DR. MONTSERRAT FERNANDEZ-RIVAS (Orcid ID : 0000-0003-1748-2328)

MR. ISMAEL GÓMEZ GARCÍA (Orcid ID : 0000-0001-9756-2373)

MR. ALEJANDRO GONZALO-FERNÁNDEZ (Orcid ID : 0000-0002-5835-1896)

DR. SABINE DÖLLE-BIERKE (Orcid ID : 0000-0002-3339-0709)

DR. RICCARDO ASERO (Orcid ID : 0000-0002-8277-1700)

PROF. FRÉDÉRIC DE BLAY (Orcid ID : 0000-0001-5678-2214)

MS. MAREEN DATEMA (Orcid ID : 0000-0003-2646-9467)

PROF. LK POULSEN (Orcid ID : 0000-0002-1730-847X)

PROF. MARGITTA WORM (Orcid ID : 0000-0002-3449-1245)

PROF. GRAHAM C ROBERTS (Orcid ID : 0000-0003-2252-1248)

DR. PAUL TURNER (Orcid ID : 0000-0001-9862-5161)

Article type : Original article

Development and validation of the Food Allergy Severity Score

Short title: Food Allergy Severity Score

Authors

Montserrat Fernández-Rivas, MD PhD[1], Ismael Gómez García, MSc[2], Alejandro Gonzalo-Fernández, MSc[2], Manuel Fuentes Ferrer, MD PhD[3], Sabine Dölle-Bierke, PhD[4], Guadalupe Marco-Martín, MD PhD[5], Barbara K. Ballmer-Weber, MD PhD[6], Riccardo Asero, MD[7], Simona Belohlavkova, MD[8], Kirsten Beyer, MD PhD[9], Frédéric de Blay, MD[10], Michael Clausen, MD[11], Mareen R. Datema, PhD[12], Ruta Dubakiene, MD[13],

Kate E.C. Grimshaw, PhD[14], Karin Hoffmann-Sommergruber, PhD[15], Jonathan O'B Hourihane, MD[16], Monika Jedrzejczak-Czechowicz , MD PhD[17], André C. Knulst, MD PhD[18], Tanya Kralimarkova, MD PhD[19], Thuy-My Le, MD PhD[18], Nikolaos G. Papadopoulos, MD PhD[20], Todor A. Popov, MD PhD[21], Lars K. Poulsen, PhD[22], Ashok Purohit, MD[23], Suranjith L.Seneviratne, MD[24], Angela Simpson, MD PhD[25], Atanasios Sinaniotis, MD[26], Mirjana Turkalji, MD PhD[27], Sonia Vázquez-Cortés, MD[5], Rosialzira N. Vera-Berrios, MD[2], Antonella Muraro, MD PhD[28], Margitta Worm, MD PhD[4], Graham Roberts, MD PhD[29], Ronald van Ree, PhD[30], Cristina Fernández-Pérez, MD PhD[31], Paul J. Turner, MD PhD[32], E.N. Clare Mills, PhD[33].

Manuscript Acceptance Date: 10-Oct-2021

Affiliations:

1 Allergy Department, Hospital Clínico San Carlos, Facultad de Medicina, Universidad Complutense (UCM), IdISSC, Madrid, Spain. ARADyAL.
2 Grupo de Investigación en Alergia, IdISSC, Madrid, Spain.
3 Unidad de Apoyo a la Investigación, Preventive Medicine Department, Hospital Clínico San Carlos, IdISSC, and Universidad Alfonso X El Sabio, Madrid, Spain.
4 Division of Allergy and Immunology, Department of Dermatology, Venerology and Allergology, Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany.
5 Allergy Department, Hospital Clínico San Carlos, IdISSC, Madrid, Spain. ARADyAL.
6 Department of Dermatology, University Hospital Zurich, Zurich, Switzerland; Clinic for Dermatology and Allergology, Kantonsspital St Gallen, St Gallen, Switzerland.
7 Ambulatorio di Allergologia, Clinica San Carlo, Paderno Dugnano, Milan, Italy.
8 Medical Faculty and Faculty Hospital Pilsen, Pilsen, Czech Republic.
9 Department of Pediatric Respiratory Medicine, Immunology and Intensive Care Medicine; Charité Universitätsmedizin Berlin; Berlin, Germany
10 Chest Diseases Department, Strasbourg University Hospital, Federation of Translational Medicine, University of Strasbourg, Strasbourg, France.
11 Children's Hospital and Department of Allergy, Landspitali, University Hospital, Reykjavik, Iceland.

12 Department of Experimental Immunology, and Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Center, Amsterdam, The Netherlands.
13 Medical Faculty, Vilnius University, Vilnius, Lithuania.
14 Department of Dietetics, Salford Royal NHS Foundation Trust, Salford, UK; Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK.
15 Department of Pathophysiology and Allergy Research, Medical University of Vienna, Austria.
16 Royal College of Surgeons in Ireland, and Children's Health Ireland at Temple Street, Dublin, Ireland
17 Department of Allergy and Immunology, Medical University of Lodz, Lodz, Poland.
18 Dept. Dermatology/Allergology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands.
19 Clinic of Allergy and Asthma, Medical University in Sofia, Sofia, Bulgaria.
20 Allergy Dpt, 2nd Pediatric clinic, University of Athens, Athens, Greece; Division of Infection, Immunity & Respiratory Medicine, University of Manchester, Manchester, UK.
21 University Hospital Sv. Ivan Rilski, Sofia, Bulgaria.
22 Allergy Clinic, Copenhagen University Hospital at Herlev-Gentofte, Copenhagen, Denmark.
23 Allergy Division, Chest Disease Department, University Hospital of Strasbourg, Strasbourg, France
24 Department of Clinical Immunology and Allergy, Central Manchester and Manchester Children's University Hospitals NHS Trust, Manchester, UK; Institute of Immunity and Transplantation, Royal Free Hospital and University College London, London, UK.
25 Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Centre, and Manchester University NHS Foundation Trust, Manchester, UK.
26 Allergy Department, 2nd Pediatric Clinic, University of Athens; Allergy Department, Sotiria Chest Diseases Hospital, Athens, Greece.

27 Srebrnjak Children's Hospital Zagreb, Catholic University of Croatia Zagreb, and Medical Faculty Osijek, J.J.Strossmayer University, Croatia
28 Food Allergy Referral Centre Veneto Region, Department of Women and Child Health, Padua General University Hospital, Padua, Italy.
29 NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton; Clinical and Experimental Sciences and Human Development in Health, Faculty of Medicine, University of Southampton, Southampton; and The David Hide Asthma and Allergy Research Centre, St Mary's Hospital, Isle of Wight, UK.
30 Departments of Experimental Immunology and of Otorhinoraryngology, Amsterdam University Medical Centers, Amsterdam, The Netherlands.
31 Preventive Medicine Department, Hospital Clínico San Carlos, IdISSC, Madrid; Preventive Medicine Department, Complejo Hospitalario Universitario Santiago de Compostela, Santiago de Compostela, Spain.
32 National Heart & Lung Institute, Imperial College London, London, UK.
33 Division of Infection, Immunity and Respiratory Medicine, Manchester Academic Health Sciences Centre, Manchester Institute of Biotechnology, University of Manchester, Manchester, UK.

Authors' ORCID

Montserrat Fernández-Rivas https://orcid.org/0000-0003-1748-2328
Ismael Gómez García https://orcid.org/0000-0001-9756-2373
Alejandro Gonzalo-Fernández https://orcid.org/0000-0002-5835-1896
Manuel Fuentes Ferrer https://orcid.org/0000-0002-5177-1441
Sabine Dölle-Bierke https://orcid.org/0000-0002-3339-0709
Guadalupe Marco Martín https://orcid.org/0000-0003-1754-3274
Barbara K. Ballmer-Weber https://orcid.org/0000-0002-4136-5036
Riccardo Asero https://orcid.org/0000-0002-8277-1700
Simona Belohlavkova https://orcid.org/0000-0003-2125-6614
Kirsten Beyer https://orcid.org/0000-0003-1859-0419
Frédéric de Blay https://orcid.org/0000-0001-5678-2214
Mareen R. Datema https://orcid.org/0000-0003-2646-9467
Kate E.C. Grimshaw https://orcid.org/0000-0003-3649-7963

