Definitive classification criteria for autoinflammatory recurrent fevers

Marco Gattorno[1*] MD, Michael Hofer[2, 3*] MD, Silvia Federici[4] MD, PhD, Federica Vanoni[5] MD, Francesca Bovis[6] MSc, PhD, Ivona Aksentijevich[7] MD, Jordi Anton[8] MD, PhD, Juan Ignacio Arostegui[9] MD, Karyl S. Barron[10] MD, Eldad Ben-Chetrit[11] MD, Paul Brogan[12] MD, Luca Cantarini[13] MD, Isabella Ceccherini[14] PhD, Fabrizio De Benedetti[15] MD, Fatma Dedeoglu[16] MD, Erkan Demirkaya[17] MD, Joost Frenkel[18] MD, Raphaela Goldbach-Mansky[19] MD, Ahmet Gul[20] MD, Veronique Hentgen[21] MD, Hal M. Hoffman[22] MD, Tilmann Kallinich[23] MD, Isabelle Koné-Paut[24] MD, Jasmin Kuemmerle-Deschner[25] MD, Helen J. Lachmann[26] MD, Ronald M. Laxer[27] MD, Avi Livneh [28] MD, Laura Obici[29] MD, Seza Ozen[30] MD, Dorota Rowczenio[26] MD, Ricardo Russo[31] MD, Yael Shinar[28] MD, Anna Simon[32] MD, Natasa Toplak[33] MD, Isabelle Touitou[34] MD, Yosef Uziel[35] MD, Marielle van Gijn[36] PhD, Dirk Foell[37] MD, Claudia Garassino[4], MD, Daniel L. Kastner[10], Alberto Martini[38] MD, Maria Pia Sormani[6, 39,] PhD, Nicolino Ruperto[40] MD, MPH, for the Eurofever Registry and the Paediatric Rheumatology International Trials Organisation (PRINTO).

[1] IRCCS Istituto Giannina Gaslini, UOSD Centro Malattie Autoinfiammatorie e Immunodeficienze, Genoa, Italy
[2] Unité Romande d'Immuno-Rhumatologie Pediatrique, CHUV, University of Lausanne, Lausanne, Switzerland
[3] Unité Romande d'Immuno-Rhumatologie Pediatrique, University Hospital of Geneva, Geneva, Switzerland
[4] Istituto Giannina Gaslini, Clinica Pediatrica e Reumatologia, Genoa, Italy
[5] Ospedale Regionale di Bellinzona e Valli, Dipartimento di Pediatria, Bellinzona, Switzerland
[6] University of Genova, Department of Health Sciences (DISSAL), Section of Biostatistics, Genoa, Italy
[7] National Human Genome Research Institut, Inflammatory Disease Section, Bethesda, MD, United States
[8] Hospital Sant Joan de Déu, Universitat de Barcelona, Unidad de Reumatología Pediátrica, Barcelona, Spain
[9] Hospital Clínic-IDIBAPS, Department of Immunology, University of Barcelona, Barcelona, Spain
[10] NIH-NIAID, Division of Intramural Research, Bethesda, MD, United States
[11] Hadassah-Hebrew University Hospital, Ein Kerem, Rheumatology Unit, Jerusalem, Israel
[12] UCL GOS Institute of Child Health , London, United Kingdom
[13] University of Siena, Department of Medical Sciences, Surgery and Neurosciences, Rheumatology Unit, Siena, Italy.
[14] Istituto Giannina Gaslini, Genetica Medica, Genoa, Italy
[15] IRCCS Ospedale Pediatrico Bambino Gesù, Division of Rheumatology, Rome, Italy
[16] Harvard Medical School, Boston Children's Hospital, Division of Immunology, Rheumatology Program, Boston, MA, USA
[17] Western University Children's Hospital, London Health Sciences Centre, London, ON, Canada
[18] Wilhelmina Kinderziekenhuis, Department of Pediatric Immunology and Rheumatology, Utrecht, Netherlands
[19] National Institute of Allergy and Infectious Diseases, National Institutes of Health, Translational

