-associated periodic fever syndrome;
further validated in a large group of additional patients from the Eurofever Registry, showing a very high specificity when compared with previous literature criteria. As noticed, most of the diagnoses refuted by the new criteria had been in patients with either non-confirmatory or negative genetic tests results. It is therefore conceivable that the present new criteria will be more stringent in the classification of patients, by excluding a substantial proportion of patients carrying variants of unknown origin. The classification criteria we propose are accurate for the discrimination of one form of autoinflammation from another in the context of the six conditions considered herein, but very much have to be applied judiciously, after careful consideration of confounding diseases, as highlighted in table 2. These classification criteria are therefore intended for use for clinical, epidemiological or translational studies, but not for routine diagnostic purposes in individual patients.17 That said, the purely clinical classification criteria might guide molecular testing approaches for individual cases, although this point requires future validation. One possible limitation of the present study is the lack of comparison groups including possible confounding conditions (chronic infections, neoplasms, immune deficiencies, autoimmune disease and metabolic diseases) presenting sometimes with a recurrent disease course. In daily practice confounding diseases with a true recurrent disease course are rather infrequent outside HRF and PFAPA, while the most challenging group of confounding conditions are the large emerging group of patients with uRF, many of whom may have a true monogenetic cause other than the four genetic diseases considered herein. For these reasons, the different HRFs have been used as controls for each individual condition with PFAPA and uRF as genetically negative controls. The panel of experts unanimously decided that the presence of elevation of acute phase reactants during disease flares (recorded at least in one occasion) should be considered as mandatory preliminary criterion for the use of the new classification criteria.14 Some other relevant pathognomonic laboratory examinations, such as urinary mevalonic acid in MKD, were not available in the Eurofever Registry, probably reflecting the fact that it is not widely available for testing routinely. As such the panel recommended the importance for the diagnostic workup, for example, with intracellular MVK enzyme activity and/or urinary mevalonic acid in MKD,18 particularly for patients with convincing phenotypes but non-confirmatory genotype for MKD. Similarly, the response to some specific treatments (such as colchicine in FMF or anti-interleukin (IL)-1 in CAPS) or ethnic background (for FMF) could certainly be considered as additional elements to be considered in daily practice, especially for patients with non-confirmatory genotype, but are not good discriminators of the different forms of autoinflammatory disease considered herein. In conclusion, the present work allowed the proposal of novel evidence-based classification criteria for HRF and PFAPA with a high specificity. The use of these classification criteria is highly recommended for inclusion of patients in translational and clinical studies, including clinical trials, and should not be misused as diagnostic criteria.17 The possible identification of new genetic entities in the heterogeneous group of undefined periodic fevers could require an update of the criteria in the future.

Author affiliations

1 UOSD Centro Malattie Autoinflammatorie e Immunodeficienze, IRCCS Istituto Giannina Gaslini, Genoa, Italy
2 Department of Paediatrics, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland
3 University Hospital of Geneva, Geneva, Switzerland
4 Clinica Pediatrica e Reumatologia, IRCCS Istituto Giannina Gaslini, Genova, Italy
5 Department of Pediatrics, Ospedale Regionale di Bellinzona e Valli, Bellinzona, Switzerland
6 Department of Health Sciences (DISSAL), University of Genova, Genova, Italy
7 Inflammatory Disease Section, National Human Genome Research Institute, Bethesda, Maryland, USA
8 Hospital Sant Joan de Déu, Universitat de Barcelona, Barcelona, Catalunya, Spain
9 Immunology Department, CDB, Hospital Clinic/DIABAPS, Barcelona, Spain
10 Division of Intramural Research, NIH-NIAID, Bethesda, Maryland, USA
11 Rheumatology Unit, Hadassah-Hebrew University Hospital, Ein Kerem, Jerusalem, Israel
12 Institute of Child Health, University College London, London, UK
13 Department of Medical Sciences, University of Siena, Siena, Italy
14 Genetica Medica, IRCCS Istituto Giannina Gaslini, Genoa, Italy
15 Division of Rheumatology, IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy
16 Pediatric Rheumatology, Harvard University Children’s Hospital, Boston, Massachusetts, USA
17 Unit of Pediatric Rheumatology, Western University Children’s Hospital, London, Ontario, Canada
18 Department of Pediatrics, Wilhelmina Kinderziekenhuis, Utrecht, The Netherlands
19 Translational Autoinflammatory Disease Studies Unit, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA
20 Division of Rheumatology, Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey
21 National Referral Centre for Auto-Immuno-inflammatory Diseases and Inflammatory Amyloidosis, CEREMAITA, Versailles Hospital, Le Chesnay (Paris), France
22 Department of Pediatrics, University of California at San Diego, San Diego, California, USA
23 Pediatric Rheumatology, University of California San Diego, La Jolla, San Diego, California, USA
24 Pediatric Rheumatology, Charite University Medicine Berlin, Berlin, Germany
25 Department of Paediatric Rheumatology and CEREMAI, Hôpital de Bicêtre, National Reference Centre for Auto-Immuno-inflammatory Diseases, Le Kremlin-Bicêtre, Paris, France
26 Department of Pediatrics, University Hospital Tuebingen, Tuebingen, Germany
27 Division of Medicine, UCL Medical School, Royal Free Campus, National Amyloidosis Centre, London, UK
28 Unit of Pediatric Rheumatology, Hospital for Sick Children, Toronto, Ontario, Canada
29 Sheba Medical Center, Tel Aviv, Israel
30 Fondazione IRCCS Policlinico San Matteo, Centro per lo Studio e la Cura delle Aminolisi Sistemiche, Pavia, Italy
31 Department of Pediatrics, Hacettepe University, Ankara, Turkey
32 Servizio di Immunologia/Reumatologia, Hospital de Pediatria Juan P Garrahan, Buenos Aires, Argentina
33 Heller Institute of Medical Research, Sheba Medical Center, Ramat Gan, Israel
34 Department of General Internal Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands
35 Department of Allergology, Rheumatology and Clinical Immunology, University Children’s Hospital, Ljubljana, Slovenia
36 National Referral Centre for Auto-Immuno-inflammatory Diseases and Inflammatory Amyloidosis, CEREMAITA, Centre Hospitalier Regional Universitaire de Montpellier, Montpellier, France
37 Department of Pediatrics, Meir Medical Centre, Kfar Saba, Israel
38 Department of Pediatrics, Tel Aviv University, Tel Aviv, Israel
39 Department of Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands
40 Department of Pediatrics, Universitätsklinikum Münster, Münster, Germany
41 IRCCS Istituto Giannina Gaslini, Genoa, Italy
42 Ospedale Policlinico San Martino IRCCS, Genoa, Italy
43 Clinica Pediatrica e Reumatologia, PRINTO, IRCCS Istituto Giannina Gaslini, Genoa, Italy

Correction notice This article has been corrected since it published Online First. Table 4 has been amended.

Acknowledgements The authors are in debt with Dr Daniel Lovell for his precious support during the Consensus Conference and all PRINTO’s staff for the precious technical support.

Contributors MG, MH and NR coordinated the study, analysed the data and drafted the manuscript. SF, FV and CG analysed the data. FB and MPS performed the technical support.

Funding The project has been funded by E-Rare-3 project (NISAIAD, grant 003037603). Eurofever was supported by the Executive Agency For Health and Consumers (EAHC, Project No 2007332) and by Istituto G Gaslini. Novartis and SOBI provided unrestricted grants for the Consensus Conference.

Competing interests None declared.
REFERENCES