Karin Hoffmann-Sommergurber https://orcid.org/0000-0002-8830-058X
Jonathan O'B Hourihane https://orcid.org/0000-0003-4997-9857
Monika Jedrzejczak-Czechowicz https://orcid.org/0000-0003-2292-2770
André C. Knulst https://orcid.org/0000-0002-1056-3179
Tanya Z. Kralimarkova https://orcid.org/0000-0001-7023-1574
Nikolaos G. Papadopoulos https://orcid.org/0000-0002-4448-3468
Todor A. Popov https://orcid.org/0000-0001-5052-5866
Lars K. Poulsen https://orcid.org/0000-0002-1730-847X
Suranjith L.Seneviratne https://orcid.org/0000-0002-6548-5673
Angela Simpson https://orcid.org/0000-0003-2733-6666
Mirjana Turkalj https://orcid.org/0000-0002-5339-861X
Antonella Muraro https://orcid.org/0000-0002-5026-5862
Margitta Worm https://orcid.org/0000-0002-3449-1245
Graham Roberts https://orcid.org/0000-0003-2252-1248
Ronald van Ree https://orcid.org/0000-0003-0767-0894
Cristina Fernández Pérez https://orcid.org/0000-0001-9853-6257
Paul J. Turner https://orcid.org/0000-0001-9862-5161
Clare Mills https://orcid.org/0000-0001-7433-1740

Corresponding author

Montserrat Fernández-Rivas, MD PhD

Allergy Dept., Hospital Clínico San Carlos

c/ Prof. Martin Lagos s/n; 28040 Madrid, Spain

Phone: + 34 913303010; Fax: +34 913303011

Email: mariamontserrat.fernandez@salud.madrid.org

Acknowledgments

The authors wish to thank Prof. Marek L. Kowalski (deceased June 22nd, 2021) for his contribution in the EuroPrevall and iFAAM projects; Audrey DunnGalvin, Lynn Regent, Angel Sánchez, and Sabine Schnadt for their participation in the discussions on severity

scoring within the iFAAM project, and Irene Serrano for the technical review of the FASS tool.

The development and evaluation of FASS was undertaken as part of the European Union funded project iFAAM (312147), and the subsequent validation was done within the ARADyAL Research Network (RD16/0006/0009) and the BIOGRIAL-SEVERAL (PI19/01095) project cofounded by the Instituto de Salud Carlos III, Ministry of Science of the Spanish Government and FEDER (European Regional Development Fund).

Authors' contributions

MFR lead the score development and validation, and wrote the manuscript; MFR, ENCM, AM and GR lead the iFAAM-EAACI collaboration; MFR, GMM, BBW, KB, MC, MRD, KECG, KHS, JOBH, LKP, AM, MW, GR, RVR, PJT, ENCM were involved in the expert consensus for the score development; IGG developped nFASS; IGG and AGF performed the statistical analysis under the guidance of MFF, CFP and MFR; IGG developped the software tool for scoring severity, and AGF revised and refined it; MFR, SDB, GMM, BBW, RA, SB, KB, FB, MC, MRD, RD, KECG, KHS, JOBH, MJC, ACK, TK, TML, NGP, TAP, LKP, AP, SLS, AS, AS, MT, SVC, RNVB, MW, GR, RVR, CFP, PJT and ENCM contributed to the acquisition of data of the different cohorts; PJT lead the MaxDiff survey of health professionals; all the authors critically revised the manuscript and approved the final version.

Conflicts of interest

Dr Fernández-Rivas reports research grant to her institution from EU FP7 for the iFAAM project related to the work reported in this manuscript; research grants to her institution from Spanish Government (MINECO, ISCIII), and Aimmune Therapeutics and Diater, consultancy fees from Aimmune Therapeutics, Novartis and SPRIM, fees for participation in DSMB from DBV, lecture fees from Aimmune Therapeutics, ALK, Allergy Therapeutics, Diater, GSK, HAL Allergy and Thermofisher Scientific, outside the submitted work.

Dr. Beyer reports grants/research supports from Aimmune Therapeutics, Danone/Nutricia/Milupa, DBV, Hipp, Hycor, Infectopharm and honoraria or consultation fees from Aimmune Therapeutics, Bencard, Danone/Nutricia/Milupa, DBV, Hipp, Hycor, Infectopharm, Jenapharma, Mylan/Meda, Nestle, Novartis, ThermoFisher outside of the submitted work

Dr de Blay reports advisory board membership of Aimmune Therapeutics, Stallergènes Greer, Mundipharma, Novartis, Regeneron, DVB, Sanofi, ALK, Boehringer, and AstraZeneca, outside the submitted work.

Dr Hourihane reports research funding from Aimmune Therapeutics, DBV Technologies, Johnson&Johnson, Dublin Skin and Cancer Hospital Charity, National Children's Research Centre, Dublin, Ireland, Temple St Hospital Foundation Charity and Clemens von Pirquet Foundation; advisory board membership of Aimmune Therapeutics; speaker bureaux of Aimmune Therapeutics, DBV Technologies and Danone-Nutricia, outside the submitted work.

Dr Papadopoulos reports research Support from Gerolymatos Int, Capricare, Nutricia, Vian; Speaker/Chairperson fees from HAL, Menarini/Faes Farma, MSD, Biomay, Novartis, Nutricia, Sanofi, Boehringer Ingelheim, Mylan/Meda, Asit Biotech; Advisory board fees from HAL, Menarini/Faes Farma, Novartis, Nutricia, GSK, AstraZeneca, Mylan/Meda, outside the submitted work.

Prof. Worm declares the receipt of honoraria or consultation fees by the following companies: ALK-Abelló Arzneimittel GmbH, Mylan Germany GmbH, Leo Pharma GmbH, Sanofi-Aventis Deutschland GmbH, Regeneron Pharmaceuticals, DBV Technologies S.A, Stallergenes GmbH, HAL Allergie GmbH, Allergopharma GmbH & Co.KG, Bencard Allergie GmbH, Aimmune Therapeutics UK Limited, Actelion Pharmaceuticals Deutschland GmbH, Novartis AG, Biotest AG., AbbVie Deutschland GmbH & Co. KG and Lilly Deutschland GmbH, outside the submitted work.

Prof. Roberts reports funding to his institution by the European Commission and the UK Food Standards Agency to undertake research as part of iFAAM, for the reported work in this manuscript.

Dr van Ree reports research grant from EU FP7 for the iFAAM Project for the reported work in this manuscript; research grants from EU FP7, Dutch Science Foundation, and Health Holland, a research contract from Agany Inc., consulting fees from HAL Allergy BV, Citeq BV, Angany Inc., lecture fees from HAL Allergy BV, ALK and Thermofisher Scientific, support for attending meetings from HAL Allergy BV, advisory borad fees from HAL Allergy BV and Agany Inc., stock options from Agany Inc., outside the submitted work.

Dr Turner reports grants from European Commission H2020 iFAAM collaboration for the reported work in this manuscript; grants from UK Medical Research Council, NIHR/Imperial BRC and UK Food Standards Agency; advisory board fees from ILSI Europe, Aimmune Therapeutics, and AllerGenis, outside the submitted work.

Dr. Mills owns founder shares in Reacta Biotech Ltd, a company developing oral food challenge diagnostics for food allergy. Funding for research and consultancy is received from the company through the University of Manchester.

All other authors declare no conflicts of interest.

ABSTRACT

Background: The heterogeneity and lack of validation of existing severity scores for food allergic reactions limit standardization of case management and research advances.

Objective: To develop and validate a severity score for food allergic reactions.

Methods: Following a multidisciplinary experts consensus it was decided to develop a food allergy severity score (FASS) with ordinal (oFASS) and numerical (nFASS) formats. oFASS with 3 and 5 grades were generated through expert consensus, and nFASS by mathematical modeling. Evaluation was performed in the EuroPrevall outpatient clinic cohort (8232 food reactions) by logistic regression with request of emergency care and medications used as outcomes. Discrimination, classification and calibration were calculated. Bootstrapping internal validation was followed by external validation (logistic regression) in 5 cohorts (3622 food reactions). Correlation of nFASS with the severity classification done by expert allergy clinicians by Best-Worst Scaling of 32 food reactions was calculated.

Results: oFASS and nFASS map consistently, with nFASS having greater granularity. With the outcomes emergency care, adrenaline and critical medical treatment, oFASS and nFASS had a good discrimination (receiver operating characteristic area under the curve [ROC-AUC]>0.80), classification (sensitivity 0.87-0.92, specificity 0.73-0.78) and calibration. Bootstrapping over ROC-AUC showed negligible biases (1.0×10^{-6}-1.23×10^{-3}). In external validation nFASS performed best with higher ROC-AUC. nFASS was strongly correlated (R 0.89) to best-worst scoring of 334 expert clinicians.

Conclusion: FASS is a validated and reliable method to measure severity of food allergic reactions. The ordinal and numerical versions that map onto each other are suitable for use by different stakeholders in different settings.