Autoinflammatory Disease Studies, Laboratory of Clinical Immunology and Microbiology (LCIM), Bethesda, MD, United States
[20] Istanbul University, Faculty of Medicine, Department of Internal Medicine, Division of Rheumatology, Istanbul, Turkey
[21] National Referral Centre of Auto-Inflammatory Diseases and inflammatory amyloidosis, - CEREMAIA, department for pediatrics, Versailles Hospital, Le Chesnay (Paris), France
[22] University of California at San Diego, Rady Children's Hospital, San Diego, CA, United States
[23] Charité University Medicine Berlin, Pediatric Penumology and Immunology, Berlin, Germany.
[24] National Referral Centre of Auto-Inflammatory Diseases and inflammatory amyloidosis, CEREMAIA, CHU de Biĉetre, APHP, University of Paris Sud, Department of Pediatric Rheumatology, le Kremlin Biĉetre, France
[25] University Hospital Tuebingen, Department of Pediatrics and autoinflammation reference center Tuebingen, Division of Rheumatology, Germany
[26] Royal Free Campus, National Amyloidosis Centre, London, United Kingdom
[27] University of Toronto and The Hospital for Sick Children, Departments of Pediatrics and Medicine, Toronto, ON, Canada
[28] Sheba Medical Center, Heller Institute, Ramat Gan, Israel
[29] Fondazione IRCCS Policlinico San Matteo, Centro per lo Studio e la Cura delle Amiloidosi Sistemiche, Pavia, Italy
[30] Hacettepe University, Department of Pediatric Rheumatology, Ankara, Turkey
[31] Hospital de Pediatria Juan P. Garrahan, Servicio de Inmunologia/Reumatologia, Buenos Aires, Argentina
[32] Netherlands Radboud University Medical Center, Department of Internal Medicine, Radboudumc Expertisecenter for Immunodeficiency and Autoinflammation, Nijmegen, Netherlands
[33] University Children's Hospital, University Medical Center Ljubljana, Department of Allergology, Rheumatology and Clinical Immunology, Ljubljana, Slovenia
[34] Stem Cells, Cellular Plasticity, Regenerative Medicine and Immunotherapies, INSERM, Univ Montpellier, Department of Medical Genetics, Rare Diseases and Personalized Medicine, National Referral Centre of Auto-Inflammatory Diseases and inflammatory amyloidosis, CEREMAIA CHU Montpellier, Montpellier, France
[35] Meir Medical Centre, Pediatric Rheumatology Unit, Department of Pediatrics, Kfar Saba and Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
[36] University Medical Centre Utrecht, Department of Genetics, Utrecht, Netherlands
[37] Universitätsklinikum Münster, Klinik für Pädiatrische Rheumatologie und Immunologie, Münster, Germany
[38] IRCCS Istituto Giannina Gaslini, Genoa, Italy
[39] Ospedale Policlinico San Martino IRCCS, Genoa, Italy
[40] IRCCS Istituto Giannina Gaslini, Clinica Pediatrica e Reumatologia, PRINTO, Genoa, Italy

*Both Authors equally contributed to the study

Corresponding author: Marco Gattorno, MD, Centro Malattie Autoinfiammatorie e Immunodeficienze, Clinica Pediatrica e Reumatologia, Istituto G. Gaslini, Via G. Gaslini 5, 16147

Genoa, Italy. Telephone: +39-010-56362793 fax: +39-010-393324, e-mail: marcogattorno@gaslini.org

Key words: Classification criteria, inherited periodic fevers, PFAPA, TRAPS, mevalonate kinase deficiency, CAPS

Short running title: Novel classification criteria for inherited recurrent fevers and PFAPA

Abstract

Background. Different diagnostic and classification criteria are available for hereditary recurrent fevers (HRF): Familial Mediterranean fever (FMF), tumor necrosis factor receptor-associated periodic fever syndrome (TRAPS), mevalonate kinase deficiency (MKD), cryopyrin-associated periodic syndromes (CAPS) and for 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 the patients' diagnosis (consensus ≥80%). Step 3. Statistical analysis for the selection of the best candidate classification criteria. Step 4. Nominal Group Technique (NGT) consensus conference for the selection of the final classification criteria with 33 panelists, to discuss and select the final classification criteria. Step 5. Cross-sectional validation of final classification criteria.

Results. The panelists achieved consensus to classify 281/360 (78%) patients (32 CAPS, 36 FMF, 56 MKD, 37 PFAPA, 39 TRAPS, 81 undefined recurrent fever). Consensus was reached for 2 sets of criteria for each HRF, one including genetic and clinical variables, the other with clinical variables only, plus new criteria for PFAPA. The 4 HRF criteria demonstrated sensitivity 0.94-1; and specificity 0.95-1; for PFAPA, criteria sensitivity and specificity were 0.97 and 0.93 respectively. Validation of these criteria in an independent dataset of 1018 patients demonstrated accuracy ranging from 0.81 to 0.98.