Word count 3998

Key words

Allergic reactions, anaphylaxis, food allergy, severity, score

Abbreviations/acronyms

AUC: area under the curve

BWS: Best-Worst Scaling

BR: bronchial

BSACI: British Society for Allergy & Clinical Immunology

CMT: critical medical treatment

CV: cardiovascular

EAACI: European Academy of Allergy and Clinical Immunology

FA: food allergy

FASS: Food Allergy Severity Score

HCSC: Hospital Clinico San Carlos

iFAAM: Integrated Approaches to Food Allergen and Allergy Risk Management

IVF: intravenous fluids

Lx: laryngeal

NORA, Network for Online Registration of Anaphylaxis

NS: nervous system

NPV: negative predictive value

OAS: oral allergy symptoms

oFASS: ordinal FASS

oFASS-3: ordinal FASS with 3 grades

oFASS-5: ordinal FASS with 5 grades

OR: odds ratio

PPV: positive predictive value

Q1: first quartile

Q3: third quartile

ROC: receiver operating characteristic

SAFE: Plant food allergies: field to table strategies for reducing their incidence in Europe

Se: sensitivity

SE: standard error

SEAIC: Spanish Society of Allergy and Clinical Immunology

Sp: specificity

TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis

WAO: World Allergy Organization

95%CI: 95% confidence interval

INTRODUCTION

Food allergy (FA) has become a significant medical problem in the last decades[1,2]. Food allergic reactions can affect different organs and systems, and present with a wide range of severity from mild transient oral symptoms, to severe and even fatal anaphylaxis. In fact, FA is one of the main causes of anaphylaxis in the community[3-7]. The severity of reactions varies between individuals, and within repeated reactions in the same individual. Severe reactions are not predictable, and allergists have limitations to accurately identify the patients at greatest risk of life-threatening reactions[8].

The treatment of acute allergic reactions is driven by severity. The longer-term management is also guided by the severity of the previous reaction(s), and/or by the presumed risk of having a future severe reaction[9]. With increasing risk (or risk perception) more rigorous avoidance of their problem food is needed, and patients should carry rescue medication including an adrenaline autoinjector. The uncertainty and fear of the severity and outcome of any future reaction has a profound negative impact on patients' health-related quality of life[10]. Current FA risk management strategies are focused on clinical, regulatory and industrial collaborations to develop reference doses that will or not trigger an allergic reaction and to use them to guide precautionary allergen labeling, but the severity of the reaction is not considered in the process[11,12]. Similarly, novel interventions in FA such as allergen immunotherapy or anti-IgE therapy aim to reduce the risk of accidental reactions by increasing the amount of allergen tolerated, but the severity of the reactions is not included as a primary outcome[13-16].

Severity is therefore a key parameter in FA that needs to be measured as accurately as possible. Several scoring systems have been proposed by different groups to grade severity of anaphylaxis [17-24], and of allergic reactions induced by foods[25-33], drugs[18-20], insect venoms[21-24] or allergen immunotherapy[34-37]. All the instruments are organ based, most of them generated using expert opinion with only one using a Delphi methodology[33]. The great majority classifies severity using ordinal scales (3 to 6 grades) that are not equivalent making comparisons difficult[38-40]. Furthermore, none of the current systems has been validated (i.e., the performance has never been assessed).

The European Union-funded project Integrated Approaches to Food Allergen and Allergy Risk Management (iFAAM) in collaboration with a task force of the European Academy of

Allergy and Clinical Immunology (EAACI) critically reviewed the available systems and proposed in an EAACI Position Paper an approach to develop a system to measure severity of allergic reactions[41]. Here we report the development and validation of the Food Allergy Severity Score (FASS), aligned with the EAACI Position Paper. The report follows the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement[42].

METHODS

Development of the Food Allergy Severity Score (FASS)

Premises that FASS must meet

In line with the iFAAM-EAACI expert consensus[41] FASS had to meet the following premises:

- Have an anatomical basis: as the number of organs/systems involved increase, severity increases.
- Consider laryngeal (Lx), bronchial (BR), cardiovascular (CV) and nervous system (NS) involvement as potentially life-threatening, even if isolated, and thus more severe than that of other organs/systems.
- Use variables (symptoms/signs) that are easily/routinely recorded.
- Be applicable to all patient populations (all ages, all foods).
- Be used in different countries.
- Have ordinal (oFASS) and numerical (nFASS) formats that map consistently, to facilitate use by different stakeholders in different scenarios.

The development of FASS needs to identify the organs/systems affected and the specific symptoms and signs associated to their involvement in an allergic reaction, since these are the variables used to build the score. The identification of organs and systems based on anatomy does not pose a problem. However, there are multiple ways to describe symptoms and signs of allergic reactions and a common international terminology is lacking. To overcome this, the symptoms as described in the PRACTALL consensus[43] have been taken as reference.

Ordinal Food Allergy Severity Score (oFASS)

The oFASS was built by experts' consensus of different stakeholders (expert allergy clinicians, epidemiologists, basic scientists, representatives of food allergic patients) after several rounds of discussion at iFAAM project meetings. It has two versions of five (oFASS-5) and three categories (oFASS-3) (Table I). We use the term oFASS to refer to both oFASS-5 and oFASS-3. In oFASS-5 grade 1 includes reactions restricted to the oral

cavity. Grades 2 to 5 may include oral symptoms, but other target organs are affected. Grading is based on the organ/system involved regardless of the type or number of specific symptoms present of that organ/system. Grades 2 and 3 include skin, eye/nose, digestive and/or uterine involvement, either 1 or more than 1 of them, respectively. Lx and/or BR involvement (even isolated) classifies a reaction as Grade 4, and CV and/or NS involvement (even isolated) as grade 5. In grades 4 and 5, other target organ/systems of lower grades may be affected. oFASS-3 is a simplified version, where mild corresponds to grade 1, moderate to grades 2 and 3, and severe to grades 4 and 5 of the oFASS-5.

Numerical Food Allergy Severity Score (nFASS)

The nFASS model was constructed from scratch and had to satisfy several conditions: i) the nFASS must consistently map onto the oFASS (i.e. higher oFASS implies higher nFASS, and a given nFASS value can only correspond to one level of oFASS); ii) a higher number of symptoms within the same organ/system should increase the score.

The nFASS is computed in 3 steps as shown in the example of Table II. Each organ/system has an assigned exponent, ε_o, (-1 for oral; 0 for GI, eye/nose, skin; 2 for BR, LX; 4 for CV/NS) and each symptom has an equivalence with the PRACTALL reference and an assigned weight, λ_s (Table S1). The small individual weights of the symptoms (λ<0.1), combined with the organ/system exponents (from -1 to 4) that multiply the expression, guarantee correct mapping from the nFASS to the oFASS. Further details in the Online Supporting Information and Table S2.

Evaluation cohort

The evaluation was done in the EuroPrevall outpatient clinic dataset that comprises 2112 patients and 8232 immediate food allergic reactions (characteristics in Tables S3-S4, and the equivalence of symptoms with the PRACTALL reference in Table S5). Patients reporting reactions to foods were selected in 12 allergy clinics across Europe, and evaluated following the same protocol previously described[44]. There was no exclusion based on age, type of food or severity of reaction.

External validation cohorts

The external validation cohorts covered infants, children and adults (n= 2930) selected across Europe, with immediate allergic reactions (n=3622) to any type of food and of any severity (Table S3). These cohorts included: patients from the EuroPrevall general population survey phase 3 (147 subjects, 635 food reactions)[44,45], patients recruited in the iFAAM project (iFAAM cohort, 356 patients, 453 reactions)[46,47]; food induced anaphylaxis from the Network for Online-Registration of Anaphylaxis (NORA) (1959 subjects, 2020 anaphylaxis) [4,7]; infants and toddlers with egg and milk allergies recruited at Hospital Clinico San Carlos (Madrid, Spain) (HCSC cohort, 83 children, 129 food reactions); and apple allergic patients recruited in the EU-funded SAFE project (385 subjects, 385 apple reactions)[48]. The equivalence of symptoms with the PRACTALL reference is presented in Table S5, and further information on the cohorts in the Supporting Information.

Predictors and outcomes

The predictors are the 3 versions of FASS: oFASS-3, oFASS-5 and nFASS. In order to assess the ability of FASS (the acronym FASS comprises all three versions) to reflect reaction severity, we selected in the evaluation cohort outcomes (or severity indicators) that reflected management decisions taken by patients, carers and/or health professionals in the routine assistance of the 8232 food reactions included in the cohort. We assumed that a patient who requested emergency care assistance for a food reaction had a perception of higher severity than a patient who did not. Similarly, we assumed that a reaction that was treated with medication/s was considered more severe than a reaction not treated at all. We therefore selected from the information collected in the record forms of the validation cohort the following outcomes: request of emergency care, use of any medication, antihistamines, corticosteroids and adrenaline. Any medication includes all the drugs used to treat a reaction. An outcome named critical medical treatment (CMT) was built which included use of intravenous fluids (IVF), vasopressors, oxygen and/or mechanical ventilation. In the external validation cohorts we used the outcomes for which FASS exhibited the best discrimination, classification and calibration in the evaluation cohort. Outcomes different from the ones used in the evaluation phase could not be included in the external validation[42].

Missing data

We excluded food reactions with missing information on the outcomes in the evaluation and validation cohorts and no imputation was performed (i.e. complete case analysis). Information on missing values is presented in Tables S6-S7.

Survey of allergy healthcare professionals to appraise the performance of FASS

Separately, we undertook a global survey of allergy healthcare professionals to rate the severity of different food-induced allergic reactions, using 'Best-Worst Scaling' (BWS) which avoids user scale bias[49]. In brief, an online "MaxDiff" survey was developed in which respondents were asked to rate the severity of different pairs of allergic reaction scenarios and choose the pair that, in their opinion, reflected the maximum difference in severity. The pairs of allergic reaction scenarios were selected from a total of 32 vignettes providing a wide spectrum of reaction severity (Table S8). The survey was administered by an independent market Research company (ResearchNow, UK), with potential respondents contacted through the EAACI, Spanish Society of Allergy and Clinical Immunology (SEAIC), British Society for Allergy & Clinical Immunology (BSACI), and World Allergy Organization (WAO). Responses were voluntary, anonymous and confidential (detailed information in Stafford et al[50]).

Statistical analysis

Descriptive statistics included frequency and percent for qualitative variables, and minimum, maximum, median, first and third quartiles (Q1, Q3) for numerical variables. The Cochran-Armitage test for trend was used to analyse the frequency of outcomes across the oFASS levels.

In the evaluation of FASS logistic regression models were generated for all outcomes. Models are presented with reference value, odds ratios (ORs) with their 95% confidence intervals (95%CI), and p-value for the Wald's test. FASS performance was assessed by examining discrimination, classification and calibration. Discrimination was quantified by calculating the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and its 95%CI. Classification measures included sensitivity (Se), specificity (Sp), positive and negative predictive values (PPV, NPV). For the calibration, predicted and observed (real) probabilities of the outcomes were calculated. The Hosmer-Lemeshow

test for agreement was not applied because it is sensitive to grouping (thus, non-applicable to oFASS) and sample size (large N in evaluation cohort)[42].

Internal validation was performed by bootstrap with 100 replicates applied over ROC-AUC. Bias and standard errors (SE) were estimated. External validation was done analyzing the predictive performance of FASS models in the external validation cohorts by logistic regression. ROC-AUC and 95%CI were calculated.

The responses to the MaxDiff survey were modelled and a "preference" score (representing severity) determined for each scenario[50]. These scores were compared to nFASS using Spearman's R correlation.

All statistical analyses were performed using R 4.0.3, and Python 3.8.5. Significant level was set at p= 0.05.

Ethical considerations

The Institutional Review Board of HCSC (Madrid, Spain) confirmed that ethical approval was not required to the development and validation of FASS. The online survey of allergy healthcare professionals did not require ethical approval, but was approved by EAACI, SEAIC, BSACI and WAO. ResearchNow follows the UK Market Research Society's Code of Conduct.

RESULTS

Evaluation of FASS

The FASS scores were implemented in the 8232 food reactions of the EuroPrevall outpatient clinic dataset (Tables S3-S4). The frequency of the different symptoms is presented in Figure S1, and the severity distribution in Table III. According to oFASS-3, 35.8% of reactions were mild, 34.5% moderate and 29.7% severe. The majority (78.8%) of moderate reactions were of Grade 2 (one organ/system affected) accounting for 27.2% of total reactions. The most frequent severe reactions (81.1%) were of Grade 4 (Lx/BR) and represented 24.1% of total reactions. The nFASS values ranged from 1.07 to 7.75. All mild-grade 1 reactions had an nFASS value of 1.07 because they were all collected under one category (oropharyngeal symptoms) in the EuroPrevall record forms. The nFASS values of moderate and severe reactions ranged from 2.01 to 3.98, and from 4.07 to 7.75, respectively. The correct mapping of nFASS onto oFASS (Table III) allows to understand the meaning of a given numerical score.

Logistic regression models were computed for oFASS-3, oFASS-5 and nFASS with any medication, adrenaline, corticosteroids, antihistamine, CMT and request of emergency care as outcomes. The frequency of the severity indicators increased progressively (Cochran-Armitage test for trend, $p<0.01$) as the severity of reactions increased with oFASS-3 and oFASS-5 (Table S9). A positive association with all the severity indicators was found for the 3 versions of FASS, implying that a higher grade of oFASS-3 and oFASS-5, and an increment in the nFASS value (OR given per 1 point increment) was associated with higher probability of use of medications and request of emergency care to treat the reaction (Table IV). As presented in Table IV and Figure 1-A, all models had ROC-AUC>0.70 which is a requirement in model development[51]. There is a trend for a progressive increase of the ROC-AUC from oFASS-3 to oFASS-5 and nFASS. Furthermore the ROC-AUC of the models with adrenaline (0.83 to 0.90), CMT (0.83 to 0.91) and emergency care (0.78 to 0.84) tended to be higher than those of either any medication (0.73 to 0.76), corticosteroids (0.76 to 0.80) or antihistamine (0.71 to 0.73).

Classification measures of the models were calculated (Table S10). With the outcomes adrenaline and CMT, Se ranged from 0.87 to 0.92 and Sp from 0.73 to 0.78. With emergency care Se was around 0.70 and Sp 0.75. The PPVs ranged from 0.10 to 0.63

depending on the prevalence of the severity indicators in the cohort (with higher outcome prevalence, higher PPV).

Predicted and observed (real) probabilities of the severity indicators show that oFASS-3, oFASS-5 and nFASS are well calibrated (Table S11).

Internal validation

As a mechanism to account for overfitting of the FASS models, a bootstrap internal validation over the ROC-AUC was undertaken. The bias and SE results for all the outcomes (Table S12) show negligible overestimation biases (1.0×10^{-6}-1.23×10^{-3}).

External validation

The FASS scores were implemented in the food reactions of the external validation cohorts and the severity distributions are presented in Table III. In the NORA cohort close to 90% were grade 4 and 5 reactions in oFASS-5, and 7.02% had systemic reactions with involvement of 2 or more organs (grade 3), which is consistent with the inclusion criteria in NORA[4,7]. In contrast, in the SAFE cohort on apple allergic patients, close to 60% of patients had exclusively mild-grade 1 oropharyngeal symptoms, and severe reactions only appeared in 9% of subjects. It can be seen that nFASS mapped correctly onto oFASS for the 5 external validation cohorts.

In the external validation we used the best 3 severity indicators of the evaluation in terms of discrimination, classification and calibration: adrenaline, CMT and emergency care. The frequency of these outcomes increased significantly in all the cohorts as severity increased classified with oFASS (Table S13). The predictive performance of FASS was evaluated using logistic regression. The ROC-AUC with their 95%CI are presented in Table V and those of nFASS in Figure 1-B. Overall, there was a trend for lower ROC-AUC for oFASS-3 and higher for nFASS in all the cohorts. The ROC-AUC for the 3 FASS models were lower in NORA compared with the other cohorts.

Comparison of nFASS to the BWS ranking of allergy healthcare professionals

The responses of 334 allergy healthcare professionals were modelled following BWS methodology and a BWS score was assigned to each of the 32 food reaction scenarios. The reaction scenarios include information on the symptoms/signs and on the treatment

given (or not), and the final outcome of the reaction. The severity of these reactions was also scored with FASS (Table S8). A strong correlation (Spearman R 0.89) was found between BWS and nFASS values (Figure S2), even if FASS severity scoring was derived solely on the basis of symptoms.

DISCUSSION

We have developed and validated a severity score of food allergic reactions with ordinal (oFASS-3 and oFASS-5) and numerical (nFASS) formats that map consistently. They have different granularity, and are intended to be used by different stakeholders in different settings. After a comprehensive validation following the TRIPOD statement[42] (Table S14), FASS has proven to be a reliable and well calibrated method to describe severity of food allergic reactions. This was shown in the EuroPrevall outpatient clinic cohort used in the evaluation, and was confirmed by bootstrap internal validation, and by external validation in 5 different cohorts that cover the whole spectrum of FA. Additionally, nFASS is strongly correlated to the classification of severity of food reaction scenarios done by expert allergy clinicians, further supporting the good performance of FASS.