Conclusion. Eurofever proposes a 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) have led to a completely new line of research. The SAIDs are caused by exaggerated activation of the innate immune system, in the absence of high-titer autoantibodies or antigen-specific T-cells[1;1]. Recurrent (or periodic) fevers are characterized 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 fevers (HRF). Familial Mediterranean fever (FMF) is caused by mutations of *MEFV* [2;3]; mevalonate kinase deficiency (MKD), by mutations of the mevalonate kinase gene (*MVK*) [4;5]; tumor necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS), by mutations of type I TNF receptor (*TNFSRF1A*)[6]; cryopyrin-associated periodic syndromes (CAPS), by mutations of *NLRP3* [7] [8]. More common forms of recurrent fever syndromes include PFAPA (periodic fever, aphthosis, pharyngitis and adenitis) syndrome, which is a multifactorial disorder [9]. So far, several clinical diagnostic and classification criteria have been proposed for HRF [10-15] and PFAPA [9;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 day-to-day clinical practice, and the two terms are often (wrongly) used interchangeably [ref]. 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 (for example 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, whilst excluding important disease mimics based on derivation and validation in cohorts that include important disease mimics. Consequently, the American College of Rheumatology (ACR) emphasised that classification criteria typically "include manifestations that are characteristics of the disease in question and occur with less frequency or are absent in other conditions"; and diagnostic criteria "tend to focus on listing or determining the combination of findings that need to be present in order to be certain that a particular disease is present". As such, classification criteria cannot be used as diagnostic criteria.

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 dataset of patients in the Eurofever registry.

Methods

A multistep process using consensus formation techniques (Delphi and Nominal Group Technique) and statistical evaluations on real patients was used to develop and test the classification criteria[17] (Supplementary Figure 1 and supplementary material), based on a methodological framework used successfully in previous studies in rheumatology [18-23].

Step 1. Selection of clinical, laboratory and genetic candidate variables.

A panel of 162 international adult and pediatric 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[24] and PFAPA[25].

Step 2. Classification of patients from the Eurofever registry.

After selection (Supplementary Figure 2), a random sample of 360 patients, 60 patients for each disease (FMF, TRAPS, MKD, CAPS, PFAPA, and undefined recurrent fevers, uRF) was selected from the Eurofever registry[26] . The inclusion criteria for the enrollement in the registry have been previously described [26]

Twenty-five international experienced clinicians/researchers and 8 geneticists (total 33 panellists) in the field of SAID blinded on patients original diagnosis were invited to participate in a multi-round secured web process to classify each of the 360 patients into one of six mutually exclusive diseases[27]. 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 Supplementary material) foresaw the following steps:

1) Evaluation through an 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.

2) Computer generation of more than 30,000 new candidate sets of classification criteria through linear combinations of genetic and clinical variables with improper linear modelling. Additionally, 11 sets of criteria were derived from the literature [9-16] , or proposed by members of the panel based on their expertise.

3) Identification of the top performing criteria through ranking according to the Akaike Information Criterion (AIC), with best model having the lowest AIC.

Step 4. Nominal Group Technique (NGT) consensus conference for the selection of the final classification criteria.

The Consensus Conference was held in Genoa, Italy, on March 18-21, 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 upon 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 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, post-consensus, 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 Supplementary Table 1.

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

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

Familial Mediterranean fever (FMF); Cryopyrin-associated periodic syndromes (CAPS); Mevalonate kinase deficiency (MKD); Tumor necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS); Periodic fever, aphthosis, pharyngitis and adenitis (PFAPA); Undefined recurrent fevers (uRF).

A total of 100 different genotypes were reported in the 360 patients included in the classification process as reported in Supplementary Table 2.

Nine CAPS and 2 TRAPS patients 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 PFAPA patients and 14 uRF patients (Supplementary Table 3).

Step 2. Classification of patients from the Eurofever registry

In the first two rounds evaluation of clinical data by clinicians (blinded to genetic analysis) resulted in consensus ≥ 80% for a total of 216/360 (60%) of patients (Figure 1); consensus was reached in 51 MKD, 43 TRAPS, 29 FMF, 29 CAPS, 26 PFAPA and 38 URF patients. Similarly evaluation of demographic and genetic data by geneticists (blinded to clinical data) in two separate rounds reached consensus on 319/360 (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 281/360 (78%) as follows: 56 (95%) MKD, 39 (76%) TRAPS, 37 (70%) PFAPA, 36 (71%) FMF, 32 (63%) CAPS, and 81 (85%) uRF (Figure 1, Supplementary Table 4). K (concordance coefficient) agreement between the panel reference standard classification and the original patients 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.