For the validation we selected as severity indicators (outcomes) the request of emergency care or the medications given to control the reactions, since they reflected management decisions taken by the patients themselves or by the clinicians providing medical care. The severity indicators were those collected in the forms of the evaluation cohort, with the CMT outcome built to identify the most severe ones. The variability in the management of allergic reactions and anaphylaxis and the underuse of adrenaline in relation to the recommendations[9,52] might lead to a limitation of the selected outcomes. We have shown that the frequency of the outcomes significantly rose with increasing severity of reactions graded with oFASS in the evaluation cohort (Table S9). The discrimination, classification and calibration of FASS was better for the outcomes of use of adrenaline, CMT and request of emergency care selected for the external validation (Figure 1-A, Tables IV, S10, S11), suggesting that they are more appropriate severity indicators than use of any medication, antihistamine and corticosteroids. Also in the external validation cohorts the 3 selected outcomes significantly increased as severity increased by oFASS (Table S13). We thus believe our results support the adequacy of adrenaline, CMT and request of emergency care as severity indicators. Additionally, several studies[52-54] reporting that adrenaline administration by health professionals increases with increasing severity of the allergic reactions support the selection of adrenaline as an outcome. This has been shown in the pre-hospital management of

anaphylaxis by ambulance crew in Manchester, UK[53], in two pediatric emergency departments of Marseille, France[54], and in the management of severe allergic reactions done by health professionals of 10 European countries of the NORA network[52]. With the external validation cohorts we found an overall good predictive performance, especially of nFASS, with some limitations in the NORA cohort. Due to the fact that almost 90% of the reactions included in NORA were severe/grade 4-5 (Table III), the discrimination capacity of oFASS is limited (ROC-AUC<0.7) (Table V). The best performance is found with nFASS for CMT (ROC-AUC 0.7) due to its higher granularity, and because CMT may be a more adequate outcome than adrenaline or request of emergency care for these anaphylaxis patients. Further work is needed to assess the performance of FASS with other outcomes for anaphylaxis (i.e. number of adrenaline doses, intravenous adrenaline, intensive care admission), but since this information was not collected in the evaluation cohort, we could not analyze them in the external validation.

Like other scoring systems of allergic reactions FASS has an anatomical basis [17-29,31-37], but includes some novel approaches in scoring severity, present already in the (more conventional) oFASS.

Isolated oral symptoms (OAS) are included separately in oFASS-3 (mild) and oFASS-5 (grade 1), because they are a local reaction, and frequently the only clinical manifestation of FA as can be seen in Tables II and IV. Actually, 30% of all the food reactions (3545 out of 11854) of the evaluation and validation cohorts would not have been properly classified without including such category. In the currently available grading systems applied in FA, OAS is either absent[25,26,29,32] or included with symptoms of other organs under the same grade[27,28,31,33].

In oFASS when the number of organ/systems involved increase, the severity increases, but reactions will only be considered severe (Table I) if Lx, BR, CV and/or NS are involved, even if isolated. When reactions involve the eye/nose, skin, GI tract, uterus, they are considered moderate in oFASS-3. The oFASS-5 provides more detail by dividing moderate reactions into grades 2 (1 target organ) and 3 (>1 organ), and severe reactions into grades 4 (Lx/BR) and 5 (CV, NS). So with oFASS-5 it can be seen in the evaluation and validation cohorts (Table III) that the most frequent reactions among the moderate ones are those with only one organ/system affected (grade 2), and among the severe

ones those with respiratory involvement (grade 4). The severe grade in oFASS-3 and grades 4 and 5 in oFASS-5 capture all the potentially life-threatening anaphylaxis[55, 56].

The anatomical approach to classify severity of oFASS-3 is a simple, easy-to-remember system to identify the potentially life-threatening anaphylaxis, and prompt the early use of adrenaline. It can thus be used to educate patients and their carers, healthcare professionals, and even non-healthcare professionals who may face food allergic reactions at work (e.g. restaurants, schools). At a population level it may be useful for education and raising awareness of food anaphylaxis, and thus of interest for public health authorities. oFASS-5 may be more informative for allergy healthcare professionals who need to document reactions with more detail, but even so, it is still a simple system easy-to-remember and use. In summary, oFASS-3 may be of highest value for patients and non-healthcare professionals, and both oFASS-3 and oFASS-5 for all healthcare professionals in their clinical practice. Furthermore, oFASS-3 and oFASS-5 have shown to be reliable in measuring severity, with a trend for a better performance of oFASS-5 (Tables IV-V) related to its higher level of detail.

In contrast to other ordinal systems we have not considered in oFASS the intensity of the symptoms/signs of the organ/system affected, or whether they are subjective or objective. This was done for the sake of simplicity and the intended use aforementioned. When more detail is required nFASS can be applied. As an example, a reaction with “skin involvement” will be classified as moderate with oFASS-3 and grade 2 with oFASS-5, regardless of whether the skin involvement is “mild pruritus” or “intense pruritus with generalized erythema, urticaria and angioedema”. However, the nFASS will score 2.36 for the later and 2.01 for mild pruritus, and it is thus able to differentiate between these two reactions.

nFASS seems suitable for scoring reaction severity in clinical research settings, and could help in the risk management of FA by including beside reference doses the severity parameter. We are currently exploring its usefulness in observational and intervention studies, analyzing the effect of cofactors and food allergen immunotherapy on the severity of food allergic reactions.

Further validation of FASS in other retrospective and, more importantly, prospective cohorts in both clinical and research settings, is needed to refine this tool and confirm its

reliability. FASS is based on the symptoms and organs affected which are not specific of food allergic reactions. Similar symptomatology and organ involvement can be observed in allergic reactions elicited by insect stings and drugs, and we are currently working in the adaptation and validation of FASS to allergy triggers other than foods.

In order to facilitate the use of FASS we have developed a software tool written in R language that allows to implement oFASS-3, oFASS-5 and nFASS in any dataset of food allergic reactions. The FASS tool is able to read any file in table format from Excel, SPSS and STATA. The software tool, a tutorial and an example are available at Zenodo http://doi.org/10.5281/zenodo.4836276.

Through the FASS dissemination, widespread use and further refinement, we can achieve the goal of improving education and immediate care decisions of patients and clinicians, help advance research, and guide the food industry and the health authorities.

REFERENCES

1) Nwaru BI, Hickstein L, Panesar SS, et al. Prevalence of common food allergies in Europe: a systematic review and meta-analysis. Allergy 2014;69:992-1007. DOI: 10.1111/all.12423

2) Dunlop JH, Keet CA. Epidemiology of Food Allergy. Immunol Allergy Clin North Am 2018;38:13-25. DOI: 10.1016/j.iac.2017.09.002

3) Panesar SS, Javad S, de Silva D, et al. The epidemiology of anaphylaxis in Europe: a systematic review. Allergy 2013;68:1353-61. DOI: 10.1111/all.12272

4) Worm M, Moneret-Vautrin A, Scherer K, et al. First European data from the network of severe allergic reactions (NORA). Allergy 2014; 69: 1397–1404. DOI:10.1111/all.12475

5) Turner PJ, Gowland MH, Sharma V, et al. Increase in anaphylaxis-related hospitalizations but no increase in fatalities: an analysis of United Kingdom national anaphylaxis data, 1992-2012. J Allergy Clin Immunol 2015;135:956-63. DOI: 10.1016/j.jaci.2014.10.021

6) Tejedor-Alonso MA, Moro-Moro M, Mosquera Gonzalez M, et al. Increased incidence of admissions for anaphylaxis in Spain 1998–2011. Allergy 2015; 70: 880–3. DOI:10.1111/all.12613

7) Grabenhenrich LB, Dölle S, Moneret-Vautrin A, et al. Anaphylaxis in children and adolescents: The European Anaphylaxis Registry. J Allergy Clin Immunol 2016;137:1128-37. DOI: 10.1016/j.jaci.2015.11.015

8) Turner PJ, Baumert JL, Beyer K, et al. Can we identify patients at risk of life-threatening allergic reactions to food? Allergy 2016;71:1241-55. DOI: 10.1111/all.12924

9) Muraro A, Werfel T, Hoffmann-Sommergruber K, et al. Diagnosis and management of food allergy. Allergy 2014; 69: 1008–25. DOI. 10.1111/all.12429

10) DunnGalvin A, Blumchen K, Timmermans F, et al. APPEAL-1: A multiple-country European survey assessing the psychosocial impact of peanut allergy. Allergy. 2020;75:2899-2908. DOI: 10.1159/000375106

11) Crevel RW, Baumert JL, Baka A, Houben GF, Knulst AC, Kruizinga AG, et al. Development and evolution of risk assessment for food allergens. Food Chem Toxicol 2014;67:262-76. DOI: 10.1016/j.fct.2014.01.032

12) Dubois AEJ, Turner PJ, Hourihane J, et al. How does dose impact on the severity of food-induced allergic reactions, and can this improve risk assessment for allergenic foods?: Report from an ILSI Europe Food Allergy Task Force Expert Group and Workshop. Allergy 2018;73:1383-92. DOI: 10.1111/all.13405

13) Pajno GB, Fernandez-Rivas M, Arasi S, et al. EAACI Guidelines on allergen immunotherapy: IgE-mediated food allergy. Allergy. 2018;73:799-815. DOI:10.1111/all.13319

14) PALISADE Group of Clinical Investigators, Vickery BP, Vereda A, Casale TB, et al. AR101 Oral Immunotherapy for Peanut Allergy. N Engl J Med 2018 22;379:1991-2001. DOI: 10.1056/NEJMoa1812856

15) Fleischer DM, Greenhawt M, Sussman G, et al. Effect of Epicutaneous Immunotherapy vs Placebo on Reaction to Peanut Protein Ingestion Among Children With Peanut Allergy: The PEPITES Randomized Clinical Trial. JAMA 2019;321:946-55. DOI: 10.1001/jama.2019.1113

16) Nicolaides RE, Parrish CP, Bird JA. Food Allergy Immunotherapy with Adjuvants. Immunol Allergy Clin North Am 2020;40:149-173. Doi: 10.1016/j.iac.2019.09.004

17) Brown SG. Clinical features and severity grading of anaphylaxis. J Allergy Clin Immunol 2004;114:371-6. DOI: 10.1016/j.jaci.2004.04.029

18) Ring J, Messmer K. Incidence and severity of anaphylactoid reactions to colloid volume substitutes. Lancet 1977;1:466-9. DOI: 10.1016/s0140-6736(77)91953-5

19) Ring J. Anaphylactoid reactions to intravenous solutions used for volume substitution. Clin Rev Allergy 1991;9:397-414. DOI: 10.1007/BF02802316

20) Ring J, Behrendt H. Anaphylaxis and anaphylactoid reactions. Classification and pathophysiology. Clin Rev Allergy Immunol 1999;17:387-99. DOI: 10.1007/BF02737644

21) Mueller HL. Further experiences with severe allergic reactions to insect stings. N Engl J Med 1959;261:374-7. DOI: 10.1056/NEJM195908202610803

22) Lockey RF, Turkeltaub PC, Baird-Warren IA, et al. The Hymenoptera venom study I, 1979-1982: demographics and history-sting data. J Allergy Clin Immunol 1988;82:370-81. DOI: 10.1016/s0091-6749(05)80182-4
23) Reisman RE. Natural history of insect sting allergy: relationship of severity of symptoms of initial sting anaphylaxis to re-sting reactions. J Allergy Clin Immunol 1992;90:335-9. DOI: 10.1016/s0091-6749(05)80012-0
24) Golden DB, Kwiterovich KA, Kagey-Sobotka A, Lichtenstein LM. Discontinuing venom immunotherapy: extended observations. J Allergy Clin Immunol 1998;101:298-305. DOI: 10.1016/S0091-6749(98)70239-8
25) Hourihane JO, Kilburn SA, Dean P, Warner JO. Clinical characteristics of peanut allergy. Clin Exp Allergy 1997;27:634-9. DOI: 10.1111/j.1365-2222.1997.tb01190.x
26) Hourihane JO, Grimshaw KE, Lewis SA, et al. Does severity of low-dose, double-blind, placebo-controlled food challenges reflect severity of allergic reactions to peanut in the community? ClinExp Allergy 2005;35:1227-33. DOI: 10.1111/j.1365-2222.2005.02312.x.
27) Ewan PW, Clark AT. Long-term prospective observational study of patients with peanut and nut allergy after participation in a management plan. Lancet 2001;357:111-5. DOI: 10.1016/s0140-6736(00)03543-1
28) Sampson HA. Anaphylaxis and emergency treatment. Pediatrics 2003;111:1601-8. PMID: 12777599
29) Astier C, Morisset M, Roitel O, et al. Predictive value of skin prick tests using recombinant allergens for diagnosis of peanut allergy. J Allergy Clin Immunol 2006;118:250-6. DOI: 10.1016/j.jaci.2006.04.053
30) Cianferoni A, Garrett JP, Naimi DR, Khullar K, Spergel JM. Predictive values for food challenge-induced severe reactions: development of a simple food challenge score. Isr Med Assoc J 2012;14:24-8. PMID: 22624438
31) Hino A, Maeda T, Haneda Y, et al. [Establishment of "Anaphylaxis Scoring Aichi (ASCA)," a new symptom scoring system to be used in an oral food challenge (OFC)]. Arerugi 2013;62:968-79. PMID: 24335424
32) Niggemann B, Beyer K. Time for a new grading system for allergic reactions? Allergy 2016;71:135-6. DOI: 10.1111/all.12765

33) Dribin TE, Schnadower D, Spergel JM, et al. Severity grading system for acute allergic reactions: A multidisciplinary Delphi study. J Allergy Clin Immunol. 2021 Jan 19:S0091-6749(21)00047-6. DOI: 10.1016/j.jaci.2021.01.003.
34) Bernstein DI, Epstein T, Murphy-Berendts K, Liss GM. Surveillance of systemic reactions to subcutaneous immunotherapy injections: year 1 outcomes of the ACAAI and AAAAI collaborative study. Ann Allergy Asthma Immunol 2010;104:530-5. DOI: 10.1016/j.anai.2010.04.008
35) Cox L, Larenas-Linnemann D, Lockey RF, Passalacqua G. Speaking the same language: The World Allergy Organization Subcutaneous Immunotherapy Systemic Reaction Grading System. J Allergy Clin Immunol 2010;125:569-74,574.e1-574.e7. DOI: 10.1016/j.jaci.2009.10.060
36) Cox LS, Sanchez-Borges M, Lockey RF. World Allergy Organization Systemic Allergic Reaction Grading System: Is a Modification Needed? J Allergy Clin Immunol Pract 2017;5:58-62.e5. DOI: 10.1016/j.jaip.2016.11.009
37) Epstein TG, Liss GM, Murphy-Berendts K, Bernstein DI. AAAAI/ACAAI surveillance study of subcutaneous immunotherapy, years 2008-2012: an update on fatal and nonfatal systemic allergic reactions. J Allergy Clin Immunol Pract 2014;2:161-7. DOI: 10.1016/j.jaip.2014.01.004
38) Eller E, Muraro A, Dahl R, Mortz CG, Bindslev-Jensen C. Assessing severity of anaphylaxis: a data-driven comparison of 23 instruments. Clin Transl Allergy 2018;8:29. DOI: 10.1186/s13601-018-0215-x
39) Vidal C, Rodríguez Del Río P, Gude F, et al. Comparison of International Systemic Adverse Reactions Due to Allergen Immunotherapy. J Allergy Clin Immunol Pract 2019;7:1298-1305.e3. DOI: 10.1016/j.jaip.2018.12.006
40) Arasi S, Nurmatov U, Dunn-Galvin A, et al. Consensus on DEfinition of Food Allergy SEverity (DEFASE) an integrated mixed methods systematic review. World Allergy Organ J. 2021 Mar 11;14(3):100503. DOI: 10.1016/j.waojou.2020.100503
41) Muraro A, Fernandez-Rivas M, Beyer K, et al. The urgent need for a harmonized severity scoring system for acute allergic reactions. Allergy. 2018;73:1792-1800. DOI: 10.1111/all.13408

42) Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73. DOI: 10.7326/M14-0698

43) Sampson HA, Gerth van Wijk R, Bindslev-Jensen C, et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology-European Academy of Allergy and Clinical Immunology PRACTALL consensus report. J Allergy Clin Immunol 2012;130:1260-74. DOI: 10.1016/j.jaci.2012.10.017

44) Fernández-Rivas M, Barreales L, Mackie AR, et al. The EuroPrevall outpatient clinic study on food allergy: background and methodology. Allergy 2015;70:576-84. DOI: 10.1111/all.12585

45) Kummeling I, Mills EN, Clausen M, et al. The EuroPrevall surveys on the prevalence of food allergies in children and adults: background and study methodology. Allergy 2009;64:1493-7. DOI: 10.1111/j.1398-9995.2009.02046.x

46) Roberts G, Allen K, Ballmer-Weber B, et al. Identifying and managing patients at risk of severe allergic reactions to food: Report from two iFAAM workshops. Clin Exp Allergy 2019;49:1558-1566. DOI: 10.1111/cea.13516

47) Grabenhenrich LB, Reich A, Bellach J, et al. A new framework for the documentation and interpretation of oral food challenges in population-based and clinical research. Allergy. 2017;72:453-461. DOI: 10.1111/all.13049.

48) Fernández-Rivas M, Bolhaar S, González-Mancebo E, et al. Apple allergy across Europe: how allergen sensitization profiles determine the clinical expression of allergies to plant foods. J Allergy Clin Immunol 2006;118:481-8. DOI: 10.1016/j.jaci.2006.05.012

49) Flynn TN, Marley AAJ. Best-worst scaling: theory and methods. In: Hess S, Daly A, eds. Handbook of Choice Modelling, Cheltenham, UK: Edward Elgar Publishing; 2014. p. 178-201.