Step 3. Statistical analysis for the selection of the best classification criteria

The top variables arising from Step 1 (see methods section) were included in an 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 Supplementary Table 5 with the related OR and 95% CI. The strategy for the classification of the genotypes is described in Supplementary Table 6.

A total of 198 over >30,000 possible new sets of classification criteria (available upon request, 50 for CAPS, 45 for FMF, 44 for TRAPS, 32 for MKD and 22 for PFAPA) were retained, based on their AIC, for further evaluation at the consensus conference, together with 11 criteria from the literature (Supplementary Figure 4).

Step 4. Nominal Group Technique (NGT) consensus conference for the selection of the final classification criteria.

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 on Supplementary Table 7).

The first disease discussed was FMF. After multiple voting sessions, all 3 tables of experts, which worked independently from each other, showed a complete convergent validity selecting the same top definition number 38 (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 the indication for molecular analysis or as classification criteria in case genetic testing is not locally available (Supplementary Figure 4, session B). To this aim, definition n. 12, including clinical variables with both positive and negative association with the disease, was chosen (Table 3).

The same approach was followed for the other HRFs (CAPS, TRAPS, MKD), leading to the selection of criteria with genetic and clinical variables (n. 32 for CAPS, n. 46 for TRAPS, n. 37 for MKD) (Table 2, Supplementary Figures 6-8). As per the process to establish FMF criteria, a set of purely clinical criteria (i.e. without genetic analyses) was also selected for each HRF; namely, definitions n. 20 and n.1 for MKD and TRAPS, respectively (Table 3). For CAPS classification, the experts reached consensus on a modified version of recently published criteria [14]. The performance of the original Kummerle criteria (using 2 out of 6 criteria) in the context of the present study population displayed a good sensitivity (0.91), but a low specificity (0.80) [14]. In contrast, when the variable “musculoskeletal pain” was excluded, a higher specificity (0.94, with a sensitivity of 0.80) was achieved, if 2 out of 5 criteria are present (Table 3). The most severe form of CAPS,

CINCA/NOMID, displays a chronic rather than a recurrent disease course. CINCA patients 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 CAPS patients with chronic disease course enrolled in the Eurofever registry, the sensitivity was 100% for 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, since this is not a genetic disease). After discussion (Supplementary Figure 8), definition n. 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.

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

A patient with:

1) Evidence of elevation of acute phase reactants (ESR or CRP or SAA) in correspondence to the clinical flares
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 following criteria are met:

CAPS	FMF	TRAPS	MKD
Presence of a *confirmatory NLRP3 genotype** and **at least 1** among: • Urticarial rash • Red eye (conjunctivitis, episcleritis, uveitis) • Neurosensorial hearing loss **OR** Presence of *not confirmatory NLRP3 genotype*°* and **at least 2** among: • Urticarial rash • Red eye (conjunctivitis, episcleritis, uveitis) • Neurosensorial hearing loss	Presence of *confirmatory MEFV genotype** and **at least** 1 among: • Duration of episodes 1-3 days • Arthritis • Chest pain • Abdominal pain **OR** Presence of *not confirmatory MEFV* genotype^ and **at least 2** among: • Duration of episodes 1-3 days • Arthritis • Chest pain • Abdominal pain	Presence of *confirmatory TNFRSF1A genotype** and **at least** 1 among: • Duration of episodes ≥ 7 days • Myalgia • Migratory rash • Periorbital oedema • Relatives affected **OR** Presence of a *not confirmatory TNFRSF1A genotype*°* and **at least 2 among**: • Duration of episodes ≥ 7 days • Myalgia • Migratory rash • Periorbital oedema • Relatives affected	Presence of a *confirmatory MVK genotype** and **at least 1** among: • Gastrointestinal symptoms • Cervical lymphadenitis • Aphthous stomatitis
Sensitivity: 1 **Specificity: 1** **Accuracy: 1**	**Sensitivity: 0,94** **Specificity: 0,95** **Accuracy: 0,98**	**Sensitivity: 0,95** **Specificity: 0,99** **Accuracy: 0,99**	**Sensitivity: 0,98** **Specificity: 1** **Accuracy: 1**

* 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.

^ *in trans* compound heterozygous for one pathogenic *MEFV* variants and one VUS; biallelic VUS; or heterozygous for one pathogenic *MEFV* variant.