50) Stafford A, Bartra J, Aston A, et al. *Improving* severity scoring of food-induced allergic reactions: a global “Best-Worst Scaling” exercise. J Allergy Clin Immunol Pract 2021 *(in press).* DOI*:* 10.1016/j.jaip.2021.06.056

51) Lemeshow S, Le Gall JR. Modelling the severity of illness of ICU patients.A systems update. JAMA 1994; 272: 1049–55. PMID: 8089888

52) Grabenhenrich LB, Dölle S, Ruëff F, et al. Epinephrine in severe allergic reactions: The European Anaphylaxis Register. J Allergy Clin Immunol Pract 2018; 6: 1898-906. DOI:10.1016/j.jaip.2018.02.026

53) Capps JA, Sharma V, Arkwright PD. Prevalence, outcome and pre-hospital management of anaphylaxis by first aiders and paramedical ambulance staff in Manchester, UK. Resuscitation 2010; 81: 653-7. DOI:10.1016/j.resuscitation.2010.01.021

54) Dubus JC, Lê MS, Vitte J, et al. Use of epinephrine in emergency department depends on anaphylaxis severity in children. Eur J Pediatr. 2019;178:69-75. DOI: 10.1007/s00431-018-3246-3

55) Sampson HA, Muñoz-Furlong A, Campbell RL, et al. Second symposium on the definition and management of anaphylaxis: summary report – Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. J Allergy Clin Immunol 2006; 117:391–397. DOI: 10.1016/j.jaci.2005.12.1303

56) Turner PJ, Worm M, Ansotegui IJ, et al; WAO Anaphylaxis Committee. Time to revisit the definition and clinical criteria for anaphylaxis? World Allergy Organ J. 2019 Oct 31;12(10):100066. DOI: 10.1016/j.waojou.2019.100066

Table I. Ordinal Food Allergy Severity Score (oFASS) versions: oFASS-3 and oFASS-5

oFASS-3	oFASS-5	Organs/Systems involved			
		Oral cavity	**Skin Nose/Eye Digestive Uterus**	**Larynx Bronchi**	**Cardiovascular Nervous system**
Mild	**Grade 1**	Yes	No	No	No
Moderate	**Grade 2**	Yes/No	1	No	No
	Grade 3	Yes/No	>1	No	No
Severe	**Grade 4**	Yes/No	Yes/No	1 or both	No
	Grade 5	Yes/No	Yes/No	Yes/No	1 or both

yes: involved; no: not involved; yes/no: it can be involved or not

Food reaction: boy 5 years, 15 minutes after eating peanut presents urticaria, red eyes, nausea, wheeze, and dizziness					
Step 1					**Step 2**
Identify organ/system involved and the symptom PRACTALL equivalence to select the corresponding organ exponent and symptom weight					**Compute one organ/system contribution ($nFASS_o$)**
Symptom in the reaction	Organ System	Exponent ε	Symptom equivalence PRACTALL	Weight λ	$nFASS_o = 2^{\varepsilon_0}(1 + \lambda_1 + \ldots + \lambda_n)$
Urticaria	Skin	$\varepsilon_0 = 0$	Generalized involvement (>10)	$\lambda_1 = 0.08$	$nFASS_{Skin} = 2^{\varepsilon_0}(1 + \lambda_1) = 2^0(1 + 0.08) = 1.08$
Red eyes	Eye	$\varepsilon_1 = 0$	Intermittent rubbing of eyes	$\lambda_1 = 0.05$	$nFASS_{\mathrm{Eye}} = 2^{\varepsilon_1}(1 + \lambda_1) = 2^0(1 + 0.05) = 1.05$
Nausea	GI	$\varepsilon_2 = 0$	Complaints of nausea OR abdominal pain	$\lambda_1 = 0.03$	$nFASS_{\mathrm{GI}} = 2^{\varepsilon_2}(1 + \lambda_1) = 2^0(1 + 0.03) = 1.03$
Wheeze	Lower respiratory	$\varepsilon_3 = 2$	Wheezing: Inspiratory and expiratory wheezing to auscultation	$\lambda_1 = 0.07$	$nFASS_{\mathrm{LoResp}} = 2^{\varepsilon_3}(1 + \lambda_1) = 2^2(1 + 0.07) = 4.28$
Dizzy	NS	$\varepsilon_4 = 4$	Weak, dizzy	$\lambda_1 = 0.05$	$nFASS_{\mathrm{NS}} = 2^{\varepsilon_4}(1 + \lambda_1) = 2^4(1 + 0.05) = 16.8$
Step 3					
Sum values of all the organs/systems affected and apply transformation to reduce the scale					
$nFASS = \log_2 \left(\sum_{o \epsilon Organs} nFASS_o \right) + 2$					

$$\boldsymbol{nFASS} = \log_2 (nFASS_{Skin} + nFASS_{\mathrm{Eye}} + nFASS_{\mathrm{GI}} + nFASS_{\mathrm{LoResp}} + nFASS_{\mathrm{NS}}) + 2 = log_2(1.08 + 1.05 + 1.03 + 4.28 + 16.8) + 2 = 4.59 + 2 = \mathbf{6.59}$$

Table II. Computing nFASS

GI, gastrointestinal; NS, nervous system; LoResp, lower respiratory

Table III. Severity of food reactions in the evaluation and external validation cohorts

Cohort	*oFASS-3*		*oFASS-5*		*nFASS*		
	Level	**N (%)**	**Level**	**N (%)**	**Min**	**Max**	**Median (Q1,Q3)**
Evaluation cohort							
EuroPrevall Outpatient clinic cohort N = 8232	*Mild*	2946 (35.8%)	*Grade 1*	2946 (35.8%)	1.07	1.07	1.07 (1.07,1.07)
	Moderate	2839 (34.5%)	*Grade 2*	2236 (27.2%)	2.01	2.80	2.23 (2.11, 2.66)
			Grade 3	603 (7.3%)	3.03	3.98	3.38 (3.12, 3.44)
	Severe	2447 (29.7%)	*Grade 4*	1984 (24.1%)	4.07	5.68	4.58 (4.39, 5.01)
			Grade 5	463 (5.6%)	6.07	7.75	6.78 (6.46, 7.26)
External validation cohorts							
NORA N = 2020	*Mild*	0	*Grade 1*	0	-	-	-
	Moderate	210 (10.4%)	*Grade 2*	68 (3.36%)	2.01	2.27	2.12 (2.09,2.2)
			Grade 3	142 (7.02%)	3.04	3.74	3.12 (3.09,3.16)
	Severe	1810 (89.6%)	*Grade 4*	905 (44.8%)	4.08	5.68	4.69 (4.44,5.27)
			Grade 5	905 (44.8%)	6.11	7.64	6.57 (6.47,7.15)
HCSC Infant	*Mild*	15 (11.6%)	*Grade 1*	15 (11.6%)	1.07	1.07	1.07 (1.07,1.07)
	Moderate	101 (78.3%)	*Grade 2*	66 (51.2%)	2.07	2.74	2.65 (2.11,2.68)

Cohort N = 129			*Grade 3*	35 (27.1%)	3.07	3.44	3.12 (3.09,3.4)
	Severe	13 (10.1%)	*Grade 4*	10 (7.75%)	4.09	4.89	4.48 (4.42,4.64)
			Grade 5	3 (2.32%)	6.2	7.35	6.78 (6.49,7.06)
EuroPrevall general population N = 635	Mild	291 (45.8%)	*Grade 1*	291 (45.8%)	1.07	1.07	1.07 (1.07,1.07)
	Moderate	196 (30.9%)	*Grade 2*	170 (26.8%)	2.01	2.74	2.65 (2.18, 2.67)
			Grade 3	26 (4.1%)	3.03	3.43	3.37 (3.37,3.41)
	Severe	148 (23.3%)	*Grade 4*	138 (21.7%)	4.07	5.48	4.53 (4.24,5.17)
			Grade 5	10 (1.57%)	6.11	6.8	6.75 (6.3,6.76)
SAFE N = 385	Mild	230 (59.7%)	*Grade 1*	230(59.7%)	1.07	1.07	1.07 (1.07,1.07)
	Moderate	120 (31.2%)	*Grade 2*	103 (26.8%)	2.07	2.68	2.65 (2.65,2.65)
			Grade 3	17 (4.41%)	3.09	3.4	3.39 (3.39,3.39)
	Severe	35 (9.09%)	*Grade 4*	34 (8.83%)	4.26	4.79	4.54 (4.26,4.54)
			Grade 5	1 (0.2%)	6.46	6.46	6.46 (6.46, 6.46)
iFAAM N = 453	*Mild*	63 (13.9%)	*Grade 1*	63 (13.9%)	1.07	1.07	1.07 (1.07,1.07
	Moderate	210 (46.4%)	*Grade 2*	159 (35.1%)	2.01	2.75	2.13 (2.11,2.59)
			Grade 3	51 (11.3%)	3.04	3.96	3.13 (3.09,3.38)
	Severe	180 (39.7%)	*Grade 4*	111 (24.5%)	4.07	5.63	4.52 (4.23, 4.69)
			Grade 5	69 (15.2%)	6.07	7.55	6.56 (6.39,6.81)

Min, minimum; Max, maximum; Q1, first quartile; Q3, third quartile.

Table IV. Model evaluation in the EuroPrevall outpatient clinic cohort: discrimination of logistic regression models with oFASS-3, oFASS-5 and nFASS.