See Suppl.TableS6 for glossary. Familial Mediterranean fever (FMF); Cryopyrin-associated periodic syndromes (CAPS) ; Mevalonate kinase deficiency (MKD);Tumor necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS)

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

PFAPA	CAPS	FMF	TRAPS	MKD
At least **7 out of 8**: *Presence* • Pharyngotonsillitis • Duration of episodes 3-6 days • Cervical lymphadenitis • Periodicity *Absence* • Diarrhoea • Chest pain • Skin rash • Arthritis	Presence of at least **2 of 5***: • Urticarial rash • Cold/stress triggered episodes • Sensorineural hearing loss • Chronic aseptic meningitis • Skeletal abnormalities (epiphyseal overgrowth/frontal bossing)	At least **6 out of 9**: *Presence* • Eastern Mediterranean ethnicity • Duration of episodes 1-3 days • Chest pain • Abdominal pain • Arthritis *Absence* • Aphthous stomatitis • Urticarial rash • Maculo-papular rash • Painful lymph nodes	**Score ≥ 5 points:** *Presence* • Fever>=7 days (2 points) • Fever 5-6 days (1 point) • Migratory rash (1point) • Periorbital edema (1 point) • Myalgia (1 point) • Positive family history (1 point) *Absence* • Aphthous stomatitis (1 point) • Pharyngotonsillitis (1 point)	Presence of **at least 3 of 6**: • Age at onset <1 years • Gastrointestinal symptoms • Painful lymph nodes • Aphthous stomatitis • Triggers • Maculo-papular rash
Sensitivity: 0,97 **Specificity: 0,93** **Accuracy: 0,99**	**Sensitivity: 0,80** **Specificity: 0,91** **Accuracy: 0,85**	**Sensitivity: 0,91** **Specificity: 0,92** **Accuracy: 0,97**	**Sensitivity: 0,87** **Specificity: 0,92** **Accuracy: 0,96**	**Sensitivity: 0,91** **Specificity: 0,82** **Accuracy: 0,92**

* Modified by Kuemmerle-Deschner et al (ref 14). See Suppl.TableS6 for glossary and Suppl. See Table 2 for pre-requisite criteria
Familial Mediterranean fever (FMF); Cryopyrin-associated periodic syndromes (CAPS) ; Mevalonate kinase deficiency (MKD);Tumor necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS); Periodic fever, aphthosis, pharyngitis and adenitis (PFAPA);

Globally, convergent validity among the 3 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 2 out of 3 tables) was reached, with the need for a final plenary voting session (Supplementary Figures 4-8 and Supplementary Table 8).

Cross-validation of 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 dataset of 1018 patients extracted from the Eurofever registry (Supplemenatry 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) (Supplemenatry Table 10).

Table 4. Performance of the new classification criteria to discriminate different recurrent fevers from control conditions in the validation dataset of patient extracted from the Eurofever registry (N = 1018).

CAPS clinical + genetics	TP: 98/1013 TN: 877/1013 FP: 0/1013 FN: 38/1013	4490	0.72	1	0.96	0.86	27/38 pts score 0 (71,05%) 10/38 pts score 1 (26,32%) 1/38 pts score 2 (2,63%)	
CAPS clinical	TP: 82/925 TN: 806/925 FP: 12/925 FN: 25/925	220,3	0.77	0.99	0.96	0.88	3/25 pts score 0 (13,04%) 8/25 pts score 1 (34,78%) 12/25 pts score 2 (52,17 %)	Score 0: 0/3 Score 1: 1/8 Score 2: 11/12
FMF clinical + genetics	TP: 304/1010 TN: 664/1010 FP: 3/1010 FN: 39/1010	1725,3	0.89	1	0.96	0.94	12/39 pts score 0 (30,77%) 26/39 pts score 1 (66,67%) 1/39 pts score 2 (2,56%)	
FMF clinical	TP: 283/940 TN: 568/940 FP: 39/940 FN: 50/940	82,4	0.85	0.94	0.91	0.89	3/50 pts score 0 (6,12%) 26/50 pts score 1 (53,06%) 20/50 pts score 2 (40,82 %)	Score 0: 0/3 Score 1: 8/26 Score 2: 19/20
MKD clinical + genetics	TP: 45/1005 TN: 944/1005 FP: 0/1005 FN: 16/1005	5209,1	0.74	1	0.98	0.87	2/16 pts score 0 (12,5%) 14/16 pts score 1 (87,5%)	
MKD clinical	TP: 43/818 TN: 617/818 FP: 144/818 FN: 14/818	13,2	0.75	0.81	0.81	0.78	2/14 pts score 1 (15,38%) 11/14 pts score 2 (84,62 %)	Score 1: 0/2 Score 2: 10/11
TRAPS clinical + genetics	TP: 73/1000 TN: 900/1000 FP: 1/1000 FN: 26/1000	2526,9	0.74	1	0.97	0.87	6/26 pts score 0 (23,08%) 20/26 pts score 1 (76,92%)	
TRAPS clinical	TP: 52/940 TN: 813/940 FP: 33/940 FN: 42/940	30,4	0.55	0.96	0.92	0.76	27/42 pts score 1 (71,05%) 11/42 pts score 2 (28,95 %)	Score 1: 8/27 Score 2: 11/11
PFAPA clinical	TP: 149/1001 TN: 752/1001 FP: 24/1001 FN: 76/1001	61,4	0.66	0.97	0.9	0.82		

TP: true positive, TN: true negative; FP: false positive; FN: false negative; OR : Odd ratio; AUC: area under the curve. For explanation of the score 0-1-2, see supplementary figure S4.

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 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 ACR (refs) the proposed classification criteria have selected clinical and genetic findings able to identify the diseases and separate from other counfounding conditions. These criteria will facilitate the clinical research in the field of SAIDs; may also help the clinician in current clinical practice, but are not meant to be clinical criteria for proper diagnosis.

The advent of the so-called next generation sequencing (NGS) 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, beside the identification of clearly pathogenic variants, in many circumstances the genetic results are not univocal (i.e. low penetrance variants or variants of unknown significance, monoallelic variants in autosomal recessive diseases, presence of variants in more than one gene). 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 sub-committee of the INSAID project has established the pathogenicity of each currently known variant associated to HRF.[28]

A differential approach for the interpretation of the bi-allelic 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 a MKD in the absence of bi-allelic 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 [29;30] and therefore FMF could be classified with no genetic confirmatory test but in the presence of a consistent clinical phenotype.
The same possibility was also considered for the two autosomal dominant diseases (CAPS and TRAPS) carrying variants of unknown pathogenic significance (such as R92Q and P46L for *TNFRSF1A*, or V198M for *NLRP3*) [31-34] for whom only the presence of some very specific clinical variables would support the proper classification of the patients.

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 i) identify patients with recurrent fevers that would need to undergo genetic testing for molecular confirmation; ii) search for possible somatic mosaicism using next generation sequencing in patients with a clear phenotype, but negative Sanger sequencing results; iii) classify patients (for example 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 [16]. Following the consensus meeting the new sets of criteria were further validated in a large group of additional patients from the Eurofever registry, showing a very high specificity when compared to previous literature criteria. It was notable that most of the patients with unclassified disease despite application of the new criteria had non-confirmatory, or negative genetic testing . Such patients might therefore benefit from earlier molecular investigation with next-generation sequencing platforms extending genetic screening beyond the four genetic diseases considered herein.

The classification criteria we propose are accurate for the discrimination of one form of autoinflammation from another in the context of the 5 diseases 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 or epidemiological studies, but not for routine diagnostic purposes in individual patients. That said, the purely clinical classification criteria might guide molecular testing approaches for individual cases, although this point requires future validation.

The main limitation of present study is the aforementioned lack of disease control patients in the EuroFever database with important confounding conditions (chronic infections, neoplasms, immune deficiencies, autoimmune disease, and metabolic diseases) that can present with systemic inflammation and thus mimic HRF. Inclusion of such patients would definitely be required to develop diagnostic criteria, and in the context of these proposed classification criteria absence of such a disease control group serves as an important caveat against using these classification criteria for routine diagnostic purposes. A particularly 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 4 genetic diseases considered herein. More widespread application of next-generation sequencing increases the diagnostic yield for such patients, over that which can be achieved by targeted sequencing of just *MEFV*, *NLRP3*, *TNFRSF1A*, and *MVK*. Thus it is likely that in the future such uRF will be able to be classified more accurately.

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 intracellular mevalonate kinase activity, 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 work-up e.g. with intracellular MVK enzyme activity and/or urinary mevalonic acid in MKD [36], 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-IL-1 in CAPS) or ethnic background (for FMF) could certainly be considered as additional

diagnostic indication 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.

Acknowledgement: The project has been funded by E-rare-3 project (INSAID, 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 held in Genoa, March 18-21, 2017. 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.

Key messages

What is already known about this subject?

- Hereditary recurrent fever syndromes are genetic disorders secondary to mutations in genes involved in the innate immune response

- A number of classification or diagnostic criteria have been developed in the past. 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.

What does this study add?

- We developed and validate new evidence-based classification criteria for HRF and PFAPA, combining international expert consensus, statistical evaluation of real patients from a large dataset of patients in the Eurofever registry

- The new classification criteria combine clinical manifestations with genotype

How might this impact on clinical practice?