Severity Indicator (outcome)	oFASS-3			oFASS-5			nFASS*	
	Level	OR (95%CI)	ROC-AUC (95%CI)	Level	OR (95%CI)	ROC-AUC (95%CI)	OR (95%CI)	ROC-AUC (95%CI)
Any medication N = 3309 (40.5%)#	*Moderate :Mild*	6.7 (5.9, 7.6)	0.73 (0.72,0.75)	*Grade 2 : Grade 1*	5.8 (5.1, 6.6)	0.75 (0.74,0.76)	1.7 (1.7, 1.8)	0.76 (0.75, 0.77)
				Grade 3 : Grade 1	11.3 (9.3, 13.8)			
	Severe :Mild	11.5 (10, 13.2)		*Grade 4 : Grade 1*	9.6 (8.4, 11.1)			
				Grade 5 : Grade 1	29.6 (22.8, 38.4)			
Adrenaline N = 278 (3.37%)#	*Moderate: Mild*	2.4 (1, 5.8)	0.83 (0.80,0.85)	*Grade 2 : Grade 1*	2.3 (0.9, 5.8)	0.88 (0.85,0.90)	2.8 (2.6, 3.1)	0.9 (0.88, 0.93)
				Grade 3 : Grade 1	2.8 (0.8, 9.7)			
	Severe :Mild	49.2 (23.2, 104.5)		*Grade 4 : Grade 1*	25.7 (12, 55.3)			
				Grade 5 : Grade 1	188.9 (87.7, 407.1)			
Corticosteroids N = 993 (12.2%)#	*Moderate :Mild*	10.6 (7.4, 15.1)	0.76 (0.74,0.78)	*Grade 2 : Grade 1*	9.4 (6.5, 13.5)	0.79 (0.78,0.80)	1.9 (1.8, 2)	0.8 (0.78, 0.81)
				Grade 3 : Grade 1	15.4 (10.2, 23)			
	Severe :Mild	31.2 (22, 44.3)		*Grade 4 : Grade 1*	21.8 (15.3, 31.1)			
				Grade 5 : Grade 1	103.6 (70.4, 152.3)			
Antihistamine N = 3012	*Moderate :Mild*	6.3 (5.5, 7.2)	0.71 (0.70,0.72)	*Grade 2 : Grade 1*	5.4 (4.7, 6.3)	0.73 (0.72,0.74)	1.6 (1.6, 1.7)	0.73 (0.72, 0.75)
				Grade 3 : Grade 1	10.5 (8.6, 12.8)			

(36.9%)#	*Severe :Mild*	9.0 (7.9, 10.3)		*Grade 4 : Grade 1*	7.5 (6.5, 8.7)			
				Grade 5 : Grade 1	21.7 (17.1, 27.6)			
CMT N = 331 (4.05%)#	*Moderate :Mild*	39.2 (5.4, 285.9)	0.83 (0.81,0.85)	*Grade 2 : Grade 1*	26.8 (3.6, 199.6)	0.89 (0.87,0.91)	2.9 (2.6, 3.1)	0.91 (0.9, 0.93)
				Grade 3 : Grade 1	86.4 (11.5, 650.2)			
	Severe: Mild	403.9 (56.7,2878.3)		*Grade 4 : Grade 1*	190.5 (26.6,1364.2)			
				Grade 5 : Grade 1	1811 (252.8, 12979)			
Emergency Care (request of) N= 928 (11.3%)#	*Moderate :Mild*	21.9 (12.7, 37.5)	0.78 (0.76,0.80)	*Grade 2 : Grade 1*	17.6 (10.2, 30.4)	0.81 (0.80,0.83)	2.8 (2.6, 3.1)	0.84 (0.82, 0.85)
				Grade 3 : Grade 1	39.2 (22.2, 69.3)			
	Severe :Mild	75.4 (44.3, 128.4)		*Grade 4 : Grade 1*	49.3 (28.8, 84.3)			
				Grade 5 : Grade 1	292 (167.3, 509.6)			

N, number of food reactions of the evaluation cohort in which the severity indicator is present; #, percent of food reactions excluding those in which the severity indicator is missing (missing values shown in Table S6); CMT, critical medical treatment; OR, odds ratio; ROC-AUC, receiver operating characteristic area under the curve.

Mild in oFASS-3 and Grade 1 in oFASS-5 were used as reference categories.

*OR in nFASS corresponds to the relative increment of probability per 1 point increment in the nFASS.

All ORs are significant (Wald's test $p < 0.05$), except for adrenaline use and grades 2 and 3 of oFASS-5

Table V. External validation: logistic regression models with oFASS-3, oFASS-5 and nFASS.

Severity indicator (outcome)		**Adrenaline** ROC-AUC (95%CI)	**CMT** ROC-AUC (95%CI)	**Emergency care** ROC-AUC (95%CI)
NORA cohort N = 2020	*N(%)#*	400 (22.35%)	226 (12.6%)	1411 (69.88%)
	oFASS-3	0.53 (0.49,0.58)	0.54 (0.48,0.6)	0.53 (0.48,0.57)
	oFASS-5	0.59 (0.55,0.63)	0.65 (0.6,0.69)	0.57 (0.54,0.61)
	nFASS	0.62 (0.59,0.65)	0.7 (0.66,0.73)	0.58 (0.56,0.61)
HCSC Cohort N = 129	*N(%)#*	4 (4.87%)	3 (2.3%)	37 (28.68%)
	oFASS-3	0.96 (0.91,1.0)	0.96 (0.89,1.0)	0.63 (0.46,0.8)
	oFASS-5	0.98 (0.93,1.0)	0.98 (0.94,1.0)	0.65 (0.52,0.79)
	nFASS	0.97 (0.94,1.0)	0.98 (0.94,1.0)	0.68 (0.59,0.78)
Europrevall general population N = 635	*N(%)#*	1 (0.15%)	0	15 (2.37%)
	oFASS-3	0.89 (0.66,1.0)	Not estimable	0.74 (0.61,0.87)
	oFASS-5	0.89 (0.68,1.0)		0.79 (0.65,0.93)
	nFASS	0.94 (0.91,0.96)		0.77 (0.64,0.89)
SAFE cohort N = 385	*N(%)#*	Information not collected		19 (5.02%)
	oFASS-3			0.81 (0.69,0.92)
	oFASS-5			0.82 (0.71,0.94)
	nFASS			0.84 (0.74,0.95)
iFAAM cohort	*N(%)#*	52 (11.76%)	49 (11.08%)	35 (7.91%)

N = 453	*oFASS-3*	0.66 (0.60,0.72)	0.75 (0.69,0.80)	0.67 (0.60,0.74)
	oFASS-5	0.70 (0.64,0.77)	0.76 (0.70,0.82)	0.72 (0.64,0.80)
	nFASS	0.72 (0.65,0.78)	0.78 (0.72,0.84)	0.74 (0.66,0.82)

#, percent of food reactions excluding those in which the severity indicator is missing (missing values shown in Table S7)

ROC-AUC, receiver operating characteristic area under the curve; 95% CI, 95% confidence interval

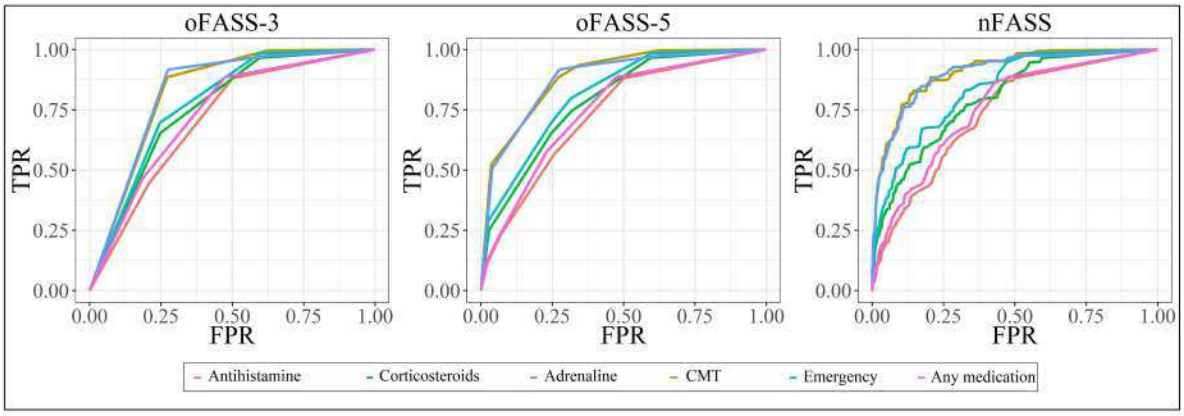
Figure 1:

A. Model evaluation: ROC curves for oFASS-3, oFASS-5 and nFASS in the EuroPrevall outpatient clinic cohort

B: External validation: ROC curves for the severity indicators with nFASS

TPR, true positive rate; FPR, false positive rate; ROC, receiver operating characteristic, CMT, critical medical treatment.

A.



B.