The use of these classification criteria is recommended for inclusion of patients in translational and clinical studies; but they cannot be used as diagnostic criteria

Legend to Figures

Figure 1. Flow chart of the consensus nominal group technique for classification of patients from the Eurofever registry. Familial Mediterranean fever (FMF); Cryopyrin-associated periodic syndromes (CAPS); Mevalonate kinase deficiency (MKD);Tumor necrosis factor (TNF) receptor-associated periodic fever syndrome (TRAPS); Periodic fever, aphthosis, pharyngitis and adenitis (PFAPA).

Reference List

(1) Masters SL, Simon A, Aksentijevich I, Kastner DL. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease (*). *Annu Rev Immunol* 2009; 27:621-668.

(2) The French FMF Consortium. A candidate gene for familial Mediterranean fever. *Nature Genetics* 1997; 17(1):25-31.

(3) The International FMF Consortium. Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial Mediterranean fever. *Cell* 1997; 90(4):797-807.

(4) Drenth JP, Cuisset L, Grateau G, Vasseur C, van de Velde-Visser SD, De Jong JG et al. Mutations in the gene encoding mevalonate kinase cause hyper-IgD and periodic fever syndrome. International Hyper-IgD Study Group. *Nat Genet* 1999; 22(2):178-181.

(5) Houten SM, Kuis W, Duran M, De Koning TJ, van Royen-Kerkhof A, Romeijn GJ et al. Mutations in MVK, encoding mevalonate kinase, cause hyperimmunoglobulinaemia D and periodic fever syndrome. *Nature Genetics* 1999; 22(2):175-177.

(6) McDermott MF, Aksentijevich I, Galon J, McDermott EM, Ogunkolade BW, Centola M et al. Germline mutations in the extracellular domains of the 55 kDa TNF receptor, TNFR1, define a family of dominantly inherited autoinflammatory syndromes. *Cell* 1999; 97(1):133-144.

(7) Hoffman HM, Mueller JL, Broide DH, Wanderer AA, Kolodner RD. Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome. *Nat Genet* 2001; 29(3):301-305.

(8) Aksentijevich I, Nowak M, Mallah M, Chae JJ, Watford WT, Hofmann SR et al. De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID): a new member of the expanding family of pyrin-associated autoinflammatory diseases. *Arthritis Rheum* 2002; 46(12):3340-3348.

(9) Marshall GS, Edwards KM, Butler J, Lawton AR. Syndrome of periodic fever, pharyngitis, and aphthous stomatitis. *J Pediatr* 1987; 110(1):43-46.

(10) Livneh A, Langevitz P, Zemer D, Zaks N, Kees S, Lidar T et al. Criteria for the diagnosis of familial Mediterranean fever. *Arthritis and Rheumatism* 1997; 40(10):1879-1885.

(11) Yalcinkaya F, Ozen S, Ozcakar ZB, Aktay N, Cakar N, Duzova A et al. A new set of criteria for the diagnosis of familial Mediterranean fever in childhood. *Rheumatology (Oxford)* 2009; 48(4):395-398.

(12) Sohar E, Gafni J, Pras M, Heller H. Familial Mediterranean fever. A survey of 470 cases and review of the literature. *Am J Med* 1967; 43(2):227-253.

(13) Simon A, van der Meer JWM, Vesely R, Myrdal U, Yoshimura K, Duys P et al. Approach to genetic analysis in the diagnosis of hereditary autoinflammatory syndromes. *Rheumatology* 2006; 45(3):269-273.

(14) Kuemmerle-Deschner JB, Ozen S, Tyrrell PN, Kone-Paut I, Goldbach-Mansky R, Lachmann H et al. Diagnostic criteria for cryopyrin-associated periodic syndrome (CAPS). *Ann Rheum Dis* 2016.

(15) Federici S, Sormani MP, Ozen S, Lachmann HJ, Amaryan G, Woo P et al. Evidence-based provisional clinical classification criteria for autoinflammatory periodic fevers. *Ann Rheum Dis* 2015; 74(5):799-805.

(16) Thomas KT, Feder HM, Lawton AR, Edwards KM. Periodic fever syndrome in children. *Journal of Pediatrics* 1999; 135(1):15-21.

(17) Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A et al. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken)* 2015; 67(7):891-897.

(18) Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997; 40(7):1202-1209.

(19) Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; 38(6):727-735.

(20) Ruperto N, Meiorin S, Iusan SM, Ravelli A, Pistorio A, Martini A. Consensus procedures and their role in pediatric rheumatology. *Curr Rheumatol Rep* 2008; 10(2):142-146.

(21) Ozen S, Pistorio A, Iusan SM, Bakkaloglu A, Herlin T, Brik R et al. EULAR/PRINTO/PRES criteria for Henoch-Schonlein purpura, childhood polyarteritis nodosa, childhood Wegener granulomatosis and childhood Takayasu arteritis: Ankara 2008. Part II: Final classification criteria. *Ann Rheum Dis* 2010; 69(5):798-806.

(22) Ravelli A, Minoia F, Davi S, Horne A, Bovis F, Pistorio A et al. 2016 Classification Criteria for Macrophage Activation Syndrome Complicating Systemic Juvenile Idiopathic Arthritis: A European League Against Rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. *Ann Rheum Dis* 2016; 75(3):481-489.

(23) Ravelli A, Minoia F, Davi S, Horne A, Bovis F, Pistorio A et al. 2016 Classification Criteria for Macrophage Activation Syndrome Complicating Systemic Juvenile Idiopathic Arthritis: A European League Against Rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. *Arthritis Rheumatol* 2016; 68(3):566-576.

(24) Federici S, Vanoni F, Ben-Chetrit E, Cantarini L, Frenkel J, Goldbach-Mansky R et al. An International Delphi Survey for the Definition of New Classification Criteria for Familial Mediterranean Fever, Mevalonate Kinase Deficiency, TNF Receptor-associated Periodic Fever Syndromes, and Cryopyrin-associated Periodic Syndrome. *J Rheumatol* 2018.

(25) Vanoni F, Federici S, Anton J, Barron KS, Brogan P, de BF et al. An international delphi survey for the definition of the variables for the development of new classification criteria for periodic fever aphtous stomatitis pharingitis cervical adenitis (PFAPA). *Pediatr Rheumatol Online J* 2018; 16(1):27.

(26) Toplak N, Frenkel J, Ozen S, Lachmann HJ, Woo P, Kone-Paut I et al. An international registry on autoinflammatory diseases: the Eurofever experience. *Ann Rheum Dis* 2012; 71(7):1177-1182.

(27) Ruperto N, Martini A. Networking in paediatrics: the example of the Paediatric Rheumatology International Trials Organisation (PRINTO). *Arch Dis Child* 2011; 96(6):596-601.

(28) Van Gijn ME, Ceccherini I, Shinar Y, Carbo EC, Slofstra M, Arostegui JI et al. New workflow for classification of genetic variants' pathogenicity applied to hereditary recurrent fevers by the International Study Group for Systemic Autoinflammatory Diseases (INSAID). *J Med Genet* 2018; 55(8):530-537.

(29) Chae JJ, Cho YH, Lee GS, Cheng J, Liu PP, Feigenbaum L et al. Gain-of-function Pyrin mutations induce NLRP3 protein-independent interleukin-1beta activation and severe autoinflammation in mice. *Immunity* 2011; 34(5):755-768.

(30) Federici S, Calcagno G, Finetti M, Gallizzi R, Meini A, Vitale A et al. Clinical impact of MEFV mutations in children with periodic fever in a prevalent western European Caucasian population. *Ann Rheum Dis* 2012; 71(12):1961-1965.

(31) Pelagatti MA, Meini A, Caorsi R, Cattalini M, Federici S, Zulian F et al. Long-term clinical profile of children with the low-penetrance R92Q mutation of the TNFRSF1A gene. *Arthritis Rheum* 2011; 63(4):1141-1150.

(32) Naselli A, Penco F, Cantarini L, Insalaco A, Alessio M, Tommasini A et al. Clinical Characteristics of Patients Carrying the Q703K Variant of the NLRP3 Gene: A 10-year Multicentric National Study. *J Rheumatol* 2016; 43(6):1093-1100.

(33) Rowczenio DM, Trojer H, Russell T, Baginska A, Lane T, Stewart NM et al. Clinical characteristics in subjects with NLRP3 V198M diagnosed at a single UK center and a review of the literature. *Arthritis Res Ther* 2013; 15(1):R30.

(34) Kuemmerle-Deschner JB, Verma D, Endres T, Broderick L, de Jesus AA, Hofer F et al. Clinical and Molecular Phenotypes of Low-Penetrance Variants of NLRP3: Diagnostic and Therapeutic Challenges. *Arthritis Rheumatol* 2017; 69(11):2233-2240.

(35) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010; 62(9):2569-2581.

(36) Ter Haar NM, Oswald M, Jeyaratnam J, Anton J, Barron KS, Brogan PA et al. Recommendations for the management of autoinflammatory diseases. *Ann Rheum Dis* 2015; 74(9):1636-1644.